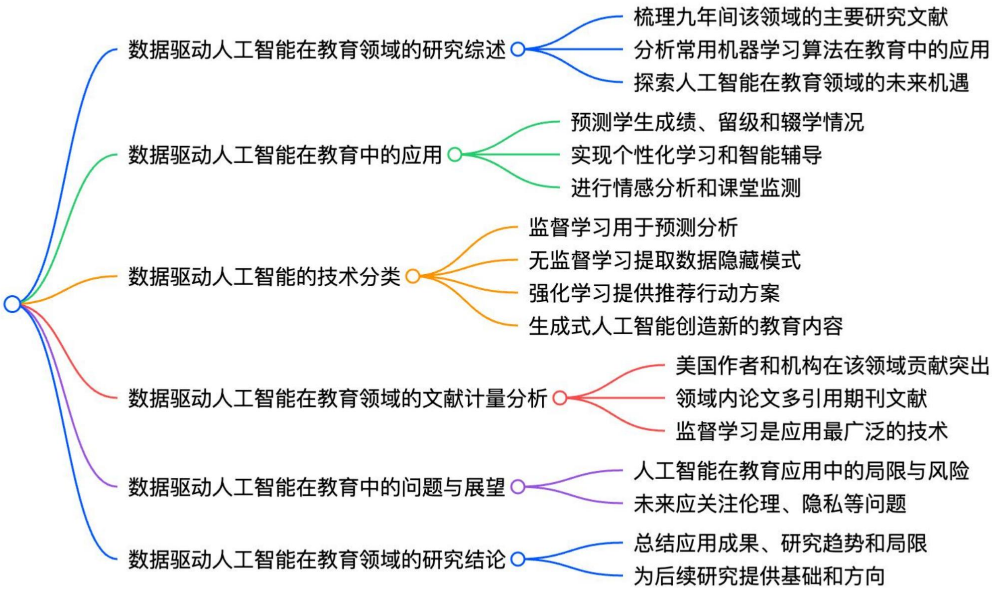
数据驱动人工智能在教育领域的研究综述
梳理九年间该领域的主要研究文献
分析常用机器学习算法在教育中的应用
探索人工智能在教育领域的未来机遇
数据驱动人工智能在教育中的应用
预测学生成绩、留级和辍学情况
实现个性化学习和智能辅导
进行情感分析和课堂监测
数据驱动人工智能的技术分类
监督学习用于预测分析
无监督学习提取数据隐藏模式
强化学习提供推荐行动方案
生成式人工智能创造新的教育内容
数据驱动人工智能在教育领域的文献计量分析
美国作者和机构在该领域贡献突出
领域内论文多引用期刊文献
监督学习是应用最广泛的技术
数据驱动人工智能在教育中的问题与展望
人工智能在教育应用中的局限与风险
未来应关注伦理、隐私等问题
数据驱动人工智能在教育领域的研究结论
总结应用成果、研究趋势和局限
为后续研究提供基础和方向

Data-Driven Artificial Intelligence in Education: A Comprehensive Review

Kashif Ahmad, Waleed Iqbal, Ammar El-Hassan, Junaid Qadir, *Senior Member, IEEE*, Driss Benhaddou, Moussa Ayyash, *Senior Member, IEEE*, and Ala Al-Fuqaha, *Senior Member, IEEE*

***Abstract*—As *education* constitutes an essential development standard for individuals and societies, researchers have been exploring the use of artificial intelligence (AI) in this domain and have embedded the technology within it through a myriad of applications. In order to provide a detailed overview of the efforts, this article pays particular attention to these developments by highlighting key application areas of data-driven AI in education; it also analyzes existing tools, research trends, as well as limitations of the role data-driven AI can play in education. In particular, this article reviews various applications of AI in education including student grading and assessments, student retention and drop-out predictions, sentiment analysis, intelligent tutoring, classroom monitoring, and recommender systems. This article also provides a detailed bibliometric analysis to highlight the salient research trends in AI in education over nine years (2014–2022) and further provides a detailed description of the tools and platforms developed as the outcome of research and development efforts in AI in education. For the bibliometric analysis, articles from several top venues are analyzed to explore research trends in the domain. The analysis shows sufficient contribution in the domain from different parts of the world with a clear lead for the United States. Moreover, students' grading and evaluation have been observed as the most widely explored application. Despite the significant success, we observed several aspects of education where AI alone has not contributed much. We believe such detailed analysis is expected to provide a baseline for future research in the domain.**

***Index Terms*—Artificial intelligence (AI) in education, e-learning, educational data mining (EDM), generative AI for education, intelligent tutoring systems (ITS), machine learning (ML) in education, personalized learning.**

Manuscript received 16 December 2022; revised 21 July 2023, 29 August 2023, and 7 September 2023; accepted 8 September 2023. Date of publication 12 September 2023; date of current version 4 January 2024. *(Corresponding author: Ala Al-Fuqaha.)*

Kashif Ahmad is with the Department of Computer Science, Munster Technological University, T12 P928 Cork, Ireland (e-mail: kashif.ahmad@mtu.ie).

Waleed Iqbal is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS London, U.K. (e-mail: w.iqbal@qmul.ac.uk).

Ammar El-Hassan is with the Department of Computer Science, Princess Sumaya University for Technology Amman, Amman 11941, Jordan (e-mail: a.elhassan@psut.edu.jo).

Junaid Qadir is with the Department of Computer Science and Engineering, College of Engineering, Qatar University, Doha 2713, Qatar (e-mail: jqadir@qu.edu.qa).

Driss Benhaddou is with the Department of Science Engineering Technology, College of Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Electrical Engineering, College of Engineering, Alfaisal University, Riyadh 11533, KSA (e-mail: dbenhadd@central.uh.edu).

Moussa Ayyash is with the Computing, Information, Mathematical Sciences and Technology, Chicago State University, Chicago, IL 60628 USA (e-mail: mayyash@csu.edu).

Ala Al-Fuqaha is with the Information and Computing Technology (ICT) Division, College of Science and Engineering (CSE), Hamad Bin Khalifa University, Doha 5825, Qatar (e-mail: aalfuqaha@hbku.edu.qa).

Digital Object Identifier 10.1109/TLT.2023.3314610

I. Introduction

In the modern world, artificial intelligence (AI) is revolutionizing the way humans live their lives. Similar to other domains, the field of education is also going through a paradigm shift through the use of AI, which can be used to unleash insights about understanding how students learn, how to personalize the learning experience of students, how to get more information to help in the decision-making process, and how to model the complex interaction between student learning, the knowledge domain, and the tools that enable students to interact with the domain? AI can be useful in addressing education-related challenges that are rooted in both the inadequacy of the traditional way of teaching the current generation and the complexity of the educational system itself. AI has a very rich history in education and different AI algorithms have been widely employed in education for different applications since the 1970s [1]. Over the past decade, the role of AI in learning has been on the radar of educational institutions, government agencies, funding agencies, and industry [2].

We use the term AI broadly as an umbrella term that subsumes methods, algorithms, and systems that learn from data [data science, statistical learning, machine learning (ML), and deep learning] or aim to create machine intelligence that can perform tasks, such as perception, reasoning, inference (such as expert systems, probabilistic graphical models, and Bayesian networks). These terms are largely used in the current convention synonymously [3], and our use of the term AI will ease exposition and reduce clutter. We make the distinction between AI and other subsumed techniques where it is important. It is worth noting though that there are various types of AI techniques, and not all of them are connectionist [i.e., based on neural networks (NN) and deep learning] [4].

The AI techniques in education can be broadly divided into two different categories, namely: 1) representational/knowledge-based AI and 2) data-driven AI [5]. The knowledge-based AI algorithms aim to employ human experts' knowledge in decision-making, such as rule-based systems. The majority of the previous efforts were based on knowledge-based AI [6]. However, recently the trend shifted toward data-driven techniques making use of data in making decisions. In this article, we focus predominantly on data-driven AI techniques in education and review the recent efforts made in the domain with a particular focus on applications and tools.

There are three main roles for AI in education including assisting teachers to deal with: 1) individual students, 2) the

1939-1382 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

whole class, and 3) the whole cohorts of students [7]. At the individual level, the focus is more on adapting teaching methods and approaches to a particular learner's needs. On the other hand, at the class level, AI aims to help teachers manage a whole class instead of individual learners [8]. Some key applications of AI in the classroom include tutoring, grading, and virtual reality (VR)-based learning to improve the teaching and learning experience in a classroom via an effective teacher and AI collaboration [9]. At the cohort level, AI aims to analyze learners' interaction with the systems and tune the learning system based on the failure and success of learners' interaction with the system. Some key applications at the cohort level include the identification of learners at risk, learners' interests, behavior, performance, and dropout prediction.

Different research communities have taken different approaches to the use of data-driven methods for addressing educational problems at different levels. For instance, the data mining research community addresses educational research problems using a Big Data approach while AI communities address research problems focusing on algorithms and methodologies as part of their efforts toward the development of interactive and adaptive learning environments. Although these fields are overlapping, these communities tend to develop distinct research areas as they have had different research histories. The knowledge discovery and data mining (KDD) research community aims to discover patterns and extract knowledge through data mining techniques. The educational data mining (EDM) community attracts interdisciplinary scientists from computer science, education, psychometrics, and other fields to analyze data acquired from the educational environment and apply data mining techniques to solve educational challenges [10]. On the other hand, the Society for Learning Analytics Research (SoLAR) community is an "interdisciplinary network of leading international researchers who are exploring the role and impact of analytics on teaching, learning, training, and development" [11]. Similarly, the International Artificial Intelligence in Education Society (IAIED) [12] is an interdisciplinary community aiming to bring researchers from different domains, such as *computer science*, education, and psychology for the promotion of interactive and adaptive learning environments. It is to be noted that AI in education is not limited to EDM, learning analytics, and ML. In fact, many other research activities are being developed by different research groups around the world to explore how AI can be utilized to solve educational problems.

A. Scope of the Survey

This article revolves around the key applications of data-driven AI techniques in education and describes the most commonly used data-driven AI techniques in different application areas of education. It also describes the tools and platforms developed in the market as outcomes of the research work achieved in different applications including: 1) student grading and evaluations and retention and drops out prediction; 2) personalized learning; 3) sentiment analysis; 4) recommendation systems in education; and 5) classrooms' monitoring and visual analysis. We also analyze research trends in AI applications in education by providing a detailed bibliometric analysis of the domain. This article also advises on the current limitations, pitfalls, and future directions of research in the domain, and how it can fill the current gaps.

B. Related Surveys

Literature reports several interesting articles analyzing different aspects of AI applications in education. There are also some surveys targeting different aspects and application areas of AI in education [13]. However, to the best of our knowledge, there is no recent detailed survey covering the domain from different perspectives. In previous works, Romero et al. [10] provided a general review of existing literature to analyze how EDM and learning analytics (LA) have been applied to educational data. Romero et al. [14] surveyed EDM literature for a decade (i.e., 1995 to 2005). Baker et al. [15] reported a detailed survey of data mining techniques used in the education sector. Charitopoulos et al. [16] provided a detailed overview of literature from 2010 to 2018 aiming at the use of soft computing methods, such as decision trees, random forests (RFs), artificial neural networks, fuzzy logic, support vector machine (SVMs), and genetic/evolutionary algorithms, in EDM and LA.

More recently, Fischer et al. [17] surveyed the existing data mining techniques in education with a particular focus on highlighting challenges in mining Big Data. Mduma et al. [18] focused on students' retention and dropout prediction techniques. Almasri et al. [19] provide a detailed survey of intelligent tutoring systems (ITS), another attractive application of AI in education, proposed from 2000 to 2018. Mousavinsab et al. [20] also provided a systematic review of literature on ITS with a particular focus on its key characteristics, applications, and evaluation methods. In another interesting survey on ITS, Malekzadeh et al. [21] provided a detailed overview of literature aiming at the development of different emotion regulation strategies to better engage learners in ITS. Tahiru [22] provided a systematic review of AI in certain applications of AI in education, such as ITS, automation of administrative tasks, and development of smart content. Al-Emran et al. [23] surveyed the Internet of Things (IoT)-based educational solutions. Matcha et al. [24] provided a systematic review of empirical studies on LA from a self-regulated learning perspective. Hooshyar et al. [25] also provided a systematic review of literature on self-regulated learning, however, they focused on the role of open learner models instead. Zaidi et al. [26], on the other hand, provided a detailed overview of online learning and the AI education market focusing on the scope, processes, and providers/competitors in the market. The authors also discussed the role and future potential of AI in online education. In another interesting relevant survey, Kim et al. [27] provided a detailed overview of pedagogical agents by analyzing their roles, current progress, and future potential in research-based pedagogical agent design.

Literature also provides several interesting surveys on recommender systems for education [28]. For instance, Khanal et al. [29] provided a systematic review of literature on ML-based recommendation systems in e-learning. Similarly, Rahayu et al. [30] provided a systematic review of literature on learning

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISON OF OUR ARTICLE AGAINST EXISTING SURVEYS

Survey	Application Domains	Tools & Platforms	Bibliometric Analysis	Generative AI	Pitfalls of AI
[31]	*Covered 2 applications*. (i) Students' Performance Prediction and (ii) Recommendation Systems	✓	Yes (2005-2010)	X	X
[15]	*Not Covered*.	X	Yes (2000-2008)	X	X
[17]	*Covered 4 applications*. (i) Students' Grading and Evaluation, (ii) Students' Retention and Dropout, (iii) Sentiment Analysis in Education and (iv) Recommendation Systems in Education	X	X	X	X
[18]	*Covered 1 application*. (i) Students' Retention and Dropout	X	X	X	X
[19]	*Covered 2 applications*. (i) Students' Performance Prediction and (ii) Intelligent Tutoring Systems	✓	X	X	X
[32]	*Covered 10 applications*. (i) Profiling and prediction, (ii) Admission decisions and course scheduling, (iii) Drop-out and retention, (iv) Student models and academic achievement, (v) ITS, (vi) Curating learning materials based on student needs, (vii) Facilitating collaboration between learners, (Viii) Assessment and evaluation, (ix) Automated grading, and (x) Evaluation of teaching	X	X	X	X
[33]	Mainly focused on the role of AI in language education, and covered relevant applications, such as writing, reading, and vocabulary acquisition	X	Yes (1990-2020)	X	X
[34]	Mainly based on EDM and covers four applications including analyzing students' motivation, attitude, behavior, learning style, and collaborative learning in e-learning environment	X	X	X	✓
[35]	Mainly focused on emotion recognition and its role in technology-based learning	X	X	X	X
This Work	*Covered 9 applications*. (i) Students' Grading and Evaluation, (ii) Students' Retention and Dropout, (iii) Personalized Learning, (iv) Students' Performance Prediction, (v) Sentiment Analysis in Education, (vi) Recommendation Systems in Education, (vii) Classroom Monitoring & Visual Analysis and (viii) Intelligent Tutoring Systems.	✓	Yes (2014-2022)	✓	✓

path recommender systems by covering key topics, such as trends, ontology, recommendation processes and techniques, contributing factors, and evaluations.

