Revista de Psicodidáctica 28 (2023) 93–103

Revista de Psicodidáctica

www.elsevier.es/psicod

Original

Impact of artificial intelligence on assessment methods in primary and secondary education: Systematic literature review

Miguel Martínez-Comesaña [a,*], Xurxo Rigueira-Díaz [b], Ana Larrañaga-Janeiro [c], Javier Martínez-Torres [d,e], Iago Ocarranza-Prado [f], and Denis Kreibel [g]

[a] Department of Mechanical Engineering, Heat Engines and Fluids Mechanics, Industrial Engineering School, CINTECX, University of Vigo (Universidade de Vigo), Maxwell s/n, 36310 Vigo, Spain
[b] CINTECX, GESSMin Group, Department of Natural Resources and Environment Engineering, University of Vigo (Universidade de Vigo), 36310 Vigo, Spain
[c] CINTECX, University of Vigo (Universidade de Vigo), Lagoas-Marcosende s/n, 36310, Vigo, Pontevedra, Spain
[d] Department of Applied Mathematics I, Telecommunications Engineering School, CINTECX, University of Vigo (Universidade de Vigo), 36310 Vigo, Spain
[e] Department of Applied Mathematics I, Telecommunications Engineering School, CITMAga, 15782 Santiago de Compostela, Spain
[f] Possible Incorporate SL, 36211 Vigo, Spain
[g] Smiletronix SL, Pl. d'Urquinaona, 6, 08010 Barcelona, Spain

ARTICLE INFO

Article history:
Available online 15 June 2023

Keywords:
Education
Artificial intelligence
Educational robots
Neural networks
Ubiquitous learning
Predictive analytics

ABSTRACT

The educational sector can be enriched by the incorporation of artificial intelligence (AI) in various aspects. The field of artificial intelligence and its applications in the education sector give rise to a multidisciplinary field that brings together computer science, statistics, psychology and, of course, education. Within this context, this review aimed to synthesise existing research focused on provide improvements on primary/secondary student assessment using some AI tool. Thus, nine original research studies (641 participants), published between 2010 and 2023, met the inclusion criteria defined in this systematic literature review. The main contributions of the application of AI in the assessment of students at these lower educational levels focus on predicting their performance, automating and making evaluations more objective by means of neural networks or natural language processing, the use of educational robots to analyse their learning process, and the detection of specific factors that make classes more attractive. This review shows the possibilities and already existing uses that AI can bring to education, specifically in the evaluation of student performance at the primary and secondary levels.

Impacto de la inteligencia artificial en los métodos de evaluación en la educación primaria y secundaria: revisión sistemática de la literatura

Palabras clave:
Educación
Inteligencia artificial
Robots educativos
Redes neuronales
Aprendizaje ubicuo
Análisis predictivo

RESUMEN

El sector educativo puede enriquecerse con la incorporación de la inteligencia artificial (IA) en diversos aspectos. El campo de la inteligencia artificial y sus aplicaciones en el sector educativo dan lugar a un campo multidisciplinar en el que confluyen la informática, la estadística, la psicología y, por supuesto, la educación. Dentro de este contexto, esta revisión pretende sintetizar las investigaciones existentes centradas en proporcionar mejoras en la evaluación del alumnado de primaria/secundaria utilizando alguna herramienta de IA. Así, nueve estudios de investigación originales (641 participantes), publicados entre 2010 y 2023, cumplen los criterios de inclusión definidos en esta revisión bibliográfica sistemática.

PII of original article:S1136-1034(23)00011-4.

* Corresponding author.
E-mail address: migmartinez@uvigo.es (M. Martínez-Comesaña).

https://doi.org/10.1016/j.psicoe.2023.06.002

Las principales aportaciones de la aplicación de la IA en la evaluación del alumnado de estos niveles educativos inferiores se centran en la predicción de su rendimiento, evaluaciones más objetivas y automatizadas mediante redes neuronales o procesamiento del lenguaje natural, el uso de robots educativos para analizar su proceso de aprendizaje y la detección de factores específicos que hacen más atractivas las clases. Esta revisión muestra las posibilidades y los usos ya existentes que la IA puede aportar a la educación, concretamente en la evaluación del rendimiento del alumnado de primaria y secundaria.

Introduction

This section presents the background and evolution of artificial intelligence (AI), its application in the educational field and the new paths within this merge. In recent years, the field of education has benefited from the advances in artificial intelligence (AI). This progress allows the structure of education to consider human and non-human actors and their respective actions on digital platforms (Bonam et al., 2020). AI has been defined by different authors; Kaelbling and Moore (1996) describes it as the ability of machines to adapt to new situations, solve problems, design plans and perform other functions that require a certain level of intelligence. The capabilities provided by the application of AI have been proven in a variety of scientific fields, such as intelligent buildings (Martínez-Comesaña et al., 2021; Troncoso-Pastoriza et al., 2022), environment (Martínez-Comesaña et al., 2022; Martínez Torres et al., 2020; Rigueira et al., 2022), finance (Jabeur et al., 2021) or chemistry (Anjos et al., 2020). In the education sector, AI has managed to grow substantially due to its algorithmic to make recommendations, predictions, decisions and learn in different contexts (Chen et al., 2022). The introduction of AI in education (AIEd) focuses on making it easier for instructors to perform their tasks more effectively and efficiently. Currently, up to 40% of teaching time is still inverted on activities that could be automated meaning AIEd has plenty of room to grow (Alam, 2021). Overall, AI has the potential to greatly enhance various educational elements or tools, including personalized learning, adaptative assessments, intelligent tutoring systems, automated grading, virtual reality and augmented reality in education, data analysis for performance prediction, language learning. and accessibility and inclusion (Beaulac & Rosenthal, 2019; Xu et al., 2019).