In contrast to the existing surveys, this article provides a broader picture of the domain by covering most of the key application areas of AI in education, such as student's grading and evaluations, students' retention and dropout prediction, students' performance, sentiment analysis, recommendation systems, classrooms' monitoring, and ITS. The article also highlights the key market players, tools, and platforms along with key challenges, potential market opportunities, future research directions, and pitfalls of AI in education. More importantly, we analyze the trends of the research in the domain from a different perspective by providing a detailed bibliometric analysis of the domain. The article also discusses the different types of AI algorithms including supervised, unsupervised, reinforcement learning (RL), and generative AI algorithms. Table I provides a comparison of this survey against the existing surveys on the topic.

C. Contributions

1) A detailed overview of existing literature in nine different application domains in which AI is deployed for education.
2) Describe and highlight recent works on the most commonly used ML algorithms adopted in literature over the years for these applications.
3) Explore and identify the future scope and market opportunities for AI researchers and developers in the education sector.
4) Analyze the publication trends of research literature taking into account a total of 4447 articles published in various top subject venues through a detailed bibliometric analysis in terms of research productivity by authors, institutions, and country, and knowledge flow across various research venues.
5) Identify the research and development companies and corporations working in the domain along with the tools and platforms available for both educational institutions and researchers.
6) Highlight the limitations, pitfalls, and open research challenges in using data-driven AI in education.

The rest of this article is organized as follows. Section II describes some key application domains of AI in education and different tools developed as part of the efforts. Section III details some key AI techniques employed in different applications of education. Section IV provides a comprehensive bibliometric analysis of existing literature. Section V provides basic insights into the domain based on our analysis of existing literature, and lists the key limitations and pitfalls of AI in education along with some potential directions of future research and open issues in the domain. Finally, Section VI concludes this article.

II. AI Applications in Education

As part of the efforts in the domain, several interesting AI-based tools and applications have been introduced to facilitate educators in different ways. Moreover, several international standards, such as Caliper, xAPI, and next generation digital learning environment (NGDLE), are proposed to provide guidelines for data collection, storage, and analytics for these applications [36]. In the next subsections, we provide an overview of existing literature on these applications as well as the tools and platforms developed as part of these efforts.

A. Students' Evaluation and Performance, Retention, and Dropout Prediction

To be able to predict a student's likely future performance, retention, and dropout can provide very powerful platforms that facilitate educational interventions and remedial actions promptly. The development of AI models for the prediction of student performance and uncovering hidden insights and patterns are some of the most salient applications and research areas in EDM and learning analytics. Several studies have been conducted in the area of academic performance analysis and prediction, including by Adejo et al. [37], who conducted an empirical investigation and comparison of several data sources, classifiers, and ensembles of classification techniques to predict the academic performance of university students. In detail, they compared and analyzed the performance of ensemble techniques combining information from different data sources against the

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

models trained on data from a single source. To this aim, several algorithms including DT, artificial neural networks (ANNs), and SVM were used and compared individually as well as in ensemble (combination) modes. Their findings support the premise that multiple data sources in combination with heterogeneous ensemble ML techniques provide efficient models for predicting student performance and also for identifying students at risk of attrition. Livieris et al. [38] also proposed an ensemble-based semisupervised approach for predicting student performance achieving sufficient accuracy in early prediction of student progress.

Deep learning techniques were also employed to tackle the challenging problem of forecasting the future performance of students. For instance, Kim et al. [39], proposed GritNet, a novel deep learning model, for the prediction of students' performance by treating it as a sequential prediction task. GritNet is mainly based on bidirectional long short-term memory (BLSTM). The authors applied the model to a group of Udacity students to predict their performance and were able to show favorable results over the logistic regression models with on-the-ground improvements in the early weeks of the course which are traditionally the most challenging to predict.

Student retention and dropout, which is a universal factor affecting both online and offline learning platforms, could be linked to student performance prediction. AI techniques have also been proven very effective in student retention and dropout prediction. As an initial effort, Aulck et al. [40] modeled student dropout based on a dataset of 32 500 demographics and transcript records at a large public institution. They conclude that early potential dropouts can be detected even with single-term transcripts thus opening the door for AI applications to predict and prevent some of the causes of dropouts.

In literature, AI modeling techniques have been applied to predict dropout rates to calculate dropout probability as well as identify the contextual, demographic, and individual factors related to learning activities such that education administrators can design effective intervention and prevention remedies. For example, Solis et al. [41] considered several factors including the year of graduation, residence while attending the lectures, gender, grants/financial aid, program, program choice, and previous academic record. The authors also evaluated the performance of various ML algorithms for the prediction of student retention rates at university levels; using comparative experiments and analysis between the effectiveness of NNs, random forest, logistic regression, and SVM, the authors found that the combination of RF using ten randomly selected candidates per division was the optimum combination for predicting student dropouts. On validation of this work, their model was able to predict dropouts with over 90% accuracy combined with a sensitivity rate of 87%. The authors concluded that for better prediction complete information from all semesters must be used in training the algorithms. The authors also recommended the inclusion of other factors, such as the students' interest in the program, their level of interest in the university, earnings/work while studying, as well as the educational level of their relatives, family support, and their attitudes and personalities.

From a different perspective, Pilkington et al. [42] conducted a qualitative study as part of funded research at a U.K. university with a sample of 75 researchers, tutors, and professors. They used a combination of "systematic, sequential, explanatory, and thematic" approaches to focus on findings from thematic analysis. They identified engagement, attendance, workload, family pressure, and mental health as factors that continue to contribute to dropout issues regardless of university engagement efforts. Apart from these factors, the sense of community, institutional social-environmental contribution, and academic integration are other critical factors contributing to students' retention and dropout [43]. It is therefore essential to go beyond basic AI modeling for predicting dropout and analyze the impact of ambient socioeconomic, psychological, demographic, and family factors to be able to conduct a determined analysis of the causes of dropout. For example, generally, the students do not have a fixed study approach in massive open online courses (MOOCs). Thus, tackling the temporal and diverse nature of MOOC data and considering the inconsistent learning/study activities of students, such as watching and rewatching a video, in dropout prediction is a very difficult task as these reasons are diverse and highly personalized. Chen et al. [44] applied visualization analytics methods and techniques (DropoutSeer) to analyze large datasets from MOOC systems to correlate ML predicted dropout rates with the learning activities of MOOC subscribers visually. The aim was to enable content designers to design more suitable engaging content and AI experts to design better predictive models [45]. This was shown to be more effective, for instance, than the process of feature identification as a critical step in the model-building process [46].

Literature shows that most of the initial efforts in the domain are based on classical ML algorithms, such as Bayesian classifiers, RFs, ANNs, and SVMs. However, similar to the other applications, recently the trend shifted toward more advanced techniques, such as DL and genetic algorithms-based techniques [47], [48]. Moreover, more recently, some explainable AI-based solutions have also been proposed for the tasks. For instance, Hasib et al. [49] proposed a local interpretable model-agnostic explanations (LIME)-based model for the interpretation of results obtained with a diversified set of AI algorithms deployed for the performance prediction of secondary school students. Moreover, the key features/factors considered in students' performance prediction in literature include students' academic record (e.g., previous grades), demographic features (e.g., age, gender, income, race), school, instructor, and parents-related features, and students' behavioral-related features [50]. A summary of some of the existing AI-based tools and platforms for grading, students' retention, drop out, and performance prediction has been provided in Table II.

B. *Personalized Learning*

Personalized learning has been subject to many simultaneous and fundamental transformations mainly due to growing students' needs, globalization, new challenges in education

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

TABLE II
SUMMARY OF TOOLS/PLATFORMS FOR GRADING, STUDENTS' RETENTION, DROPOUT, AND PERFORMANCE PREDICTION

Tool/platform	Open source	Learning Settings	Key Features
CampusLabs [51]	X	K12 and Higher Education	• Integrates data from different sources to cultivate campus intelligence and make better decisions. • Identifies students at risk • Special supports for targeted students • Strengthen educators' ability to guide students on their pathways to success.
RNL Student Retention Predictor [52]	X	Higher Education	• Poses accurate assessment capabilities and extracts actionable insights from data obtained from different sources • Its predictive analytics identify students who need special attention and help • Helps in developing strategies that increase efficiency and the impact of retention efforts • Helps administration to recruit and retain the right students
Nuro Retention [53]	X	All Levels	• Efficient predictive analytics helps educators in engaging each student • Identifies students at risk and also provides insights on the reasons for it • Assists educators in developing effective strategies for improved student graduation and retention • Can be customized to an institution's needs
Othot Retention Predictor [54]	X	No specific information	• Real-time and dynamic predictions • Identifies the individuals who need more attention • Recommends actions and devises strategies that will have the greatest impact on an individual's performance • Affordable tool showing educators where to focus resources
CampusNexus Succeed [55]	X	Higher Education	• Tracks each student's engagement and progress • Identifies and prioritizes students based on the risk level • Flags and allows a teacher to respond to alerts from other teachers
WriteToLearn [56]	X	K-12	• Automated assessments, scoring system, and reporting to teachers • An immediate feedback to students to better practice • Focuses on summary and essay writing • Teacher as well as student reporting capabilities
Quantum Adaptive Learning and Assessment [57]	X	All Levels	• Provides a question answer facilities where students can put their inquiry • Acts as a cognitive coach observing the thinking and questioning expertise of the students
Azure Cloud AI Tools for Education [58]	✓	All Levels	• Facilitates all stakeholders in the education sector including students, teachers, and administration • Generate insightful student performance analytics showcased through Microsoft PowerBI dashboards • Coding courses and tutorials • Helps students to pursue careers in technology or other fields
Lightside [59]	X	K-12	• Evaluates of students' writing • Provides feedback on the use of language, the focus of the document, organization, and evidence • Specially customized for students in grades six through 12
Proctorio [60]	X	K-12	• Fully automated exam proctoring without scheduling 24 hours a day, 7 days a week • Supports automatic ID verification • Provides admin dashboard and aggregates exam data • Ensures content protection with copy/print/download restrictions
Gradescope [61]	✓	All Levels	• Supports grading of paper-based, digital, and code assignments • Also provides insights into students' performances • Covers multiple subjects

management, and technological developments [62]. Technology implementation and inclusion notwithstanding, the traditional learning approach with a static and unidirectional model including the teacher in front of students, reading text material, and written exam-based assessments that cover all sections of the classroom uniformly is being eroded. Contemporary learning directions converge to interactive, student-focused, tailored learning models that serve each student or student group much closer with better engagement, closer interaction, improved comprehension, and wider scope coverage of learning outcomes.

The adaptability of the learning model to cater to multiple learning settings would not have been feasible just a few years ago before the wide availability and accessibility of technology, such as AI, which leverages the power of cognitive computing, which makes use of reasoning, language processing, ML, and human capabilities/input allowing better solutions and data analysis, in the support of education [63]. Shawky and Badawi [64] explored RL, which represents a branch of AI algorithms, as a cognitive computing catalyst to provide adaptive learning materials and paths in support of bespoke, learner-centered requirements. To this aim, a myriad of personal, social, and environmental factors determining and affecting learners' experience have been monitored and analyzed in different learning settings to update and suggest new learning paths. The objective is to adapt to the "most influential" of these learning factors per learner needs and learning settings. The designed smart learning platform also recommends appropriate learning material in a connected, continuous way that adapts to the changing needs of the learners. The personalized learning approach is composed of three steps. First, a learning path/setting is suggested for learners based on their state. Subsequently, the learner's state and the reward received by the suggested learning path/settings are updated using RL, which ultimately leads to effective personalized learning settings based on the learner's needs. The approach is

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

TABLE III
SUMMARY OF AI-BASED TOOLS/PLATFORMS FOR PERSONALIZED LEARNING

Tool/platform	Open Source	Learning Settings	Domain	Key Features
Third Space Learning [65]	X	K-12	Math	• Special attention to target students based on their weaknesses • Provides weekly online lessons • Personalized learning by adapting the tutor to a student's needs with weekly reporting • Provides access to premium maths resources
EnLearn [66]	X	All Levels	Generic	• Personalized content via an adaptive learning ecosystem involving students, teachers, and curriculum • Can increase content for target students • Able to identify misconceptions and remedies in the learning process
Watson Content Analytic [67]	X	All Levels	Generic	• Vocabulary learning applications • Helps teachers to track students' progress • Conducts real-time assessments and provides insights to instructors • Addresses the high-tech skills gap
Querium [68]	X	MOOCs	Generic	• Uses AI to help students with STEM skills so they can be ready for further studies • A personalized program is called StepWise and it works on smartphones and computers • Delivers personalized, bite-sized lessons and step-by-step tutoring assistance
Edly [69]	X	K-12 and Higher Education	Generic	• Supports students of all ages including K-12 and Higher education • Provides training management of different stakeholders of the education
Squirrel [70]	✓	K-12	Math, English, Physics, Chemistry	• Squirrel AI Learning offers high-quality after-school courses in subjects such as Chinese, Math, English, Physics, and Chemistry. • Provides students with a supervised adaptive learning experience
MobyMax [71]	X	K-8	all K-8 subjects	• Uses AI to pinpoint and fix learning gaps with adaptive, differentiated learning materials for all K-8 subjects. • Students can learn at their own pace with lesson plans and practice with automatically generated sheets.

evaluated in an extensive simulation setup using 100 states and actions.

AI can be used for a diversified set of tasks in personalized learning, however, in this section, we will mainly focus on ITS and recommender systems. In the next subsections, we provide an overview of these applications of AI in personalized learning.