Online education has transformed from a platform where materials were simply downloaded to include intelligent, adaptive systems that adjust based on the actions of learners and instructors to enhance the educational experience (Knox, 2020; Kuleto et al., 2021). Specifically, virtual reality significantly facilitates the learning process beyond the classical learning space, creating global classes and allowing the connection of different students in virtual classes (Bonam et al., 2020; Chen, Xie et al., 2020).

Machine learning, learning analytics (Tlili et al., 2021) and data mining are technologies closely related to education. In this sense, ML can contribute to define recommendations for students (subject or university selection) or help teachers to assess students in a faster and more reliable way (Chen, Chen et al., 2020). In this context, also known as computer-assisted education (CAE), the most widely used techniques are decision trees (Alonso-Fernández et al., 2020), inductive logic programming (Zhang et al., 2021), clustering (Tuyishimire et al., 2022) or neural networks (Kaya, 2019; Okewu et al., 2021). On the other hand, data mining can be considered as the process of pattern discovery and predictive modelling aimed at extracting hidden knowledge.

The use of AI models has had a major impact on education including improvements in efficiency, personalized and global learning, improvements in administration and in the generation of intelligent content (virtual reality, robotics, audio-visual archives or 3-D technology) (Chen, Chen et al., 2020; Chen et al., 2022). This impact can be divided into three different areas: administration, instruction, and learning. In the area of administration, faster task completion and the identification of preferences to create personalized study plans stand out. In terms of instruction, AI-supported learning enables the discovery of potential learning deficiencies to address them as early as possible, interventions tailored to the learner, and the prediction of career paths by studying data (Bonam et al., 2020). Ultimately, in the case of learning, these intelligent education systems (AIEd) are designed to enhance the added value of learning, especially machine learning technologies, which are closely related to statistical modelling and cognitive learning theory (Chen, Chen et al., 2020; Kahraman et al., 2010). Nevertheless, despite all the possibilities that AIEd generates, it is important to continue to study and research ways to implement its use effectively to better support the practice of AIEd (Richardson & Clesham, 2021).

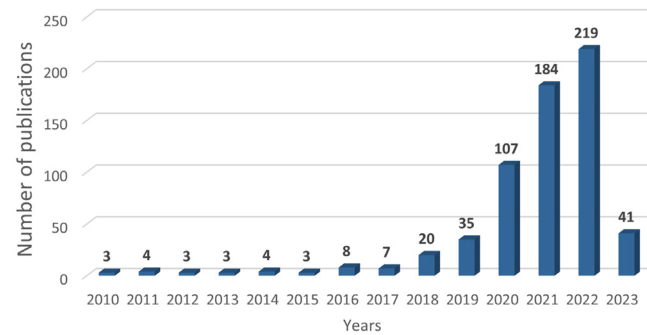

Figure 1. Annual evolution of the number of publications related to AIEd, considering the mentioned criteria search.

Artificial intelligence in education

Over the years, scientific interest in this field has progressively increased. The evolution of AIEd-related publications through time, presented in Figure 1, shows how there is an increasing or upward trend over the years. The search presented in Figure 1 was conducted on 08 March 2023 considering the search criteria presented in the systematic review of the literature shown below. It can be observed that in the last six years, and more intensively in the last three years, the number of publications is increasing a higher rate. In addition, from 2010 to 2017 there was a minimal influence of AIEd in international research. From that year onwards, taking advantage of the rise of machine and deep learning, the interest of the research community in applying these new techniques in education increases. And finally, as a result of the pandemic, this interest undergoes a great increase that is materialized in the number of publications that we observe in Figure 1. The trend, considering that only two months of 2023 have passed, is evidence that interest in the application of AI in education is already considerably and that the publications in the next years continue to rise.

Currently, AI as a field of knowledge linked to computer science is in constant development. Its main objective is the understanding and execution of intelligent tasks such as thinking, acquiring new skills and adapting to new scenarios. Sarker (2022) provides a concise classification of the various AI techniques and divides this technology into four fields. These are introduced below with their most fundamental capabilities:

- Analytical AI: it is oriented to the study and discovery of related events and patterns in the available data. Several machine learning and deep learning models are used, where neural networks are included (Sarker, 2021b, 2021a). In addition, Bayesian models and fuzzy logic are also used for uncertainty quantification (Zadeh, 2008).
- Functional AI: like analytical AI, it studies large amounts of data in order to find relationships and patterns. In this case, instead of making recommendations and presenting the results, it also makes decisions based on the results of the analysis (Aslam et al., 2021; Dowell et al., 2019; Samoilescu et al., 2019).
- Interactive AI: is aimed at automating communication in an efficient and interactive way. There are several examples of this type such as chatbots or personal voice assistants. For the development of these models, several AI techniques are necessary, including heuristic search (Martínez-Tenor et al., 2019; Pivetti et al., 2020).
- Textual AI: comprises the areas of text analysis and natural language processing. This enables text detection, dialog-to-text conversion, machine translations and the ability to generate content (Caratozzolo et al., 2022; Yunanto et al., 2019; Zhang & Zou, 2020).
- Visual AI: capable of recognizing, classifying, and sorting objects from photographs as well as extracting dominant features in video or images converted into text. This type of technology is used in computer vision or augmented reality (Chen et al., 2022).