1) Student Modeling in Intelligent Tutoring Systems (ITS): ITS systems can be differentiated from personalized learning platforms as they represent a specialized concept/component in personalized learning, and there is dedicated literature on it. ITS is a very well-developed form of the personalized learning platform and therefore we describe it in a distinct subsection here. ITS, which aims to provide immediate and customized feedback to learners, plays a major role in overcoming the growing gap between the increasing number of learners and the shortages in qualified specialist teachers globally. Many ITS systems are in active use supporting and enhancing traditional school curricula in thousands of schools in the US and beyond [72]. ITS is also very effective in predicting student cognitive needs, results, mental states, and skills and subsequently recommending the right course of action. For example, ITS is applicable in modeling student emotions [73], efficacy [74], ability to perform scientific inquiry within a virtual environment [75] and then generate recommendations automatically [76]. Although in many of the ITS techniques studied thus far, the mapping between the actions and decisions of intelligent pedagogical agent (IPA) systems on the one hand and students and teachers on the other are yet to be refined. Nonetheless, the role of AI techniques and models' interpretability is even more essential within these contexts of modern learning as it enables an IPA to justify actions and inferences. This, in turn, improves IPAS' effectiveness (providing a "why" analysis rather than merely "what"). Furthermore, this fosters user trust and confidence in the correctness and integrity of the learning system [77].

Student models play many roles within ITS including for assessment of student performance in core or soft skills or for monitoring student compliance with curriculum of school constraints during their path/plan of study. This variety of purposes shapes the fundamental design of the student model's architecture and opens up this field for quite rich research and application design. The authors in [78] analyzed student models in ITS from the perspective of their role in the architecture of an ITS and also for determining the model components that should be considered in its design. Using a conversational ITS (CIRCSIM-Tutor), the authors define the decisions that the system needs to make together with the associated information that supports these decisions. The authors recommend four types of student model blueprints that are based on information aspects and constraints of the tutoring system being analyzed.

In [79], the authors focus on the less explored aspect of ITS, namely, tailored instruction mechanisms also known as tutorial dialogue systems (TDS). TDS engages students in conceptual discourse using natural language processing techniques. Using conceptual physics as an application domain, the authors introduce a TDS that maps tutorial dialogues and student models; their (RIMAC) model dynamically builds a persistent student model that supports proactive as well as reactive decisions in service of adaptive student instruction. In the applied classroom and test pilot studies, the authors demonstrated the effectiveness of their TDS with students taking less time to complete

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
SUMMARY OF AI-BASED TOOLS/PLATFORMS FOR RECOMMENDATION SYSTEMS IN EDUCATION

Platform	Learning Settings	Key Features
Qbot [85]	MOOCs	• Tags tutors and classmates to answer students' questions • Recommends content to individual students based on their current level of knowledge • Provides students' analytics
MyEdMatch [86]	K-12	• Connects schools and teachers having shared beliefs and goals • Helps schools to recruit the best talent • Helps teachers in searching for jobs
TeacherMatch [87]	All Levels	• Uses AI for real-time analysis of the candidates in the pool maintained by the platform • Assess teachers on four factors including teachers' qualifications, teaching skills, cognitive abilities, and attitudinal factors

learning tasks than counterparts who did not utilize the system in tutorials. It was also demonstrated that "both high and low prior knowledge students learned more efficiently from a version of the tutor that dynamically updates its student model during dialogues than from a control version that included the static (poor man's) student model" [79].

2) Recommendation Systems in Education: Recommendation systems have proven highly effective in various domains, such as business, food, tourism, and entertainment [80]. Similarly, these systems have found wide application in the education sector. In education, recommendation systems serve multiple purposes, aiding various stakeholders. For instance, they recommend relevant learning materials to students. In addition, they assist teachers in professional development by identifying useful documents and resources based on input from other teachers [81]. Moreover, recommendation systems can suggest remedial actions to enhance learning quality, supporting the operational side of academic teaching [82]. In this environment, assessment data collected over several academic semesters are analyzed based on learning outcomes at the course and program levels. Historical student attainment shortcomings, typically addressed by domain experts and course coordinators, are utilized to build a pool of remedial actions (recommendations) over a span of 3 to 5 years [83]. Training data/features, such as course domain, course level, section size, and lab options, guide experts in selecting recommended remedial actions for subsequent assessments, whether formative or summative. A multilabel classification algorithm is employed to choose suitable actions for each rubric line, representing performance per group of students, from the master pool. AI provides notable strengths in this domain, offering efficiency, consistency, and fairness in the application of remedial actions. While this setup is particularly suitable for large colleges with extensive student populations and archives of structured, outcome-based assessment data spanning several years, the approach has demonstrated success and reasonable accuracy even with a smaller number of learning instances [83].

Table IV summarizes some existing recommendation systems in education for both students and teachers. Most of the existing recommendation systems aim to select/recommend relevant educational resources, learning activities, and finding peers [80]. Similar to other application domains, the majority of educational recommendation systems are based on content-based, collaborative filtering, or hybrid recommendation methods. Moreover, a diversified set of metrics is used for the evaluation of recommendation systems in the domain, such as students' grades and learning outcomes [84].

C. Sentiment Analysis in Education

Sentiment analysis involves analyzing and extracting people's opinions about a service or an entity. Often referred to as opinion mining, it is a challenging task that generally involves different phases, such as collection and storage of data as well as analysis of the data using a combination of knowledge-based and ML techniques [88]. In the context of education, sentiment analysis attempts to improve the learning process by analyzing students' feedback to better understand their opinions, emotions, and concerns and make adjustments to the content or delivery of the learning material accordingly [89]. Sentiment analysis is generally associated with the extraction of people's emotions, which in the learning context could be very beneficial as emotions may affect students' motivation and learning outcomes [90]. Thus, early identification and proper handling of students' emotions and concerns may result in a better learning experience/process.

Sentiment analysis also plays a significant role in extracting students' opinions on learning materials from social media. According to Chauhan et al. [91], students' views are continuously exchanged using different social media platforms in response to their experiences in the classroom or online learning; these provide a rich pool of data for evaluating students learning. From an MOOCs perspective, Kastrati et al. [92] demonstrated the effectiveness of sentiment analysis in automatically analyzing students' feedback. The authors also highlight the labor involved in manually assessing and annotating students' feedback. To overcome this limitation, they proposed a weakly supervised framework for aspect-level sentiment analysis specifically aiming to highlight polarity in student feedback about MOOCs. Liu et al. [93] argued that the continuous feedback of students in MOOC environments stipulates that students' emotions and learning activities be tracked for understanding learning requirements. To analyze students' emotions (i.e., *positivity*, *negativity*, and *confusion*) and concerned aspects (e.g., teaching styles and learning activities), the authors proposed the temporal emotion-aspect model (TEAM) which tracks students' emotions toward the concerned aspects and characterizes their joint distribution over time. In other words, the model associates emotions with different aspects and their evolution over a period of time. The results indicated that: 1) content-related aspects were the main emphasis with a higher likelihood of confused or negative emotions; 2) there were higher likelihoods of emotional expressions at the start and end of a semester; and 3) underachieving students were less active in emotional engagement and tended to express more confusion toward the end of a semester when compared

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

to high-achieving and medium-achieving students. Such observations could be useful to teachers for timely instructional guidance or psychological intervention that could help the learners in the future. Applications of this technology in education and other fields are already established with favorable results in education, healthcare, social media, and natural language processing domains [89].

The other key role of opinion mining/sentiment analysis in education research includes the investigation of learners' satisfaction with the available resources, their attitude toward learning, their concerns, and evaluation of teachers' performance, and teaching methods [94]. For instance, Munezero et al. [95] analyzed students' learning diaries to predict students' sentiments, emotions, and opinions about their learning experience. According to Kechaou et al. [96], knowledge and evaluation of user opinions is an essential prerequisite for the effective development of e-learning systems. To this end, an opinion mining method has been applied in their research to support e-Learning content developers to enhance the quality of provided services using three feature selection methods, namely, mutual information (MI), information gain (IG), and CHI statistics (CHI) in conjunction with hidden Markov models (HMM) and SVM-based hybrid learning methods. Experimental results indicate that opinion mining is more challenging in e-learning blogs in the presence of noise. Although this work has shown that IG constitutes the optimal potential for sentimental term selection and produced optimum accuracy in sentiment classification. More recently, Mostafa et al. [97] reviewed work in sentiment analysis related to gamification in learning, the author proposed a classifier that will analyze the sentiments of students while using gamification tools for learning in Egypt.

The majority of the recent efforts in the domain are based in higher education, where several important topics, such as the collection of relevant and meaningful data for sentiment analysis, designing sentiment analysis methodologies, and exploring the relationship between sentiment and behavioral analysis with the learners' performance and achievement, are explored [94]. The majority of the state-of-the-art solutions are based on textual information, where mostly DL techniques, such as LSTM, biLSTM, and transformers, such as bidirectional encoder representations from transformers (BERT) and robustly optimized BERT (RoBERTa) are used. However, literature also hints at the use of visual sentiment analysis techniques for analyzing and extracting students' emotions and expressions [98].

The research efforts in the domain resulted in several interesting AI-based sentiment analysis tools. Some of the existing sentiment analysis tools developed or customized for education are provided in Table V.

TABLE V
SUMMARY OF TOOLS/PLATFORMS FOR SENTIMENT ANALYSIS IN EDUCATION

Tool	Learning Settings	Key Features
Pepper [99]	MOOCs	• Can answer visitors'/customers' queries • Helps administrative staff to perform routine tasks • Engages visitors in effective conversations and provides personalized responses
NAO [100]	Preschool through grad school	• Informs and entertains visitors • Provides an optimized teaching-aid tool • Effective tool for special education ("students with disabilities such as autism, emotional and behavioral disorders" [101])
ZimGo Emotional Intelligence [102]	Early Childhood	• Recognizes and differentiates human emotions from text • Employs state-of-the-art AI and NLP techniques • Poses Contextual Analysis capabilities (i.e., considers the context of the text in emotion extraction) • Can be customized for any application
Talkwalker [103]	MOOCs	• An effective social listening and analytics tool • Helps educators to promote their college and university • Poses user-generated content detection capabilities • An influencer identification tool

D. Classroom Monitoring and Visual Analysis

Classroom monitoring and visual analysis play a key role in classroom management. It can be used for a diversified list of tasks to support teachers in different ways [104]. For instance, classroom monitoring and visual analysis of the classrooms allow for analyzing students' reactions to course contents and other learning materials and activities. It also allows for investigating the students' behavior in the classroom in terms of the time they keep and lose focus during lectures. In addition, it could also be used to analyze teachers' spatial behavior in classrooms, which has a significant impact on students' motivation and engagement in learning [105]. However, monitoring and analyzing individuals in a classroom for a long time is not an easy task. Thanks to advanced video analytics and AI algorithms, it is possible to automatically analyze students' responses, reactions, and levels of attentiveness during lectures. For instance, Narendra et al. [106] proposed a video analytic framework to monitor and investigate how long and when students keep and lose focus during lectures. Similarly, Martinez et al. [107] proposed an interactive teacher's dashboard enabling teachers to keep an eye on groups' learning activities and collaborations in a multitabletop learning environment. This real-time monitoring and analysis could overcome the limitations of conventional supervision, where the teachers can only see the final product of the students/group.

Classroom utilization and occupancy calculations, which are part of budgetary planning and strategic planning of higher education institutions, especially where real estate is a premium asset [108], is another key application of classroom monitoring and visual analysis. Students in modern offline or online degrees have many technology-driven advantages at their fingertips but equally, suffer excessive demands which often cause dropouts and classroom underutilization. Although predicting room occupancy/utilization is an age-old problem [109], the use of modern AI technology as an instrument in measuring or increasing the efficiency of room utilization is a new topic. Sutjaritthamet et al. [110] used on-campus sensor instruments to monitor classroom attendance while respecting student privacy. Several measurement approaches were evaluated in a lab experiment to identify the best sensor technology in terms of cost, accuracy, and convenience.

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

AI also has an impact on technology enhanced learning (TEL) by providing several interesting applications in many subdomains. One such area is to aid the difficult task of understanding the various dimensions of TEL in schools. One reason for this difficulty is the limitation of monitoring classrooms for a longer period to analyze teachers' teaching methods and students' learning experiences. Howard et al. [111] explored the area of observing, analyzing, and visualizing TEL classrooms over time and used sensors to collect observation data over two months. This data are presented as insights to academic administrators and teachers for reflection and corrective action to enhance student learning.

Other applications of EDM and AI for the transformation of the traditional classroom include the analysis of student facial expressions to assess their level of engagement in the classroom. Soloviev et al. [112] proposed a system that analyzes (in real-time) the data feeds from video cameras that are installed in the classroom and applies AI and facial recognition technology to recognize student emotions to determine their level of enjoyment. Although Chua et al. [113] reviewed case studies and technologies developed to collect and analyze educational data. Several aspects of the learning environment, which is a combination of the physical and digital classroom setting, are studied. Moreover, different aspects of the learning process are assessed and analyzed to quantify teaching and learning processes, student assessments are also analyzed automatically. The authors introduce data pipelines that leverage data and information collected from both physical spaces as well as digital spaces.

The majority of the current efforts in classroom monitoring and visual analysis rely on state-of-the-art DL algorithms, such as CNNs, convolutional encoder/decoder networks, and hybrid models by combining CNNs and RNNs [114]. Moreover, the key tasks involved in the process generally include facial landmark extraction, face segmentation, head pose estimation, and facial expression recognition. As part of the efforts, several AI-based tools have been developed that can help in classrooms in several ways, such as security, marking attendance rolls, and emotional and movement monitoring for better classroom dynamics analysis. Table VI summarizes some of the existing classroom monitoring systems.

TABLE VI
SUMMARY OF AI-BASED TOOLS/PLATFORMS FOR CLASSROOMS' MONITORING AND VISUAL ANALYSIS

Tool/platform	Learning Settings	Key Features
Jibble Attendance Platform [115]	All Levels	• Provides an accurate attendance mechanism with biometric verification • Tracks attendance with Phones and Tablets • Prevents cheating with the use of photos, facial recognition, and GPS • Generates automatic attendance sheets and reports with actionable insights
LoopLearn [116]	All Levels	• Provides secure and efficient roll marking facilities by allowing designated school staff only to access the tool • Generates automatic attendance sheets and reports with actionable insights • Can be customized to the needs of other departments of a school, such as sports, peripatetic, and excursions by adding additional features
Secure Accurate Facial Recognition (SAFR) [117]	K-12	• A general-purpose AI-based facial recognition • The tool is customized for K-12 schools with facial recognition of students' parents to allow them to enter the school
AirClass [118]	K-12	• Analyze students' response to a lecture automatically • Detects whether students' eyes are opened or closed during a lecture • Also analyzes students' interests and commitment to learning through facial emotion recognition and analysis.
Captemo: Emotion Recognition [119]	All Levels	• A general-purpose tool that analyzes customers' experience through emotional intelligence • Embedded with state-of-the-art facial and emotion recognition algorithms • Supports both: continuous and on-demand monitoring capabilities
BliPPAR [120]	All Levels	• Improves students' creativity, interactivity, and engagement with any subject with the help of computer vision and augmented reality. • Provides better visualization of complex topics • Helps in creating interactive learning materials

III. TECHNIQUES

Literature on data-driven AI in education based on the nature of the AI algorithms can be roughly divided into four main categories, namely, 1) aupervised ML; 2) unsupervised ML; 3) RL; and 4) generative AI. In the next subsections, we provide a brief description of each of these categories.

A. Supervised Learning

The majority of the works on AI in education rely on supervised learning, as detailed in Section IV. Supervised learning aims at function approximation or curve fitting by finding a relation/function $f : x \to y$ using a training set $\{x, y\}$. Though the efficiency of supervised learning largely depends on the availability and quality of training data, it is a far more accurate learning strategy compared to its counterparts [121]. Supervised ML algorithms can further be divided into several categories at different hierarchies. A complete taxonomy of supervised learning techniques can be found in [121]. Some well-known techniques include RF, conditional RFs (CRFs), SVMs, decision trees, NNs, logistic and linear regressions, belief networks, naive Bayes, and Markov random fields and Markov models.

In education, supervised learning is mostly used in predictive analysis, such as grading [122], retention, and dropout prediction [123]. For instance, Majeed et al. [122] proposed several supervised learning techniques for students' grade prediction. In detail, around 2500 students' records were collected from a degree-awarding institution to train different supervised learning algorithms including naive Bayes and K-nearest neighbor classifiers. Similarly, in [124], several supervised learning algorithms including decision tree-based algorithms, naive Bayes, k-NN, linear models, and deep learning, are employed for the identification of students at risk using around 15 825 samples from the Budapest University of Technology and Economics.

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

One of the main limitations of a supervised learning-based strategy in the education sector is the lack of quality datasets, as detailed in Section V. In order to overcome these limitations, a modified form of supervised learning, namely, semisupervised learning aiming to exploit partially labeled train sets for classification tasks, has been introduced. For instance, Livieris et al. [38] proposed a semisupervised learning-based framework for secondary school students' performances.

B. Unsupervised Learning

Unsupervised learning, which aims to discover or extract patterns of regularities and irregularities in a set of observations, has also been widely exploited in educational data analysis. Unsupervised algorithms process and discover hidden patterns in input samples without needing any training samples, and thus, are easy to implement and deploy in an application. Unsupervised ML algorithms can be mainly divided into two categories, namely, 1) clustering and 2) dimensionality reduction techniques, which are further divided into subgroups. A detailed taxonomy of unsupervised ML algorithms has been provided in [121]. Clustering algorithms aim to divide a collection of samples into clusters or segments while dimensionality reduction algorithms are used to extract a small set of relevant features for building a reliable model.

In literature, unsupervised learning—especially clustering algorithms— has been mostly used in EDM to extract useful information for a diverse set of applications from raw data [125]. Some of the applications of EDM in which clustering has been proven very effective include students' performance prediction [126], students' profiling and modeling [127], recommendation systems [128], enrollment management [129], constructing course contents [130], and analyzing students' behavior [131]. Similarly, dimensionality reduction algorithms, such as principal component analysis (PCA) and linear discriminant analysis (LDA), have also been employed in educational data analysis. For instance, Borges et al. [132] employed PCA for students' performance prediction and data analysis.

C. Reinforcement Learning

In the beginning, RL was mostly restricted to robotics and game theory, however, more recently it has been deployed in other application domains as well [121]. A significant portion of literature, especially the work presented in top venues, as detailed in Section IV, is based on RL. RL provides a set of recommended actions to maximize reward in a particular situation/application. RL differs from supervised learning in several ways. For instance, supervised learning algorithms are trained on class labels to predict a class while RL algorithms are trained on a reward signal and predict/recommend an action to solve a particular problem. Moreover, RL performs a task in a sequential way where the input depends on the previous decision.

Similar to supervised and unsupervised learning, RL algorithms can be divided into different categories. RL algorithms can mainly be categorized as Markovian or evolutionary. A complete taxonomy of RL can be found in [121].

In education, RL has been mainly used for generating feedback for students on time series data [133], modeling students' learning style [134], personalized learning [135], adaptive tutorial modeling [136], and improving students' problem-solving capabilities [137].

D. Generative AI

Generative AI is a cutting-edge field within AI that focuses on creating models capable of generating new content, such as text, images, and even music. Unlike traditional ML techniques like supervised learning, unsupervised learning, and RL, which primarily focus on pattern recognition and optimization, generative AI models aim to simulate human creativity and produce original outputs in various forms including text, audio, images, and videos. In recent years, generative AI has gained immense popularity, thanks to the emergence of pretrained large language models (LLMs) trained on vast amounts of Internet data using large-scale parallel computing with new transformer-based techniques [138]. These pretrained models, often referred to as foundation models [139], are designed to be versatile and applicable to a wide range of tasks that were not explicitly specified during their training. They can be utilized with prompts to enable zero-shot learning, where they provide direct answers, or one-shot/multishot learning, where prompts include examples. Furthermore, these models can be fine-tuned for more specific applications later on. Prominent examples of foundation models include BERT, T5, and the GPT family, including GPT3 and GPT4. The widespread adoption of generative AI has been fueled by the popularity of platforms like ChatGPT, which leverages a conversational interface built on top of the GPT3.5 model, capturing the public's attention and imagination. These foundation models have revolutionized entire fields such as NLP by eliminating the need for developing bespoke models. Instead, general-purpose models are built once and can be utilized in various contexts for a multitude of tasks.

This advancement holds great promise in the field of education, particularly since the introduction of ChatGPT on top of GPT3.5 and GPT4, which has captured the interest of the education community since the release of ChatGPT near the end of the year 2022. Educators have recognized the transformative potential of this technology in enabling personalized education through a conversational interface. Recent studies and research have delved into the implications of generative AI and LLMs in education, highlighting both the promise and the potential risks associated with these technologies [140], [141], [142]. The generative AI models are used for several tasks in the education sector, such as sentiment analysis and feedback [143], interactive chat-based teacher training [144], and student assessments [145] etc.

IV. Bibliometric Analysis

In the bibliometric analysis, we analyze the research trends in AI in education. Such analysis is an integral part of the research evaluation methodology in different domains [146]. We believe the bibliometric analysis of the domain over the last few years could be useful for the community. It indicates recent trends in the domain. To this aim, we have collected 4447

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

TABLE VII
STATISTICS OF THE DATASET USED FOR THE BIBLIOMETRIC ANALYSIS

Venue	# Articles
Journal of Educational Data Mining	150
International Conference of Educational Data Mining	922
ACM Conference on Learning at Scale	510
International Conference on Learning Analytics	540
International Conference on Artificial Intelligence in Education	743
International Conference on Intelligent Tutoring Systems	342
Journal of Learning Analytics	267
British Journal of Educational Technology	787
International Journal of Artificial Intelligence in Education	336

articles from the top venues including five conferences and four journals in the domain. The choice of these venues for the study is motivated by a significant portion of literature covered in these venues. It is to be noted that our search includes the following keywords: EDM, ML for education, ITSs, intelligent tutor, AI tutor, AI for education, ML for education, intelligent classroom, generative AI for education, generative AI for learning, large language models for education, student-agent discourse analysis, ChatGPT-powered education, GPT tutor, transformer models for virtual tutor, and large language models for educational settings. Finally, we checked whether the resultant articles in our search queries incorporated data-driven AI techniques as their methodology and kept only those articles in our dataset.

Some statistics of the data used in the analysis are provided in Table VII, which include 922 articles from the International Conference of EDM [11], 150 articles from the Journal of Educational Data Mining (JEDM) [147], 510 articles from the ACM Conference on Learning at Scale (L@S) [148], 540 articles from the International Conference on Learning Analytics and Knowledge (ILAK) [149], 743 articles from International Conference on Artificial Intelligence in Education (AIED) [150], 342 articles from International Conference on ITS [151], 267 articles from Journal of Learning Analytics (LAK) [152], 787 articles from British Journal of Educational Technology (BJET) [153], and 336 articles from International Journal of Artificial Intelligence in Education (IJAIED) [154]. These numbers from the top venues are expected to provide a reasonable generalization of the research trends in the domain. The data were obtained from various sources, including ACM Digital Library [155], Scopus [156], and CrossRef [157]. Data from the CrossRef repository were scraped using Harzing's Publish or Perish utility [158]. These libraries index a complete set of research articles from the aforementioned venues and allow users to extract data based on different features including author name, affiliation name, source venue name, years, and funding status. We extracted these papers published on the abovementioned venues using filters on source venue names irrespective of location or affiliation of the authors. This approach gave us a comprehensive dataset of papers covering our considered venues. Finally, we did a manual check to confirm that the extracted papers do not include any papers from the venues outside our considered list. In detail, we analyze several factors, namely: 1) authors-based productivity analysis; 2) institution and country-based productivity analysis; 3) knowledge flow by highlighting the cross-references of different venues; 4) the relationship between the applications and techniques; 5) relationship between applications and venues; and 6) the relationship between techniques and venues.

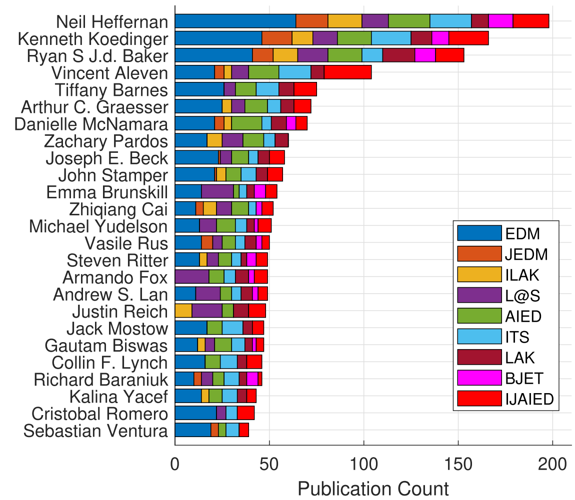

Fig. 1. Authors with the highest publication count during 2014–2022. *USA-based authors appear to be prominent in this list.*

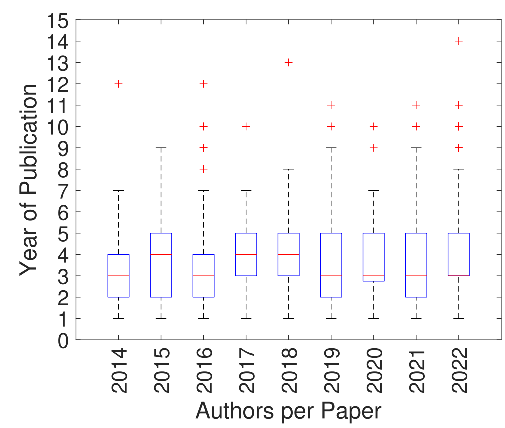

Fig. 2. Distribution of authorship during 2014–2022 in our dataset. *Although a median number of authors remained more or less constant throughout the mentioned time period but spread of co-authorship increased in recent years.*

A. Authors-Based Productivity Analysis

Author productivity is one of the common methods to evaluate significant entities. By consulting the work of top authors in a domain, the directions of a research domain can be easily determined. Fig. 1 shows the most published authors in the field of AI in education. We observe that authors from USA-based organizations are significantly contributing to the field of AI in education. Neil Heffernan from Worcester Polytechnic Institution, Kenneth Koedinger from Carnegie Mellon University, and Ryan S.J.D. Baker from the University of Pennsylvania are collectively ranked as the top three authors in the field of EDM. Most of the subsequent authors on the list also belong to USA-based institutions.

We also observe the trends of coauthorship in AI in education in Fig. 2. The average number of authors during 2014–2022 remains more or less constant at three authors per paper but the spread of authorship has increased in recent years. Some papers have experienced a higher number of coauthorship as well,

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

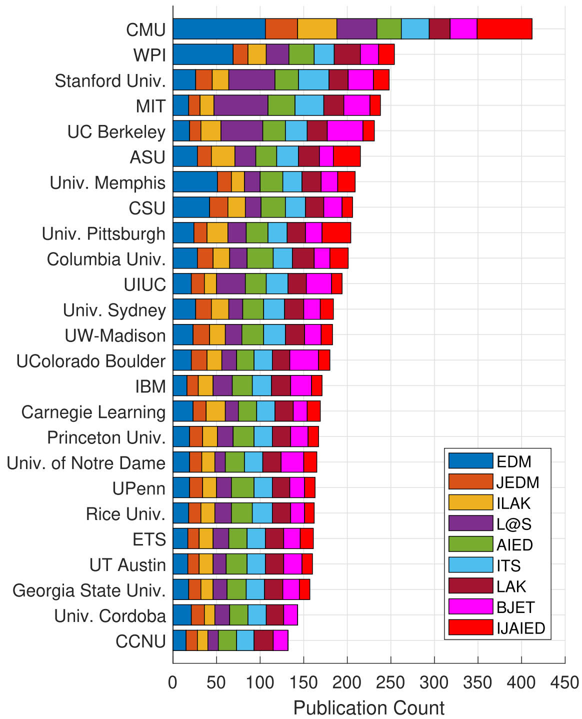

Fig. 3. Institutions with the highest number of publications in our dataset during 2014–2022. *Almost all of the top publishing institutions are from USA.*

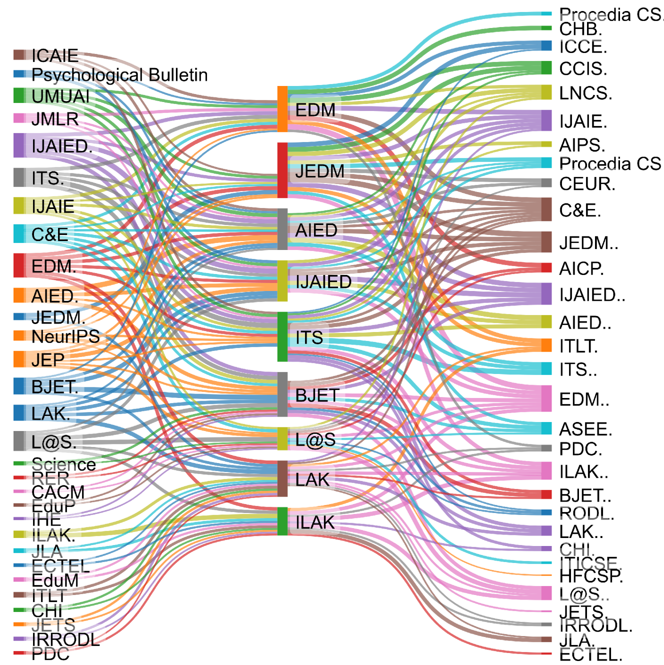

Fig. 5. Distribution references and citations in AI in education venues during 2014–2022. The left input shows the conferences that are referenced by our dataset; the right output shows which papers cite publications in our dataset. *Major sources of references and citations in our dataset are from journals.*

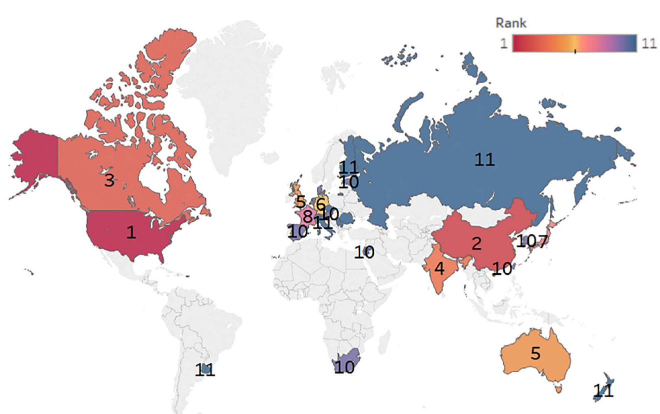

Fig. 4. Rank of countries based on their publication count in our dataset during 2014–2022. *USA emerges as top contributor followed by China, Canada, India, and Australia/U.K.*

e.g., in 2022, the maximum number of authors of a paper was 14 authors.