Each of these types of AI has the ability to provide solutions to real problems and their applications in education are explored in the following sections. Furthermore, the field of artificial intelligence in education (AIEd) encompasses three branches of knowledge: (1) computer science; (2) statistics; and, (3) education. In addition to these three areas, the interdisciplinarity of this field is enriched by contributions from cognitive psychology and neuroscience. As a result of this intersection, there are three subfields that underpin the applications of artificial intelligence in education: (a) data mining for education; (b) learning analytics; and, (c) computer-assisted education.

Data mining applied to education consists of the analysis of educational information through the use of statistical, machine learning and deep learning algorithms (Romero & Ventura, 2010). It focuses on the development of models to understand how students learn and to identify the conditions under which they perform better, as well as to obtain valuable information about the learning phenomenon (Baepler & Murdoch, 2010). Studies in data mining for education include techniques such as statistics or visualization in addition to web data mining involving the use of clustering, classification and deep text mining techniques (Luckin et al., 2016; Romero & Ventura, 2010).

The learning analysis field is defined as the collection, analysis, measurement and presentation of results based on data obtained about students and their context, the main objective being to better understand and optimize learning and the environment in which it occurs (Baepler & Murdoch, 2010; Romero et al., 2013). The techniques most used in learning analytics are statistics, visualization, discourse analysis, social connection analysis and the development of logic models. In addition, learning analytics is more focused on describing data and presenting results, while data mining for education focuses on describing and comparing its various technologies.

Computer-aided education (CAE) is defined as the use of these machines in education to provide assistance and instructions to teachers. In the early stages of its development CAE systems were isolated tools running on computers independently, without AI being able to act on tasks such as student modeling, subject adaptation, or personalization. With the introduction of the Internet, new educational web platforms appeared, and the use of AI was encouraged to achieve more personalized environments for each student on the web and more intelligent at the educational level. Examples of this methods are the intelligent tutoring systems (Mostow & Beck, 2006; Ventura, 2017), the learning management systems (Romero et al., 2008), adaptive multimedia systems (Merceron & Yacef, 2004), examination systems (Romero et al., 2013) and ubiquitous learning environments (Ventura, 2018).

The new phase of the AIEd

The future of education is highly correlated with the future of AI. The increase in the consumption of AI technologies brings with it an increase in the number of people who are developing AI. Thus, innovation and development in this field has never been faster (Luckin et al., 2016). In the following, we will present some developments in AIEd, which have just been incorporated or may be incorporated in the near future, with the aim of improving education:

- It helps students to acquire the so-called *21st century skills* (Van Laar et al., 2017). These skills include communication, collaboration, citizenship, digital literacy, creativity, critical thinking or problem solving and are more related to economic and social developments than skills sought in past years more related to an industrial process. AIEd provides tools for a detailed analysis and evaluation of the development of these activities in students. In addition to changes in the knowledge transmitted, taking into account that AI can be considered as the fourth revolution of the human being, education in the coming years will be introducing changes adapting to this new context (online enrolment and payment, digital books, online exams and classes linking students from all over the world, etc.) (Hans & Crasta, 2019).
- Changes in evaluation. The use of technologies is enabling the collection of big data. In the near future, the sophistication of learning analytics will be complemented by AI techniques to provide just-in-time information. Data from digital teaching and learning experiences provide new insights. These datasets can be analysed not only for correct or incorrect answers but for understanding why the learner arrived at that answer. Furthermore, with new technologies there will be no need for *stop and test*. Instead of classic assessments that are based on a test with a small sample of everything that has been taught, with AIEd assessments will be based on meaningful learning activities (a game or collaborative work) where all learning is analysed (Chassignol et al., 2018).
- AI and AIEd are interdisciplinary fields. AIEd takes advantage of new knowledge in disciplines such as psychology or educational neuroscience to better understand the learning process and to be able to build more accurate models in predicting student progress, motivation or perseverance (Zhang & Aslan, 2021). This requires creating associations that bring together AI developers, educators and student researchers (Luckin & Cukurova, 2019). There are several examples where this synergy is beginning to be exploited, ranging from a platform, known as *CENTURY Tech*, to narrow the performance gap between advantaged and disadvantaged students based on findings in cognitive science and neuroscience (Luckin & Cukurova, 2019), to the knowledge fbthat

learning can be enhanced when connected to an uncertain reward (Luckin et al., 2016).

- Generation of permanent learning partners. These partners can be based on cloud data and be accessible from a multitude of devices. In this way, instead of teaching all possible subjects, the partner can rely on intelligent AIEd systems specialised, or even experts, in the subject needed by the student. These devices can make the learner focus on critical points such as inference or prediction, leaving aside simpler tasks such as calculations or editing. In addition, they can also function as tools to present data in a šmartŵay by helping learners to think deeply and/or find underlying implications in the data (Hwang et al., 2020).