B. Institution and Country-Based Productivity Analysis

This subsection deals with the varying research trends of AI in education in different institutions and countries. Fig. 3 shows the most publishing institutions in the field of AI in education. Almost all of the top institutions are from the USA, which shows the significant research contributions in this domain by the USA.

Fig. 4 shows the rank of a contributing country in the field of AI in education using a global heat map. The United States is in the highest position in the field of AI in education in terms of publication count. Other top countries include China, Canada, India, and the United Kingdom in AI in education.

C. Knowledge Flow

First, we extract the references from all papers and create a citation graph, as we are curious to understand how venues in AI in education cite each other. Fig. 5 is a Sankey diagram that shows the fraction of papers that AI in education papers reference (left), as well as the other papers that in turn cite the papers in our dataset (right).

Interesting patterns emerge from this analysis. Most noteworthy is the bias for citing papers from the same venue. For example, 31% of the references in papers for AI in education are from other papers previously published in the AI in education conferences. In contrast, a far more diverse set of conferences and journals cites articles from venues considered in our dataset. 51% of the papers in our dataset are cited by journals, rather than conferences. Major citers of papers in our dataset include Computer and Education and LNCS (which subsumes many proceedings) besides the venues where the paper is published.

All that said, it is clear that several other publication venues feature heavily in the bibliographies of papers of our dataset, and these are dominated by journals rather than conferences.

D. Relationships Between Applications and Techniques

It is also important to provide readers with an overview of AI techniques employed in different applications of AI in education. To this aim, in Fig. 6 we provide the statistics of four main

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

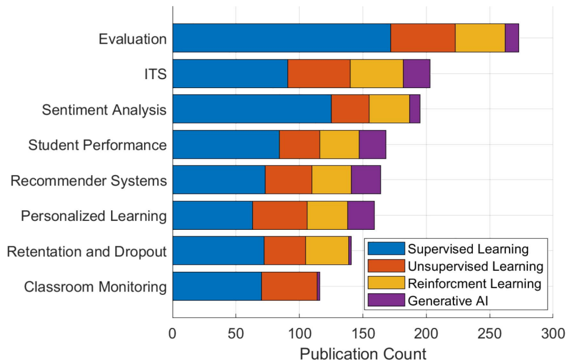

Fig. 6. Relationship between applications and AI techniques.

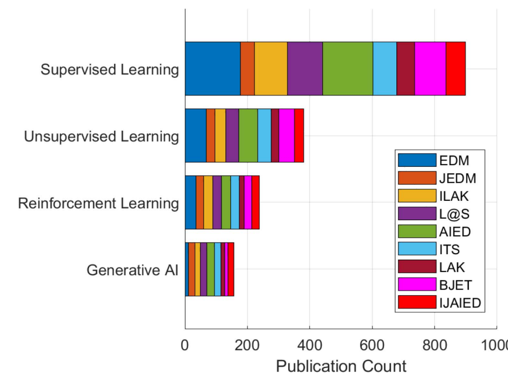

Fig. 8. Statistics of AI papers at different venues and techniques.

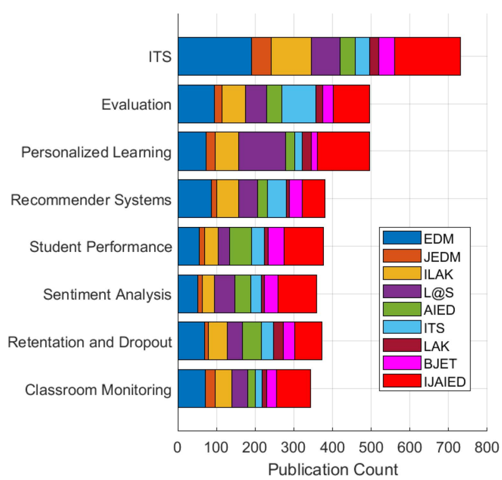

Fig. 7. Statistics of AI papers at top venues in terms of applications.

categories of AI techniques using data-driven AI methodologies in terms of the number of publications in different applications in top venues. As can be observed, in most of the applications, supervised or semisupervised techniques of learning have been employed suggesting the availability of the annotated data in the majority of the applications. Unsupervised learning techniques have also been widely employed in some of the applications, such as e-learning, student evaluation, ITS, and personalized learning. Similarly, RL has also been employed in several works on ITS, student evaluation and retention, and dropout prediction. We also observe the growing number of works on ITS, recommender systems, and personalized learning are using generative AI.

E. Relationship Between Applications and Venues

Fig. 7 provides the statistics of some interesting applications of AI in education in terms of the number of papers published on each in the leading venues. The most popular application of AI in education is in developing ITSs followed by its use for evaluation and personalized learning.

F. Relationship Between Techniques and Venues

Fig. 8 shows the statistics of the papers published based on the four types of data-driven AI methods, namely: 1) supervised learning, 2) unsupervised learning, 3) RL, and 4) generative AI. As can be seen in the figure, algorithms from each of the four categories have been deployed in educational data analysis works presented in the aforementioned venues in our dataset. The most common technique type by far is supervised learning followed by unsupervised learning and RL, and then generative AI.

V. Discussion: Insights, Pitfalls, Future Research, and Open Issues

In this section, we provide some key insights from literature on data-driven AI in education, limitations of AI in education, and future research directions and open issues.

A. Insights

Some key insights from literature on the use of AI in education discussed in the article are summarized as follows.

1) *Success of AI in Education:*

a) *Changing roles of humans and machines in education:* Technology advancements in AI might eventually change the roles of teachers substantially with their traditional duties of knowledge dissemination changing to becoming coaches focusing on assessment, mentoring, and monitoring [159]. To this aim, they will need to develop new skills, such as an in-depth understanding of the new education system offered by modern technologies and AI [160].

b) *AI is a learning catalyst:* AI-powered systems are expected to enhance learning as catalysts, particularly for children. These systems will intelligently adapt to individual learning needs, addressing areas, such as writing, reading, social skills, and soft skills [161].

c) *Teacher and AI Collaboration:* The ever-increasing role of AI in education is expected to help fill the gaps in learning and teaching allowing teachers to perform more efficiently than ever. AI's support of teachers in personalization, evaluation, and testing allows teachers to spend more time on tasks that are beyond machines AI and require human

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

capabilities [162]. Leveraging the capabilities of AI-driven technology and teachers is expected to result in a better learning and teaching environment [163].

d) *Learning for all:* AI plays a crucial role in providing improved learning resources for students with special needs. Efforts have been made to develop AI-based educational tools specifically designed for disabled students, resulting in more affordable and enhanced learning environments and materials [164]. This approach offers a better learning experience for special-needs children without the need for costly and time-consuming therapy sessions. Furthermore, AI-driven machine translation tools are assisting second-language students in overcoming language barriers. The revolutionizing impact of AI is anticipated in the field of special education as well [165].

e) *Balanced use of robot teachers:* Successful examples of robot teachers like Elias, a language teacher in Finland, and Jill Watson, a virtual teaching assistant in the US [166], highlight their reliability in covering topics consistently and offering positive feedback to human teachers for innovative teaching [167]. However, challenges arise from their lack of human touch, creativity, discipline enforcement, and the special teacher–pupil bond. While AI is expected to outperform humans in various activities [168], the dynamic and inspirational qualities of human teachers remain vital [169]. AI technology can support human teachers in understanding student needs, but caution is needed to prevent the loss of human traits and the potential shortage of quality teaching staff in developed countries.

2) Limitations of AI in Education/Learning: There are several aspects of education where AI alone cannot contribute much. The limitations and pitfalls of AI in education can be mainly divided in terms of technological and social aspects. The technological pitfalls of AI in education are either due to conceptual/algorithmic limitations or because of the training data. Some of the pitfalls are listed as follows.

a) *Failure in the extraction of interpretable and actionable insights:* AI alone is not enough to fully understand and extract interpretable and actionable insights from educational data to improve students' learning. For instance, in [170], several case studies have been reported where simply AI-based predictions are not enough to understand and improve the learning process. The authors, rather propose an explanatory learning model by employing human–computer interaction (HCI) and AI (i.e., model interpretation approaches at the interpretation stage [171]) techniques to derive insights from the student's learning experiences and suggest how the technology could be made more useful for the learners.

b) *Failure in generation of course content:* All AI techniques can do is to recommend a particular chapter/course content to a student at a timestamp (i.e., alter the sequence of the course materials). According to Popovic [172], presenting the same material in a different sequence has little impact on the learner's performance, and the real game-changer is the generation of course content on the fly, which is a very challenging task.

c) *Limitations of robotic teachers:* While content and learning analytics have greatly contributed to the development of personalized educational content, concerns remain regarding the clarity and flexibility of teaching methods employed by virtual robots serving as tutors or teaching assistants. Traditional teachers play a vital role in motivating students to learn and excel in their courses, but robots lack the capacity for such interpersonal engagement [173].

d) *Lack of training data:* The strength of AI techniques comes from training data, which has a significant impact on their prediction capabilities [174]. However, it is very challenging to acquire a sufficient amount of training samples for AI algorithms in a sensitive and high-stakes environment, such as the education sector, where one cannot afford any risk with students [172], [175].

e) *High risk due to biased data:* AI algorithms need precise and sound data to be more effective. A high risk of biases is involved with AI in education, where it is very probable to reach false conclusions due to inaccurate predictions that may benefit a more advantaged group of learners [176].

f) *Testing and evaluation of AI-systems*: Education is one of the critical applications where several risks are associated with the deployment of AI-based solutions [139]. In order to develop users' trust in AI systems in education, the solutions need to be properly formulated, trained, and evaluated on representative and real-world data before deployment, which is a challenging process.

g) *Security concerns:* The increasing dependence on AI will also lead to serious privacy concerns [177]. The institutions would need to focus not only on quality but also on data privacy. According to Calhoun Williams [172], in schools, data need to be carefully handled and the administrations need to be ready for AI from a policy standpoint.

AI in education has societal drawbacks beyond algorithm limitations [178]. Deploying AI at the school level can lead to children's technology addiction, negatively affecting their health and personalities [179]. In addition, technology use in learning limits interaction with peers and teachers, potentially causing isolation [8]. Moreover, AI implementation increases dependence on expensive technology and raises power consumption, burdening school budgets and impeding access to quality education for the underprivileged [180]. Furthermore, AI in education may contribute to joblessness. Considering these factors, careful considerations are necessary when deploying AI in education, including defining the aspects, processes, and levels of AI integration in the sector.

B. Future Research Directions and Open Issues

In this section, we provide some potential directions for future research in the domain.

1) Identifying Ethical and Privacy Issues: Developing ethically sound AI algorithms for education poses challenges due to varying editions of ethical practices. It is crucial to mitigate biases and privacy concerns when analyzing and identifying patterns in student data. For example, having access to students'

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

online search behaviors may have a long-term negative impact. Therefore, AI researchers need to look for ways to tame their algorithms and analytics when it comes to analyzing data and detecting patterns. There are already some efforts in this direction [181]. For instance, Holmes et al. [182] recommended considering multiple ethical aspects, such as fairness, accountability, transparency, bias, autonomy, agency, and inclusion, in the deployment of AI tools in education. Similarly, Akgun et al. [183] discussed, analyzed, and made recommendations for addressing ethical challenges in K-12 settings allowing the stakeholders to benefit from AI in education in an ethical and responsible way. More recently the emergence of auto-text generation tools, such as ChatGPT, posed more ethical challenges associated with the use of AI in education. Mhlanga [184] provided a detailed overview of challenges associated with the responsible and ethical use of AI and ChatGPT in education. Despite these and similar efforts, the ethical and responsible use of AI and relevant technologies in education is an open area of research and several aspects still need to be explored. For example, very few practical proposals/codes of ethics ensuring fair and responsible use of AI for education can be found in literature [185].

2) Students' Evaluation and Performance Prediction:

a) *Minimizing biased evaluation:* One of the main challenges that each educator face is how to minimize personal biases when it comes to grading and evaluation. This stems from the fact that human behaviors are hard to predict when it comes to relationships and judgment [186]. Relying on AI can reasonably help in protecting against internal biases by offering insight into students' performance based on data. However, designing AI techniques that can help in minimizing biased evaluation while keeping teachers' attitudes in mind is not easy and requires careful consideration. Moreover, the majority of the attempts in literature relied on the traditional black-box models, which do not provide any explanation of their outcome/prediction. In a critical application like education, such black-box models are not adequate [187]. Explainable AI for education is one of the potential future research directions in the domain.

b) *Predicting student's career success:* AI can play a crucial role in detecting concerning patterns and identifying students at risk of dropping out. However, developing AI solutions for predicting optimal career paths and specialization areas poses challenges due to the vast array of factors influencing students, such as their backgrounds, skills, biological variations, environmental aspects, and individual needs. To effectively predict a student's career success and guide them toward the most suitable career path, a comprehensive AI software application is necessary. A notable study [188] underscores the efficacy of AI in predicting postgraduation employment.

3) Personalized Learning:

a) *Customized teaching pedagogy:* Pedagogical Models represent the methods practiced by an effective teacher to better engage the students in a challenging learning environment. AI-based solutions have been proven very effective in the domain allowing educators to develop effective pedagogical models, strategies, and methods to support individuals through data analytics. Literature reports several interesting AI-based solutions to identify teaching pedagogy better suited for individuals. For instance, Xiao et al. [189] employed AI techniques to assess and identify effective pedagogical factors leading to a better learning practice for primary school students. However, research is still needed to identify the best teaching pedagogy that suits each learner's skills and interests. A customized teaching pedagogy that offers effective adjustments could help students in grasping new concepts and course materials effectively [190].

b) *Generation of course contents on the fly:* AI techniques have been proven very effective in course content recommendations. However, it will be very interesting to investigate how AI can be employed in content generation for a particular learner, which will be a real game-changer in the education sector [172].

c) *Scheduling efficiency:* Optimal learning is connected to optimal scheduling of learning lessons and activities [191]. Knowing the best way to design optimal teaching schedules is challenging due to the contributing factors such as understanding how people learn (cognitive psychology, knowledge retention, etc.), topic, age, level of a learner, availability of qualified teachers, availability of resources such as physical space, etc. AI modeling for optimized and adaptive teaching policies when it comes to effective scheduling is an area of research. In [192], authors proposed online job scheduling using AI. Such efforts can help in understanding the need for effective scheduling of learning lessons.

4) Sentiment Analysis in Education:

a) *Better analysis of learners' feedback:* Sentiment analysis allows teachers and administrators to analyze and understand students' feedback and their learning experience in a better way. In distance learning, sentiment analysis could also be proved very effective in several ways. For instance, students' early failure could be predicted by analyzing their feelings, feedback, and learning experience in a course via sentiment analysis, which can ultimately improve the graduation rate by taking necessary remedial actions [193]. Sentiment analysis of students' feedback could also benefit personalized learning tools to further improve/stimulate students' enthusiasm for learning. The role of sentiment analysis in the education sector is expected to go beyond analyzing students' feedback on a course or a teacher.

b) *Visual sentiment analysis:* Visual sentiment analysis in education involves extracting learners' opinions and emotions from visual content, which could be very useful in various ways [194]. For instance, they could help to automatically filter and summarize visual educational content based on perceived sentiments and emotions. It could also help to perceive and stimulate students' interest in both traditional and on-screen teaching/e-learning [195].

5) Classroom Monitoring and Visual Analysis:

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

a) *Behavior modeling:* Besides student attendance, classroom monitoring technology could also be used to analyze students' behavior in classrooms for a diversified set of applications, such as losing focus during lectures, playing mobile phones, and other misbehaviors [196]. Such behavior modeling and analysis will allow monitoring of students' personality-building process subject to their consent.

b) *Technology integration in classroom:* Developing an effective model for integrating technology in education is a critical and challenging endeavor for several reasons. The diversity of technology products, learners' skills and interests, and knowledge areas add complexity to the process. While researchers primarily focus on the benefits and optimal methods of technology integration in classrooms [197], it is essential to acknowledge that there are potential negative impacts that cannot be disregarded when implementing technology [198]. Consequently, the development of an AI model for technology integration must confront these challenges and garner attention from researchers. Furthermore, the integration of NLP and AI in classrooms presents new opportunities to enhance learning and teaching effectiveness, with promising advancements already made in this field [199], [200].

6) General Issues:

a) *Risk assessment in AI-based education:* Given the serious consequences of mistakes in education, caution must be exercised when applying AI techniques, emphasizing the importance of identifying potential risks and consequences. Inaccurate or inappropriate AI recommendations can result in significant social and economic impacts, highlighting the need for AI in education researchers to incorporate a risk assessment framework to measure the implications of potential errors and mistakes.

b) *Favoring AI over traditional statistical methods:* Research is still needed to identify when AI is better than traditional statistical methods when it comes to deciding on education at different levels. Over the years, the usage of traditional statistical analytical methods has been successful (at least this is what we observed). Nevertheless, it is vital to study if AI analytics can be more successful in deciding on improving our education.

c) *Explanatory learning model:* To obtain more interpretable and actionable insights from educational data, explanatory learning models involving all the stakeholders including learners' parents and schools, etc., need to be developed [170]. AI along with HCI methods can then be jointly utilized to better analyze the data.

d) *Security implications:* AI is very dependent on data. Data in the education field are miscellaneous. Designing AI algorithms while security is very prominent and in mind is critical. This requires distinguishing between sensitive and insensitive data before jumping to apply AI techniques to educational data. Hence, researchers are in need to develop intelligent AI techniques that are ready to deal with data in classified and careful ways.

e) *Protective intelligence to secure schools:* Keeping in view the rise of shooting accidents in schools, and as also recommended in [201], schools' administration needs to regularly analyze, and evaluate students' information for potential threats, AI-based protective intelligence is one of the potential future research directions to ensure schools' safety.

VI. Conclusion

In this article, we have reviewed applications of data-driven AI in the education sector from different perspectives. On the one side, we provided a detailed overview of the existing tools and applications developed as a result of the efforts of the AI community in education. On the other hand, we highlighted the research trends in the domain over the last nine years as well as the current limitations and pitfalls of data-driven AI in education. In particular, we provided a detailed overview of existing literature in eight application domains, such as students' grading and evaluation, students' dropout, sentiment analysis, ITSs, and classroom monitoring. The efforts made in these applications resulted in several interesting tools helping students and administration in several ways. The survey also highlighted key market players, tools, and platforms in the abovementioned applications of AI in education, which are expected to provide a good starting point for beginners in the domain. We also provided an overview of the most commonly used data-driven AI strategies and techniques in different applications of AI in education. In addition, a detailed bibliometric analysis has been provided to highlight the research trends of AI in the education sector over the last few years. The bibliometric analysis shows a significant contribution from researcher USA. Moreover, students' grading and evaluation has been among the mostly explored application discussed in this article. Based on our analysis of existing literature and experience in the domain, we also identified the current limitations and pitfalls of data-driven AI in education. We believe such a detailed analysis of the domain will provide a baseline for future research in the domain.

Acknowledgment

The authors would also like to thank Dr. Sebastian Ventura for his constructive feedback, which greatly improved the manuscript's quality.

References

[1] C. Guan, J. Mou, and Z. Jiang, "Artificial intelligence innovation in education: A twenty-year data-driven historical analysis," *Int. J. Innov. Stud.*, vol. 4, no. 4, pp. 134–147, 2020.

[2] M. Bienkowski, M. Feng, and B. Means, "Enhancing teaching and learning through educational data mining and learning analytics: An issue brief," *Office Educ. Technol., US Dept. Educ.*, Washington, DC, USA, 2012.

[3] M. Mitchell, *Artificial Intelligence: A Guide for Thinking Humans*. Penguin, U.K.: Farrar, Straus and Giroux, 2019.

[4] P. Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York, NY, USA: Basic Books, 2015.

[5] "Gthe use of artificial intelligence (AI) in education," 2020, Accessed: Apr. 13, 2020. [Online]. Available: https://tinyurl.com/y44hym7y

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

[6] A. K. Goel and D. A. Joyner, “Design of an online course on knowledge-based AI,” in *Proc. AAAI Conf. Artif. Intell.*, 2016, pp. 4089–4094.

[7] “Artificial intelligence in education: Where are we now?,” 2020, Accessed: Dec. 1, 2020. [Online]. Available: https://www.youtube.com/watch?v=54-yKBv6xeU

[8] K. Holstein, “Towards teacher-AI hybrid systems,” in *Companion Proc. 8th Int. Conf. Learn. Analytics Knowl.*, 2018.

[9] “How artificial intelligence is being used in the classroom,” 2020, Accessed: Apr. 13, 2020. [Online]. Available: https://tinyurl.com/yy2oomye

[10] C. Romero and S. Ventura, “Educational data mining and learning analytics: An updated survey,” *Wiley Interdiscipl. Rev.: Data Mining Knowl. Discov.*, vol. 10, no. 3, 2020, Art. no. e1355.

[11] “Educationl data mining,” 2023, Accessed: Jul. 19, 2023. [Online]. Available: http://educationaldatamining.org/

[12] “International artificial intelligence in education society,” 2023, Accessed: Jun. 04, 2023. [Online]. Available: https://tinyurl.com/2ar673vf

[13] G.-J. Hwang and Y.-F. Tu, “Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review,” *Mathematics*, vol. 9, no. 6, 2021, Art. no. 584.

[14] C. Romero and S. Ventura, “Educational data mining: A survey from 1995 to 2005,” *Expert Syst. Appl.*, vol. 33, no. 1, pp. 135–146, 2007.

[15] R. S. Baker and K. Yacef, “The state of educational data mining in 2009: A review and future visions,” *J. Educ. Data Mining*, vol. 1, no. 1, pp. 3–17, 2009.

[16] A. Charitopoulos, M. Rangoussi, and D. Koulouriotis, “On the use of soft computing methods in educational data mining and learning analytics research: A review of years 2010–2018,” *Int. J. Artif. Intell. Educ.*, vol. 30, pp. 371–430, 2020.

[17] C. Fischer et al., “Mining Big Data in education: Affordances and challenges,” *Rev. Res. Educ.*, vol. 44, no. 1, pp. 130–160, 2020.

[18] N. Mduma, K. Kalegele, and D. Machuve, “A survey of machine learning approaches and techniques for student dropout prediction,” *Data Sci. J.*, vol. 18, no. 14, pp. 1–10, 2019.

[19] A. Almasri et al., “Intelligent tutoring systems survey for the period 2000-2018,” *Int. J. Academic Eng. Res.*, vol. 3, no. 5, pp. 21–37, 2019.

[20] E. Mousavinasab, N. Zarifsanaiey, S. R. N. Kalhori, M. Rakhshan, L. Keikha, and M. G. Saeedi, “Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods,” *Interactive Learn. Environments*, vol. 29, no. 1, pp. 142–163, 2021.

[21] M. Malekzadeh, M. B. Mustafa, and A. Lahsasna, “A review of emotion regulation in intelligent tutoring systems,” *J. Educ. Technol. Soc.*, vol. 18, no. 4, pp. 435–445, 2015.

[22] F. Tahiru, “AI in education: A systematic literature review,” *J. Cases Inf. Technol.*, vol. 23, no. 1, pp. 1–20, 2021.

[23] M. Al-Emran, S. I. Malik, and M. N. Al-Kabi, “A survey of Internet of Things (IoT) in education: Opportunities and challenges,” in *Toward Social Internet of Things (SIoT): Enabling Technologies, Architectures and Applications*. Berlin, Germany: Springer, 2020, pp. 197–209.

[24] W. Matcha, N. A. Uzir, D. Gašević, and A. Pardo, “A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective,” *IEEE Trans. Learn. Technol.*, vol. 13, no. 2, pp. 226–245, Apr.–Jun. 2020.

[25] D. Hooshyar, M. Pedaste, K. Saks, Ä. Leijen, E. Bardone, and M. Wang, “Open learner models in supporting self-regulated learning in higher education: A systematic literature review,” *Comput. Educ.*, vol. 154, 2020, Art. no. 103878.

[26] A. Zaidi, S. Beadle, and A. Hannah, “Review of the online learning and artificial intelligence education market: A report for the department of education: July 2018,” Dept. Educ., London, U.K., Rep. DfE-00112-2019, 2019.

[27] Y. Kim and A. L. Baylor, “Based design of pedagogical agent roles: A review, progress, and recommendations,” *Int. J. Artif. Intell. Educ.*, vol. 26, pp. 160–169, 2016.

[28] F. L. da Silva, B. K. Slodkowski, K. K. A. da Silva, and S. C. Cazella, “A systematic literature review on educational recommender systems for teaching and learning: Research trends, limitations and opportunities,” *Educ. Inf. Technol.*, vol. 28, no. 3, pp. 3289–3328, 2023.

[29] S. S. Khanal, P. Prasad, A. Alsadoon, and A. Maag, “A systematic review: Machine learning based recommendation systems for e-learning,” *Educ. Inf. Technol.*, vol. 25, pp. 2635–2664, 2020.

[30] N. W. Rahayu, R. Ferdiana, and S. S. Kusumawardani, “A systematic review of learning path recommender systems,” *Educ. Inf. Technol.*, vol. 28, no. 6, pp. 7437–7460, 2023.

[31] C. Romero and S. Ventura, “Educational data mining: A review of the state of the art,” *IEEE Trans. Syst., Man, Cybern., Part C (Appl. Rev.)*, vol. 40, no. 6, pp. 601–618, Nov. 2010.

[32] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, “Systematic review of research on artificial intelligence applications in higher education–where are the educators?,” *Int. J. Educ. Technol. Higher Educ.*, vol. 16, no. 1, 2019, Art. no. 39.

[33] J.-C. Liang, G.-J. Hwang, M.-R. A. Chen, and D. Darmawansah, “Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach,” *Interactive Learn. Environments*, vol. 31, pp. 1–27, 2021.

[34] A. Dutt, M. A. Ismail, and T. Herawan, “A systematic review on educational data mining,” *IEEE Access*, vol. 5, pp. 15991–16005, 2017.

[35] K. Loderer, R. Pekrun, and J. C. Lester, “Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments,” *Learn. Instruct.*, vol. 70, 2020, Art. no. 101162.

[36] B. Flanagan and H. Ogata, “Learning analytics platform in higher education in Japan,” *insKnowl. Manage. E-Learn.: An Int. J.*, vol. 10, no. 4, pp. 469–484, 2018.

[37] O. W. Adejo and T. Connolly, “Predicting student academic performance using multi-model heterogeneous ensemble approach,” *J. Appl. Res. Higher Educ.*, vol. 10, no. 1, pp. 61–75, 2018.

[38] I. E. Livieris, K. Drakopoulou, V. T. Tampakas, T. A. Mikropoulos, and P. Pintelas, “Predicting secondary school students’ performance utilizing a semi-supervised learning approach,” *J. Educ. Comput. Res.*, vol. 57, no. 2, pp. 448–470, 2019.