Contribution and organization

The aim of this research is to present and analyse the contributions of AI in education in recent years, showing concrete examples, through a systematic literature review focused on the application of AI to improve the student assessment in primary/secondary levels. In particular, this study is organized as follows: the Materials and Methods is focused on explaining how the search for articles was carried out and the criteria followed, used are presented; the Results and Discussion sections presents and analyses, respectively, the results of the systematic literature review and in the conclusion section the main findings extracted from this research are presented.

Material and method

This research is conducted following the systematic review method to answer specific questions through a replicable process (Gough et al., 2017). This process is defined based on the PRISMA Statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page et al., 2021a, 2021b). Once the study selection process has been carried out, based on specifically defined criteria, the main results of the selected papers are presented and extracted in order to answer the questions defined in the following section regarding the implementation of AI in primary/secondary school student assessment.

Research objective and search

The questions that define the research problem presented in this work are: Are there any studies concerning the application of AI for the assessment of primary/secondary school students? What type of student assessment is based on AI? What are the contributions of these applications? To answer the aforementioned questions, the following objectives are defined: (1) Identification of the main studies focused on the assessment of secondary/primary students with AI tools in recent years (2010–2023), through a systematic review; (2) Analysis of the different forms of educational assessment that are intended to be improved with the application of AI; and (3) Analysis of the actual contributions and improvements provided by the application of AI in the assessment of primary/secondary school students.

Eligibility criteria

This review has focused on research papers describing and introducing the use of AI for students assessment at primary/secondary level. The selection includes research papers published between 2010 and 2023 and considering studies in English. Moreover, once the initial search was conducted (see Search strategy section) the filters used for inclusion/exclusion of studies are presented in Chart 1 .

Chart 1
Criteria for including/excluding research papers

Inclusion	Exclusion
Published between 2010 and 2023	Published before 2010
English language	Not in English
Empirical research	Not empirical such as a review
AI applied for assessing students	Not use of AI
Level primary/secondary	Not use for assessment

Search strategy

The systemic literature review was conducted on 08 March 2023 and following the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement (Page et al., 2021a). Thus, the searches have been carried out from 2010 to 2023 and considering the following databases: ACM Digital Library, Elsevier (ScienceDirect), IEEE Xplore Digital Library, Springer, Taylor and Francis, and Wiley Online Library. These were selected as the most important online scientific libraries with free access. When searching in each of the databases, considering research and conference articles, the following sets of words were taken into account: [*Education* AND *Artificial Intelligence*] OR [*Education* AND *Machine Learning*] OR [*Education* AND *Deep Learning*]. The first search found a total of 659 studies, which after removing duplicates left 641 articles to analyse.

Study selection

The articles found based on the search strategy introduced in the previous section were independently assessed by three reviewers. First, the titles and abstracts were analysed to select the most appropriate ones. Once selected, the remaining articles were evaluated by reading their full text and checking whether they met the criteria of this systematic literature review. In addition, the management of the selected studies, together with the information of those eliminated, was carried out using spreadsheets and the Mendeley manager.

Results

Considering the total number of initial articles reviewed after eliminating duplicates (641 studies), 582 studies were eliminated for not meeting the search requirements after reading their title and abstract. Thus, the remaining 59 articles were analysed through their full text and 48 were excluded for not meeting the aforementioned criteria (see Chart 1) and 1 for not having access to it. The workflow of this process is presented in Figure 2.

Chart 2 shows that in recent years, endorsing Figure 1, the number of publications related to AIEd has increased, although the search has focused on a very specific area (student assessment in primary/secondary school). In addition, the journals where these studies have been published are recognized journals, most of them showing the highest quartiles in the Journal Impact Factor classification (i.e., Q1 or Q2). The most important aspects extracted from the selected studies are analysed below.

Discussion

In this section, we describe the main applications extracted from the studies selected and presented in the previous section based on different fields of study.

Revista de Psicodidáctica 28 (2023) 93–103

Chart 2
Summary of the selected studies, considering the year of publication, their characterization and the Journal Citation Score (JCR) of the journal