[39] B.-H. Kim, E. Vizitei, and V. Ganapathi, “GritNet: Student performance prediction with deep learning,” 2018, *arXiv:1804.07405*.

[40] L. Aulck, N. Velagapudi, J. Blumenstock, and J. West, “Predicting student dropout in higher education,” 2016, *arXiv:1606.06364*.

[41] M. Solis, T. Moreira, R. Gonzalez, T. Fernandez, and M. Hernandez, “Perspectives to predict dropout in university students with machine learning,” in *Proc. IEEE Int. Work Conf. Bioinspired Intell.*, 2018, pp. 1–6.

[42] A. Pilkington, P. Bowen, R. C. Rose, D. R. Rajasinghe, and I. Evans, “A qualitative study of student retention: The university javascript: Void (0); academic’s perspective,” *Int. J. Academic Multidisciplinary Res.*, vol. 2, no. 10, pp. 19–28, 2018.

[43] V. Tinto, *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Chicago, IL, USA: Univ. Chicago press, 1987.

[44] Kinshuk, N.-S. Chen, I.-L. Cheng, and S. W. Chew, “Evolution is not enough: Revolutionizing current learning environments to smart learning environments,” *Int. J. Artif. Intell. Educ.*, vol. 26, no. 2, pp. 561–581, 2016.

[45] S. Liu, W. Cui, Y. Wu, and M. Liu, “A survey on information visualization: Recent advances and challenges,” *Vis. Comput.*, vol. 30, no. 12, pp. 1373–1393, 2014.

[46] C. Taylor, K. Veeramachaneni, and U.-M. O’Reilly, “Likely to stop? Predicting stopout in massive open online courses,” 2014, *arXiv:1408.3382*.

[47] F. Ouyang, L. Zheng, and P. Jiao, “Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020,” *Educ. Inf. Technol.*, vol. 27, pp. 7893–7925, 2022.

[48] M. Lee, E. Croteau, A. Gurung, A. Botelho, and N. Heffernan, “Knowledge tracing over time: A longitudinal analysis,” in *Proc. 16th Int. Conf. Educ. Data Mining*, 2023, pp. 1–6.

[49] K. M. Hasib, F. Rahman, R. Hasnat, and M. G. R. Alam, “A machine learning and explainable AI approach for predicting secondary school student performance,” in *Proc. IEEE 12th Annu. Comput. Commun. Workshop Conf.*, 2022, pp. 0399–0405.

[50] K. Alalawi, R. Athauda, and R. Chiong, “Contextualizing the current state of research on the use of machine learning for student performance prediction: A systematic literature review,” *Eng. Rep.*, 2023, Art. no. e12699.

[51] “Campuslabs: Data driven innovations,” 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.campuslabs.com/campus-labs-platform/retention-and-success/

[52] “RNL student retention predictor,” 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/voh3cjk

[53] “Nuro retention,” 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.campuslabs.com/campus-labs-platform/retention-and-success/

[54] “Othot: Retention analytics,” 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/mshyxedf

[55] “Campusnexus succeed: Retention analytics,” 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.campusmanagement.com/products/student-retention-software/

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

[56] "Pearson's WriteToLearn," Pearson, London, U.K., 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.pearson.com/country-selector.html
[57] "Quantum adaptive learning," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.quantumal.com/
[58] "Azure for education," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://azure.microsoft.com/en-us/education/
[59] "Lightside labs," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/yadrjze6
[60] "Proctorio," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/yxy9e74q
[61] "GradeScope: A grading tool," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.gradescope.com/
[62] P. Sruthi and S. Mukherjee, "Byju's the learning app: An investigative study on the transformation from traditional learning to technology based personalized learning," *Int. J. Sci. Technol. Res.*, vol. 9, no. 3, pp. 5054–5059, 2020.
[63] C. Wilson and B. Scott, "Adaptive systems in education: A review and conceptual unification," *Int. J. Inf. Learn. Technol.*, vol. 34, pp. 2–19, 2017.
[64] D. Shawky and A. Badawi, "Towards a personalized learning experience using reinforcement learning," in *Machine Learning Paradigms: Theory and Application*. Cham, Switzerland: Springer, 2019, pp. 169–187.
[65] "Third space learning," Third Space Learning, London, U.K., 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://thirdspacelearning.com
[66] "Enlearn," Enlearn, Seattle WA, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.enlearn.org/
[67] "Watson incorporation," 2023, Accessed: Jul. 19, 2023. [Online]. Available: www.ibm.com/watson/education
[68] "Querium: The home master," 2023, Accessed: Jul. 19, 2023. [Online]. Available: http://querium.com/
[69] "Online learning is our business—Let us do it for you," Edly, McKinney, TX, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://edly.io
[70] "Squirrel AI learning," Squirrel AI, Shanghai, China, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://squirrelai.com/our-story/
[71] "Mobymax: Fix learning gaps," Mobymax, Pittsburgh, PA, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.mobymax.com/
[72] J. A. Kulik and J. Fletcher, "Effectiveness of intelligent tutoring systems: A meta-analytic review," *Rev. Educ. Res.*, vol. 86, no. 1, pp. 42–78, 2016.
[73] N. Bosch et al., "Detecting student emotions in computer-enabled classrooms," in *Proc. Int. Joint Conf. Artif. Intell.*, 2016, pp. 4125–4129.
[74] M. Mavrikis, "Modelling student interactions in intelligent learning environments: Constructing Bayesian networks from data," *Int. J. Artif. Intell. Tools*, vol. 19, no. 06, pp. 733–753, 2010.
[75] R. S. Baker, J. Clarke-Midura, and J. Ocumpaugh, "Towards general models of effective science inquiry in virtual performance assessments," *J. Comput. Assist. Learn.*, vol. 32, no. 3, pp. 267–280, 2016.
[76] L. Fratamico, C. Conati, S. Kardan, and I. Roll, "Applying a framework for student modeling in exploratory learning environments: Comparing data representation granularity to handle environment complexity," *Int. J. Artif. Intell. Educ.*, vol. 27, no. 2, pp. 320–352, 2017.
[77] V. Kostakos and M. Musolesi, "Avoiding pitfalls when using machine learning in HCI studies," *Interactions*, vol. 24, no. 4, pp. 34–37, 2017.
[78] Y. Zhou and M. W. Evens, "A practical student model in an intelligent tutoring system," in *Proc. 11th Int. Conf. Tools Artif. Intell.*, 1999, pp. 13–18.
[79] S. Katz, P. Albacete, I.-A. Chounta, P. Jordan, B. M. McLaren, and D. Zapata-Rivera, "Linking dialogue with student modelling to create an adaptive tutoring system for conceptual physics," *Int. J. Artif. Intell. Educ.*, vol. 31, pp. 397–445, 2021.
[80] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," *Decis. Support Syst.*, vol. 74, pp. 12–32, 2015.
[81] M. Deschênes, "Recommender systems to support learners' agency in a learning context: A systematic review," *Int. J. Educ. Technol. Higher Educ.*, vol. 17, no. 1, 2020, Art. no. 50.
[82] F. Yang and F. W. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," *Comput. Educ.*, vol. 123, pp. 97–108, 2018.
[83] A. Elhassan, I. Jenhani, and G. B. Brahim, "Remedial actions recommendation via multi-label classification: A course learning improvement method," *Int. J. Mach. Learn. Comput.*, vol. 8, no. 6, pp. 583–588, 2018.
[84] N. S. Raj and V. Renumol, "A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020," *J. Comput. Educ.*, vol. 9, no. 1, pp. 113–148, 2022.
[85] "Qbot: Help students to reach their ful potential," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/wwvyrkf
[86] "Myedmatch," Edumatch, Alexandria, VI, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.edumatch.org/
[87] "Teachermatch," Teacher Match, Chicago, IL, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://platform.teachermatch.org/signin.do
[88] S. Hiriyannaiah, G. Siddesh, and K. Srinivasa, "Real-time streaming data analysis using a three-way classification method for sentimental analysis," *Int. J. Inf. Technol. Web Eng.*, vol. 13, no. 3, pp. 99–111, 2018.
[89] S. Baral, A. F. Botelho, A. Santhanam, A. Gurung, J. Erickson, and N. T. Heffernan, "Investigating patterns of tone and sentiment in teacher written feedback messages," in *Proc. Int. Conf. Artif. Intell. Educ.*, 2023, pp. 341–346.
[90] L. Shen, M. Wang, and R. Shen, "Affective e-learning: Using "emotional" data to improve learning in pervasive learning environment," *J. Educ. Technol. Soc.*, vol. 12, no. 2, pp. 176–189, 2009.
[91] G. S. Chauhan, P. Agrawal, and Y. K. Meena, "Aspect-based sentiment analysis of students' feedback to improve teaching–learning process," in *Information and Communication Technology for Intelligent Systems*. Berlin, Germany: Springer, 2019, pp. 259–266.
[92] Z. Kastrati, A. S. Imran, and A. Kurti, "Weakly supervised framework for aspect-based sentiment analysis on students' reviews of MOOCs," *IEEE Access*, vol. 8, pp. 106799–106810, 2020.
[93] Z. Liu, C. Yang, S. Rüdian, S. Liu, L. Zhao, and T. Wang, "Temporal emotion-aspect modeling for discovering what students are concerned about in online course forums," *Interactive Learn. Environments*, vol. 27, pp. 598–627, 2019.
[94] J. Zhou and J.-m. Ye, "Sentiment analysis in education research: A review of journal publications," *Interactive Learn. Environments*, vol. 31, pp. 1–13, 2020.
[95] M. Munezero, C. S. Montero, M. Mozgovoy, and E. Sutinen, "Exploiting sentiment analysis to track emotions in students' learning diaries," in *Proc. 13th Koli Calling Int. Conf. Comput. Educ. Res.*, 2013, pp. 145–152.
[96] Z. Kechaou, M. B. Ammar, and A. M. Alimi, "Improving e-learning with sentiment analysis of users' opinions," in *Proc. IEEE Glob. Eng. Educ. Conf.*, 2011, pp. 1032–1038.
[97] L. Mostafa, "Student sentiment analysis using gamification for education context," in *Proc. Int. Conf. Adv. Intell. Syst. Inform.*, 2019, pp. 329–339.
[98] S. Z. Hassan et al., "Visual sentiment analysis from disaster images in social media," *Sensors*, vol. 22, no. 10, 2022, Art. no. 3628.
[99] "Softbank robotics," Softbank robotics, Tokyo, Japan, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.softbankrobotics.com/emea/en/
[100] "Softbank's NAO," Softbank robotics, Tokyo, Japan, 2020, Accessed: Apr. 13, 2020. [Online]. Available: https://www.softbankrobotics.com/emea/en/nao
[101] "NAO in education," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/vuhyl9h
[102] "IRONSIDE: Predict higher education student retention risk," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/2eupujt8
[103] "Talkwalker: Understand trends and react to customer opinions instantly," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/voqwqqp
[104] S. Pozdniakov, R. Martinez-Maldonado, S. Singh, H. Khosravi, and D. Gašević, "Using learning analytics to support teachers," in *Handbook of Artificial Intelligence in Education*, Cheltenham, U.K.: Edward Elgar Publishing Limited, 2023, p. 322.
[105] R. Martinez-Maldonado et al., "Moodoo the tracker: Spatial classroom analytics for characterising teachers' pedagogical approaches," *Int. J. Artif. Intell. Educ.*, vol. 32, pp. 1025–1051, 2021.
[106] J. N. Nath, Y. S. Nath, and C. Bishwadeep, "Classroom behaviour monitoring using AI," 2020.
[107] R. M. Maldonado, J. Kay, K. Yacef, and B. Schwendimann, "An interactive teacher's dashboard for monitoring groups in a multi-tabletop learning environment," in *Proc. Intell. Tutoring Syst.: 11th Int. Conf.*, 2012, pp. 482–492.
[108] F. Paci, D. Brunelli, and L. Benini, "0, 1, 2, many—A classroom occupancy monitoring system for smart public buildings," in *Proc. Conf. Des. Architectures Signal Image Process.*, 2014, pp. 1–6.

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

[109] Y. P. Raykov, E. Ozer, G. Dasika, A. Boukouvalas, and M. A. Little, "Predicting room occupancy with a single passive infrared (PIR) sensor through behavior extraction," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 1016–1027.

[110] T. Sutjarittham, H. H. Gharakheili, S. S. Kanhere, and V. Sivaraman, "Data-driven monitoring and optimization of classroom usage in a smart campus," in *Proc. IEEE/ACM 17th Int. Conf. Inf. Process. Sensor Netw.*, 2018, pp. 224–229.

[111] S. K. Howard, J. Yang, J. Ma, C. Ritz, J. Zhao, and K. Wynne, "Using data mining and machine learning approaches to observe technology-enhanced learning," in *Proc. IEEE Int. Conf. Teach., Assessment, Learn. Eng.*, 2018, pp. 788–793.

[112] V. Soloviev, "Machine learning approach for student engagement automatic recognition from facial expressions," *Sci. Pub. State Univ. Novi Pazar Ser. A: Appl. Math., Inform. Mechanics*, vol. 10, no. 2, pp. 79–86, 2018.

[113] Y. H. V. Chua, J. Dauwels, and S. C. Tan, "Technologies for automated analysis of co-located, real-life, physical learning spaces: Where are we now?," in *Proc. 9th Int. Conf. Learn. Analytics Knowl.*, 2019, pp. 11–20.

[114] Z. Trabelsi, F. Alnajjar, M. M. A. Parambil, M. Gochoo, and L. Ali, "Real-time attention monitoring system for classroom: A deep learning approach for student's behavior recognition," *Big Data Cogn. Comput.*, vol. 7, no. 1, 2023, Art. no. 48.

[115] "Jibble: An error-free attendance app with biometric verification," Jibble, Palo Alto, CA, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.jibble.io/

[116] "LoopLearn: Automated attendance keeping," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.looplearn.net

[117] "SAFR for security provides accurate real-time face recognition," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://safr.com/

[118] "Airclass," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://info.airclass.com/goodbye

[119] "Captemo: An emotion recognition platform," Capetoma, Chicago, IL, USA, 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.captemo.com

[120] "Blippar: Transform the classroom experience," Blippar, London, U.K., 2020, Accessed: Apr. 26, 2020. [Online]. Available: https://www.blippar.com/build-ar/augmented-reality-education

[121] "Taxonomy of machine learning algorithms," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/w68oagu

[122] E. A. Majeed and K. N. Junejo, "Grade prediction using supervised machine learning techniques," in *Proc. 4th Glob. Summit Educ.*, 2016, pp. 222–234.

[123] Z. Iqbal, A. Qayyum, S. Latif, and J. Qadir, "Early student grade prediction: An empirical study," in *Proc. 2nd Int. Conf. Advancements Comput. Sci.*, 2019, pp. 1–7.

[124] M. Nagy and R. Molontay, "Predicting dropout in higher education based on secondary school performance," in *Proc. IEEE 22nd Int. Conf. Intell. Eng. Syst.*, 2018, pp. 000389–000394.

[125] H. Aldowah, H. Al-Samarraie, and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: A review and synthesis," *Telematics Inform.*, vol. 37, pp. 13–49, 2019.

[126] M. Durairaj and C. Vijitha, "Educational data mining for prediction of student performance using clustering algorithms," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 4, pp. 5987–5991, 2014.

[127] G. Akçapỳnar, A. Altun, and E. Cosgun, "Investigating students' interaction profile in an online learning environment with clustering," in *Proc. IEEE 14th Int. Conf. Adv. Learn. Technol.*, 2014, pp. 109–111.

[128] S. Dwivedi and V. K. Roshni, "Recommender system for Big Data in education," in *Proc. 5th Nat. Conf. E-Learn. E-Learn. Technol.*, 2017, pp. 1–4.

[129] F. Siraj and M. A. Abdoulha, "Uncovering hidden information within university's student enrollment data using data mining," in *Proc. Third Asia Int. Conf. Modelling Simul.*, 2009, pp. 413–418.

[130] C. Romero, S. Ventura, and E. García, "Data mining in course management systems: Moodle case study and tutorial," *Comput. Educ.*, vol. 51, no. 1, pp. 368–384, 2008.

[131] M. A. Hogo, "Evaluation of e-learners behaviour using different fuzzy clustering models: A comparative study," *Int. J. Comput. Sci. Inf. Security*, vol. 7, no. 2, 2010.

[132] V. R. P. Borges, S. Esteves, P. d. N. Araújo, L. C. d. Oliveira, and M. Holanda, "Using principal component analysis to support students' performance prediction and data analysis," in *Proc. Braz. Symp. Comput. Educ. (Simpósio Brasileiro de Informática na Educação-SBIE)*, vol. 29, no. 1, 2018, pp. 1383–1392.

[133] D. Gkatzia, H. Hastie, S. Janarthanam, and O. Lemon, "Generating student feedback from time-series data using reinforcement learning," in *Proc. 14th Eur. Workshop Natural Lang. Gener.*, 2013, pp. 115–124.

[134] F. A. Dorça, L. V. Lima, M. A. Fernandes, and C. R. Lopes, "Comparing strategies for modeling students learning styles through reinforcement learning in adaptive and intelligent educational systems: An experimental analysis," *Expert Syst. Appl.*, vol. 40, no. 6, pp. 2092–2101, 2013.

[135] P. Wang, J. P. Rowe, W. Min, B. W. Mott, and J. C. Lester, "Interactive narrative personalization with deep reinforcement learning," in *Proc. Int. Joint Conf. Artif. Intell.*, 2017, pp. 3852–3858.

[136] G. Fenza, F. Orciuoli, and D. G. Sampson, "Building adaptive tutoring model using artificial neural networks and reinforcement learning," in *Proc. IEEE 17th Int. Conf. Adv. Learn. Technol.*, 2017, pp. 460–462.

[137] J. P. Rowe and J. C. Lester, "Improving student problem solving in narrative-centered learning environments: A modular reinforcement learning framework," in *Proc. Int. Conf. Artif. Intell. Educ.*, 2015, pp. 419–428.

[138] A. Vaswani et al., "Attention is all you need," in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, 2017.

[139] R. Bommasani et al., "On the opportunities and risks of foundation models," 2021, *arXiv:2108.07258*.

[140] J. Qadir, "Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education," in *Proc. IEEE Glob. Eng. Educ. Conf.*, 2023, pp. 1–9.

[141] A. Johri, A. S. Katz, J. Qadir, and A. Hingle, "Generative artificial intelligence and engineering education," *J. Eng. Educ.*, vol. 112, no. 3, pp. 572–577, 2023.

[142] Y. K. Dwivedi et al., ""So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," *Int. J. Inf. Manage.*, vol. 71, 2023, Art. no. 102642.

[143] J. Zheng, J. Wang, Y. Ren, and Z. Yang, "Chinese sentiment analysis of online education and internet buzzwords based on BERT," in *J. Phys.: Conf. Ser.*, vol. 1631, no. 1, 2020, Art. no. 012034.

[144] J. M. Markel, S. G. Opferman, J. A. Landay, and C. Piech, "GPTeach: Interactive TA training with GPT based students," in *Proc. 10th ACM Conf. Learn. @ Scale*, 2023, pp. 226–236.

[145] J. Qadir, "Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education," in *Proc. IEEE Glob. Eng. Educ. Conf.*, 2023

[146] O. Ellegaard and J. A. Wallin, "The bibliometric analysis of scholarly production: How great is the impact?," *Scientometrics*, vol. 105, no. 3, pp. 1809–1831, 2015.

[147] "JEDM: Journal of educational data mining," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://jedm.educationaldatamining.org/index.php/JEDM

[148] "Learning at scale," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://learningatscale.acm.org/las2020/

[149] "International learning analytics and knowledge conference," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://www.solaresearch.org/conference-proceedings/

[150] "International conference on artificial intelligence in education," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://link.springer.com/conference/aied

[151] "International conference on intelligent tutoring systems," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://link.springer.com/conference/its

[152] "Journal of learning analytics," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://www.solaresearch.org/journal/

[153] "British journal of educational technology," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://onlinelibrary.wiley.com/journal/14678535

[154] "International journal on artificial intelligence in education," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://www.springer.com/journal/40593/

[155] "ACM digital library," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://dl.acm.org/

[156] "SCOPUS," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://www.scopus.com

[157] "Crossref," Crossref, Lynnfield, MA, USA, 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://www.crossref.org

[158] "Harzing's publish or perish' utility," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://harzing.com/resources/publish-or-perish

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.

[159] "DZone: Will AI replace teachers?," 2023. Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/yxyl6lft
[160] L. Shen and A. Su, "The changing roles of teachers with AI," in *Revolutionizing Education in the Age of AI and Machine Learning*. Hershey, PA, USA: IGI Global, 2020, pp. 1–25.
[161] P. Lameras and S. Arnab, "Power to the teachers: An exploratory review on artificial intelligence in education," *Information*, vol. 13, no. 1, 2021, Art. no. 14.
[162] Y. Wang, P. Han, and L. Shi, "The practical study of the collaborative teaching mode of "AI-teacher" in the 5G era," in *Proc. 7th Int. Conf. Humanities Social Sci. Res.*, 2021, pp. 772–775.
[163] "How is AI used in education—Real world examples of today and a peek into the future," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://bernardmarr.com/default.asp?contentID=1541
[164] A. S. Drigas and R.-E. Ioannidou, "Artificial intelligence in special education: A decade review," *Int. J. Eng. Educ.*, vol. 28, no. 6, 2012, Art. no. 1366.
[165] "AI has already started reshaping special education," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/y36ffu6t
[166] A. K. Goel and L. Polepeddi, "Jill Watson: A virtual teaching assistant for online education," Georgia Institute of Technology, Atlanta, GA, USA, Tech. Rep., 2016.
[167] Q. S. Mahdi, I. H. Saleh, G. Hashim, and G. B. Loganathan, "Evaluation of robot professor technology in teaching and business," *Inf. Technol. Ind.*, vol. 9, no. 1, pp. 1182–1194, 2021.
[168] P. R. Spence, A. Edwards, and C. Edwards, "Attitudes, prior interaction, and petitioner credibility predict support for considering the rights of robots," in *Proc. Companion ACM/IEEE Int. Conf. Human-Robot Interaction*, 2018, pp. 243–244.
[169] M. G. Voskoglou and A.-B. M. Salem, "Benefits and limitations of the artificial with respect to the traditional learning of mathematics," *Mathematics*, vol. 8, no. 4, 2020, Art. no. 611.
[170] C. P. Rosé, E. A. McLaughlin, R. Liu, and K. R. Koedinger, "Explanatory learner models: Why machine learning (alone) is not the answer," *Brit. J. Educ. Technol.*, vol. 50, pp. 2943–2958, 2019.
[171] J. Fiacco, E. Cotos, and C. Rosé, "Towards enabling feedback on rhetorical structure with neural sequence models," in *Proc. 9th Int. Conf. Learn. Analytics Knowl.*, 2019, pp. 310–319.
[172] "Steam universe: Emerging technology driving education IT," 2020, Accessed: Apr. 13, 2020. [Online]. Available: https://tinyurl.com/w8cfole
[173] A. Alam, "Should robots replace teachers? Mobilisation of AI and learning analytics in education," in *Proc. Int. Conf. Adv. Comput., Commun., Control*, 2021, pp. 1–12.
[174] K. Ahmad, M. Maabreh, M. Ghaly, K. Khan, J. Qadir, and A. Al-Fuqaha, "Developing future human-centered smart cities: Critical analysis of smart city security, data management, and ethical challenges," *Comput. Sci. Rev.*, vol. 43, 2022, Art. no. 100452.
[175] J. T. Avella, M. Kebritchi, S. G. Nunn, and T. Kanai, "Learning analytics methods, benefits, and challenges in higher education: A systematic literature review," *Online Learn.*, vol. 20, no. 2, pp. 13–29, 2016.
[176] W. Holmes et al., "Ethics of AI in education: Towards a community-wide framework," *Int. J. Artif. Intell. Educ.*, vol. 32, pp. 504–526, 2021.
[177] J. P. Daries et al., "Privacy, anonymity, and Big Data in the social sciences," *Commun. ACM*, vol. 57, no. 9, pp. 56–63, 2014.
[178] N. Garrett, N. Beard, and C. Fiesler, "More than" if time allows" the role of ethics in AI education," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.*, 2020, pp. 272–278.
[179] P. C.-H. Soh, K. W. Chew, K. Y. Koay, and P. H. Ang, "Parents vs peers' influence on teenagers' internet addiction and risky online activities," *Telematics Inform.*, vol. 35, no. 1, pp. 225–236, 2018.
[180] S. Ahmad, S. Umirzakova, G. Mujtaba, M. S. Amin, and T. Whangbo, "Education 5.0: Requirements, enabling technologies, and future directions," 2023, *arXiv:2307.15846*.
[181] B. K. Daniel, "Big data and data science: A critical review of issues for educational research," *Brit. J. Educ. Technol.*, vol. 50, no. 1, pp. 101–113, 2019.
[182] W. Holmes and S. Anastopoulou, "What do students at distance universities think about AI?," in *Proc. 6th ACM Conf. Learn.*, 2019, pp. 1–4.
[183] S. Akgun and C. Greenhow, "Artificial intelligence in education: Addressing ethical challenges in k-12 settings," *AI Ethics*, vol. 2, 431–440, 2021.
[184] D. Mhlanga, "Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning," in *FinTech and Artificial Intelligence for Sustainable Development*. Cham, Switzerland: Palgrave Macmillan, 2023.
[185] P. R. Córdova and R. M. Vicari, "Practical ethical issues for artificial intelligence in education," in *Proc. Technol. Innov. Learn., Teach. Educ.: Third Int. Conf.*, 2023, pp. 437–445.
[186] P. Steinke and P. Fitch, "Minimizing bias when assessing student work," *Res. Pract. Assessment*, vol. 12, pp. 87–95, 2017.
[187] A. B. Arrieta et al., "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, 2020.
[188] "Predicting employment through machine learning," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/ul3lnk8
[189] Y. Xiao and J. Hu, "Assessment of optimal pedagogical factors for Canadian ESL learners' reading literacy through artificial intelligence algorithms," *Int. J. English Linguistics*, vol. 9, no. 4, pp. 1–14, 2019.
[190] "Heartbeat," 2019, Accessed: Dec. 18, 2019. [Online]. Available: https://tinyurl.com/r63t5tr
[191] "Scheduling for learning, not convenience," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/3xxz3nyu
[192] Y. Bao, Y. Peng, C. Wu, and Z. Li, "Online job scheduling in distributed machine learning clusters," in *Proc. IEEE Conf. Comput. Commun.*, 2018, pp. 495–503.
[193] L.-C. Yu et al., "Improving early prediction of academic failure using sentiment analysis on self-evaluated comments," *J. Comput. Assist. Learn.*, vol. 34, no. 4, pp. 358–365, 2018.
[194] M. R. Georgescu and I. A. Bogoslov, "Importance and opportunities of sentiment analysis in developing e-learning systems through social media," in *Proc. Dubrovnik Int. Econ. Meeting*, 2019, pp. 83–93.
[195] K. Ahmad, S. Zohaib, N. Conci, and A. Al-Fuqaha, "Deriving emotions and sentiments from visual content: A disaster analysis use case," 2020, *arXiv:2002.03773*.
[196] J. H. Lim, E. Y. Teh, M. H. Geh, and C. H. Lim, "Automated classroom monitoring with connected visioning system," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, 2017, pp. 386–393.
[197] "9 amazing benefits of technology in the classroom (18 best ways to incorporate technology)," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://www.jenreviews.com/classroom-technology/
[198] "Technology can hurt students' learning, research shows," 2023, Accessed: Jul. 19, 2023. [Online]. Available: https://tinyurl.com/ybwaea86
[199] J. Liu and J. Cohen, "Measuring teaching practices at scale: A novel application of text-as-data methods," *Educ. Eval. Policy Anal.*, vol. 43, pp. 587–614, 2021.
[200] A. Botelho, S. Baral, J. A. Erickson, P. Benachamardi, and N. T. Heffernan, "Leveraging natural language processing to support automated assessment and feedback for student open responses in mathematics," *J. Comput. Assist. Learn.*, vol. 39, no. 3, pp. 823–840, 2023.
[201] B. DeVos, K. M. Nielsen, A. M. Azar, and M Whitaker, "Final report of the federal commission on school safety. presented to the president of the United States," US Department of Education, Washington, D.C., USA, Tech. Rep., 2018.

Authorized licensed use limited to: East China Normal University. Downloaded on September 02,2025 at 15:01:54 UTC from IEEE Xplore. Restrictions apply.