Refs.	Year	Objective	Design	Limitations	Results	Level	JCR
(Santos & Boticario, 2014)	2014	Create an AI model to encourage an effective involvement of students and reduce the instructors workload.	Participants fill in questionnaires, are monitored with physiological sensors and their behaviour is recorded.	Students reached, sensors used and tests in presential format.	Sensor information can provide certainty about the actual performance of participants and help to better interpret the data collected (e.g. whether they are actually working on the task).	Secondary	Q2 (2013)
(Wiley et al., 2017)	2017	Capture the causal structure of students explanations to detect their understanding of the topic with Cognitive Science and Artificial Intelligence approaches.	Write an essay. Machine Learning, together with natural language processing, is used for Predicting Explanation Quality and Test Scores.	Difficulties in automatically capturing the quality of explanations, especially in terms of their macrostructure or causal structure	Reasonable application of the coding rules used in the manual coding system. The machine learning scores added 8% to the total explained variance over traditional contributions.	Secondary	Q3 (2021)
(Cruz-Jesus et al., 2020)	2020	Predicting the academic performance of public high school students using artificial intelligence techniques.	Compare a machine learning methods with traditional data analysis. Each input consists of a vector of variables representing a student, and the output indicates whether the student has been promoted to the next grade.	Not clear why one value or the other is predicted. This could be a major obstacle to the wide-scale adoption of AI.	AI techniques have a better accuracy than traditional techniques: RF 87%, ANN 80%, 46% SVM and 49% LR. The most critical variables are the number of unit courses taken, the number of failures and the gender of the student.	Secondary	Q2 (2021)
(Zafari et al., 2021)	2021	This study aims to identify the most relevant factors that affect students' performance by training different machine learning algorithms that classify the students in four categories: very well, good, medium, and bad	Different machine learning algorithms were used as classifiers. Dataset was obtained based on behavioural and individual information and scores were obtained from tests and online questionnaires.	Lack of appropriate educational databases that were not correctly connected. The dataset contains limited information of 459 high school students from different fields that studied in 2020-2021. Individual questionnaires were completed remotely by the students without supervision.	The authors conclude that, in order to improve students' performance, it would be necessary to make more attractive classroom activities. Also, the classroom should be student centred. Moreover, it is recommended to grade the students based on more factors than just grades (such as critical thinking, creativity, etc.). Nevertheless, grades is identify as the most influential variable for students' performance.	Secondary	Q2 (2021)
(Thanh & Tuan, 2021)	2021	Development of an adaptive level AI-based chatbot system to assess students' performance in Mathematics.	An AI chatbot was designed with an API implementing complex features, including Adaptive Testing algorithms. Experiments assessed the feasibility of using the chatbot in math education.	Limited number of topics assessed (1) and reduced number of participants (25).	Results show promising application potential in education.	Secondary	Q1 (2021)

Chart 2 (*Continued*)

Refs.	Year	Objective	Design	Limitations	Results	Level	JCR
(Hsu et al., 2021)	2021	This study aimed to develop an AI instructional tool for young students and used learning analytics to identify sequential learning behaviours.	This research paper conducted a 9-week teaching experiment integrating AI and STEM education with gifted fifth-grade students. It involved individual and cooperative learning, including app development and robot creation, culminating in a learning effectiveness test.	The current study did not fully explore the learning effects of the course due to time constraints and the limited number of research participants. The sample size was also a limitation, with only one girl among the gifted subjects, and no other factors contributing to her outstanding performance were found.	The course design effectively facilitated students' learning of image recognition and machine learning. Encouraging personal comments and prediction improved learning outcomes, emphasizing the importance of teacher-student interactions and planning in interdisciplinary courses.	Primary	Q1 (2021)
(Lamb et al., 2022)	2022	Enhance predictive models of student achievements using brain data from fNIRS in adaptive learning.	fNIRS measured students' cognitive responses across conditions, finding it a robust tool for educational settings. Neurocognitive data predicted science test responses, analysed using Random Forest model and penalized logistic regression and indicators of lost concentration.	The study's limitations include a small sample size, limited cognitive actions coverage, and lack of neurocognitive diversity representation.	Neurocognitive responses during VR and video conditions predicted content test outcomes, while null condition signals did not. Machine learning models accurately predicted correct and incorrect responses, with high success rates ranging from 69% to 85%.	Secondary	Q1 (2021)
(Thomas et al., 2022)	2022	Development of automatic models based on deep learning to predict presentation style from lecture videos and learner engagement from their emotional behaviour.	Extracting visual and verbal information from the slides in the video frames to decide the style (Clustering and classification). Use of a pre-trained model that takes a video clip as input, extracts features and predicts: engaged or distracted.	The study was conducted on a limited sample population ($n = 6$).	The presentation style model performed with an accuracy of 76% and the student engagement model resulted in an accuracy of 95% accuracy at the video level.	Secondary	Q1 (2021)
(Denes, 2023)	2023	Use of a range of AI models to investigate whether AI can be used as an alternative to exam-based grades.	Using detailed information on students past performance, the accuracy of artificial intelligence models in predicting exam grades and differences in accuracy between subjects are investigated.	A single selective school was analysed; the result may not be applicable to other institutions.	The results indicate that, for most students, predictions are accurate (MAE < 1 grade). And subject-dependent; more accurate for STEM subjects and for subjects with more students.	Secondary	Q1 (2021)

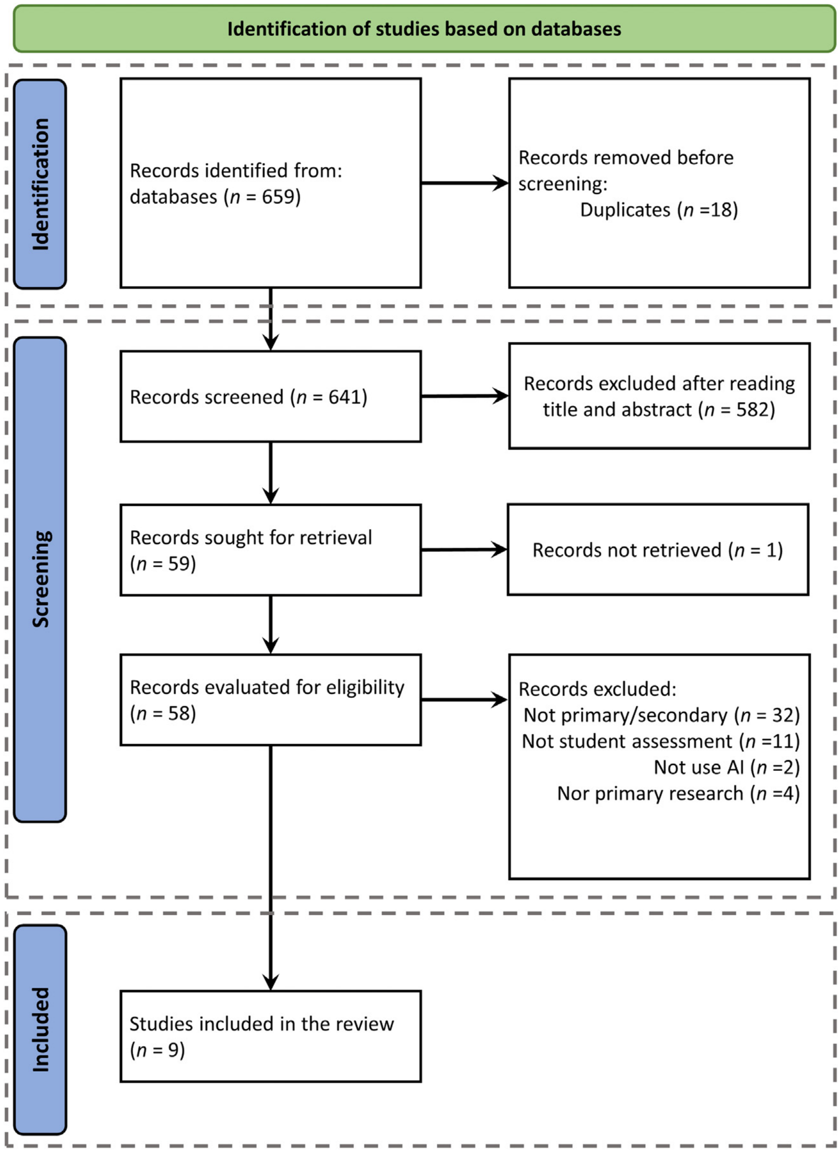

Figure 2. PRISMA diagram.

Natural language processing for language studies

From the analysis of the different studies selected (see Chart 2), it can be observed that natural language processing is a fundamental technique in student assessment. In particular, Wiley et al. (2017) present a methodology for assessing the similarity of student responses, focused on explanations essays, based on idealized target responses by means of Latent Semantic Analysis (LSA). In this way, combining ML with methods of natural language processing (NLP), they manage to improve the variance explained in the evaluation by 8%. Thus, the assessment justification becomes more technical and gains in quality. NLP facilitates vocabulary learning by reducing lexical ambiguity that is achieved by providing dictionary definitions or offering a context to the user for the word in question. Based on the results and conclusion of this study, this technique presents several benefits when it comes to its applications in education, but there are some limitations to consider as well. One limitation is the need for high-quality data to train the algorithms. This means that the data sets used for training should be accurate. Additionally, natural language processing algorithms are not always able to understand the nuances of human language, such as idiomatic expressions, sarcasm, or irony. This can lead to errors in comprehension and feedback. Another limitation is the difficulty in customizing the algorithms for individual learners, as different learners have different learning styles and needs.

Educational robots

Observing the systematic review results (see Chart 2), the educational robots show contribution in student assessment. Hsu et al. (2021) aimed to create an instructional tool for teaching AI to young students, using learning analytics to analyse sequential learning behaviours. The study involved eight gifted fifth-grade students in a nine-week teaching experiment integrating AI and STEM education. The first stage focused on individual learning of MIT App Inventor and Personal Image Classifier, while the second stage involved cooperative learning to create a robot car and play a computational thinking board game. Results showed that the course design effectively taught students image recognition model training and machine learning concepts. Students who expressed personal opinions and sought verification performed better on the learning effectiveness test. Additionally, students who predicted outcomes before executing actions excelled. Limitations included a small sample size and time constraints. Encouraging student expression and predictive thinking can enhance learning outcomes in interdisciplinary, hands-on courses. This study provides insights for improving AI education, but further research is needed with larger samples and additional factors.

In this case, it is shown that the learning based and controlling educational increase the problem-solving skills of students in real time and helps them to better understand theoretical concepts by putting them into practice. Nevertheless, there are also drawbacks to consider. One limitation is the high cost of developing and maintaining robotic systems, which can make them unaffordable for many schools and educational institutions. Another limitation is the need for skilled personnel to operate and maintain the robots, as well as to create and program the educational content. Moreover, educational robots may not be able to provide the same level of individualized instruction as human teachers, as they may lack the ability to adapt to individual learners' needs and learning styles.

Data mining for education: predicting academic performance

Predicting student performance is important for the extraction of behavioural patterns and knowledge about the problems or difficulties they face. The most common application cases are the prediction of academic performance, activity level, knowledge retention, dropout, and early detection of learning problems. Several examples of this application of AIEd can be found throughout the selected studies (see Chart 2). Denes (2023) compares several ML models to predict the grades (based on letter grades) of students, obtaining errors below one grade of distance with the real ones. Cruz-Jesus et al. (2020) also compares several ML techniques (from ANN to Support Vector Machine (SVM)) to predict whether the student will be pass to the next course using as model inputs variables such as year, gender, age, or number of failures. Their results show that some techniques can predict this event with more than 80% of accuracy and the most significant variables were the number of course completed, the number of failures and gender. In addition, Thomas et al. (2022) developed a deep learning model for estimating on the one hand, the presentation style and, on the other hand, the level of engagement in oral presentations. The results obtained show an accuracy of 95% in the best case. From another approach, explained in Santos and Boticario (2014), the performance of students can be assessed through data monitored from sensors and videos, which are fed to a ML model, in order to better understand their real performance (e.g. if they are actually working on the required task at each moment).

Moreover, Lamb et al. (2022) investigated the use of functional near infrared spectroscopy (fNIRS) data to develop accurate predictive models of student achievements in a computer-based learning environment. The study involved 40 ninth-grade students and measured cognitive responses using fNIRS during video, virtual reality (VR), and null conditions. Neurocognitive data from the prefrontal cortex were analysed using statistical methods and ML techniques. The study found that fNIRS data collected during the VR and video conditions predicted correct responses on content tests, while the null condition did not. However, the study is limited by a small sample size and focused on a specific content area and task. Further research is needed to explore diverse samples, different tasks, and the relationship between cognition, affect, behaviour performance, and hemodynamic response.

Lastly, Zafari et al. (2021) developed a ML-based framework capable of evaluating the performance of high school students during one semester in order to identify the most relevant factors that affect their success. They compared different algorithms to check which architecture proved to be the most accurate, choosing neural networks and SVM as the preferred alternative. Authors concluded the classroom needed to be more attractive to students and to grade the students based on more activities than just grades. It is recommended to include in the grading system more activities that include critical thinking and creativity. Nonetheless, this study is highly conditioned by the lack of appropriate educational databases.

In addition to all the possibilities that these applications bring, one limitation is the potential for bias in the data used to train the predictive models, which can lead to inaccurate or unfair predictions. Additionally, data mining models may not be able to account for external factors that can affect student performance, such as family circumstances, health issues, or socio-economic factors.

Dialogue analysis in computer-supported collaborative learning (CSCL)

From the analysis of the presented studies (see Chart 2), it is clear that dialogue analysis is essential for enabling computer-supported join learning because it facilitates the collaborative process and allows for tailored interventions. Thanh and Tuan (2021) describes the development of an AI-based chatbot system called Kant, which uses Adaptive Testing algorithms to assess high school students' mathematical performance. The researchers created a dataset of multiple-choice questions and built an API for the chatbot's implementation. Experiments conducted with a limited number of participants showed promising results, indicating the successful integration of Computerized Adaptive Testing (CAT) and the potential of the chatbot in education. In addition, the union of these tools with time series and semantic similarity analysis allow identifying the moments of best collaboration between participants. Although the study had limitations in terms of the number of topics assessed and participants, this research showcases the application of AI in education and the effectiveness of CAT in assessing students' mathematical achievement.

Based on the application aforementioned, one specific limitation extracted is that dialogue analysis tools are typically based on text data, which may not capture the full complexity of social interactions in CSCL environments. Non-verbal clues such as facial expressions and body language are important elements in face-to-face collaboration but are often lost in digital communication. Furthermore, automated dialogue analysis tools are not always accurate in detecting the intended meaning of students' messages, as they may miss out on nuances of language, sarcasm, or irony.

Neural networks

Regarding the research papers presented in Chart 2, several examples of neural networks applications are identified. Denes (2023) built a Multi-Layer Perceptron (MLP) for classifying grades of students, Thomas et al. (2022) used a pretrained Convolutional Neural Network (CNN) to generate estimations of student style and engagement from presentation videos and in Cruz-Jesus et al. (2020) the efficiency of an artificial neural network for assessing student success in a specific course was compared with others ML models. Neural networks have gained prominence due to their ability to make an objective assessment of the results presented by the student, which avoids possible bias on the part of the teacher. Among the limitations for neural networks is the need for large amounts of data to train the models effectively. This may not always be available, especially in smaller educational settings. Additionally, neural networks can be complex and difficult to interpret, making it challenging to understand how the model arrived at its decisions. This lack of transparency can hinder trust in the model's predictions and recommendations.

Chart 3
Common insights through a SWOT analysis

Analysis SWOT		
Interior	**Strengths** • STEM programs. • Computer-assisted instruction. • Evaluation of learning models. • Programmable agents for natural language learning.	**Weaknesses** • Biased data and algorithms. • Need for large amounts of data.
Exterior	**Opportunities** • Change from an on-stage sage approach to a teacher-guide approach. • Reduction of tedious tasks and routines	**Threats** • Fear of replacement. • Lack of preparation for efficient use of. AIEd.

SWOT analysis

All the studies mentioned in this paper (see Chart 2) have a common idea: exploiting the opportunities offered by the application of artificial intelligence in student assessment. However, the approaches presented in each of them differ in aspects such as the model used, the level of automation or the final objective. In this section they are presented in the form of a scope analysis or SWOT (Strengths, Weaknesses, Opportunities and Threats), the most important common aspects of the above-mentioned publications (see Chart 3). Through this analysis and taking into account the AIEd, strengths and weaknesses (more focused on internal issues and experience) on the one hand, and opportunities and threats (more focused on external aspects and directed towards the future) on the other hand, are identified.

As shown in Chart 3, the implementation of AI in education has greater potential in its strengths and opportunities than its weaknesses and threats. On the one hand, considering an internal approach, the strengths focus on the technological advancement and its exploitation while its weaknesses focus on the lack of data quality or room for improvement in some field of application. On the other hand, considering an external approach, the opportunities show the possible future improvements to be implemented by AIEd and the threats in the adaptability of the teaching staff and students to the mentioned novelties.

Conclusions

In the presented systematic literature review the focus was on the analysis of the application of AI in the assessment of students, specifically at primary/secondary levels, through the collections of articles published, in the most popular databases, from 2010 onwards. Considering the first objective defined in this research we found 641 papers, but after carrying out the selection criteria only nine studies present original applications of AI in student assessment at the mentioned levels. On the one hand, the main conclusion of this research is that, despite the complexity of AI, this systematic research shows the potential of AI-related tools to improve education, in particular student assessment, at lower levels such as primary or secondary. Through the nine selected studies, different models and application of AIEd have been analysed. Answering the research objective two, the main fields where AI applications have been found are the use of educational robots for improving and qualify students learning, the prediction of students performance to anticipate and try to redirect their path, the use of different AI techniques such as NLP or NN to improve the quality of the evaluation or even remove repetitive tasks to teachers. In this way, this research contributes with guidance in the implementation of AIEd for student assessment in primary/secondary education levels. In addition, in response to the third research objective, the main improvements brought by AIEd and reached with this review are more accurate predictions of student performance, more automatic and objective evaluation of student tasks (such as when they are more collaborative), and the detection of significant factors related to classes that make them attractive to students. On the other hand, the main limitation of this research was the specific area of application chosen because, so far, most of the AIEd implementations are focused on university or postdoctoral levels. However, the studies found show the great impact at all levels of AI in education. In short, this systematic literature review shows the influence of AIEd at lower levels of education, the existing research interest in this field and the real and ongoing improvements of using AI tool to improve the student assessment.

Conflicts of interests

The authors declare that they have no conflicts of interest.

Acknowledgements and funding

This research has been supported by the Ministry of Science, Innovation and Universities of the Spanish Government through grant FPU19/01187 (Miguel Martínez). The work of Ana Larrañaga has been supported by the predoctoral grant 2020 of the University of Vigo. Funding for the open access position: Universidade de Vigo/CISUG.

References

Alam, A. (2021). Possibilities and Apprehensions in the Landscape of Artificial Intelligence in Education. *2021 International Conference on Computational Intelligence and Computing Applications (ICCICA)*, 1–8. https://doi.org/10.1109/ICCICA52458.2021.9697272

Alonso-Fernández, C., Martínez-Ortiz, I., Caballero, R., Freire, M., & Fernández-Manjón, B. (2020). Predicting students' knowledge after playing a serious game based on learning analytics data: A case study. *Journal of Computer Assisted Learning*, *36*(3), 350–358. https://doi.org/10.1111/jcal.12405

Anjos, O., Martinez, M., Caldeira, I., Pedro, S., Eguía, P., & Canas, S. (2020). Application of functional data analysis and FTIR-ATR spectroscopy to discriminate wine spirits ageing technologies. *Mathematics*, *8*, 896. https://doi.org/10.3390/math8060896

Aslam, N., Khan, I., Alamri, L., & Almuslim, R. (2021). An improved early student's academic performance prediction using deep learning. *International Journal of Emerging Technologies in Learning (IJET)*, *16*(12), 108–122. https://www.learntechlib.org/p/220061

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, *4*(2), 108–126. https://doi.org/10.20429/ijsotl.2010.040217

Beaulac, C., & Rosenthal, J. S. (2019). Predicting university students' academic success and major using random forests. *Research in Higher Education*, *60*(7), 1048–1064. https://doi.org/10.1007/s11162-019-09546-y

Bonam, B., Piazentin, L., & Possa, A. D. (2020). Educación, big data e inteligencia artificial: metodologías mixtas en plataformas digitales. *Comunicar: Revista Científica Iberoamericana de Comunicación y Educación*, *28*(65), 43–52.

Caratozzolo, P., Rodriguez-Ruiz, J., & Alvarez-Delgado, A. (2022). Natural language processing for learning assessment in STEM. *2022 IEEE Global Engineering Education Conference (EDUCON)*, 1549–1554. https://doi.org/10.1109/EDUCON52537.2022.9766717

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, *25*(1), 28–47.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*(August), Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Cruz-Jesus, F., Castelli, M., Oliveira, T., Mendes, R., Nunes, C., Sa-Velho, M., & Rosa-Louro, A. (2020). Using artificial intelligence methods to assess academic achievement in public high schools of a European Union country. *Heliyon*, *6*(6), e04081. https://doi.org/10.1016/j.heliyon.2020.e04081

Denes, G. (2023). A case study of using AI for General Certificate of Secondary Education (GCSE) grade prediction in a selective independent school in England. *Computers and Education: Artificial Intelligence*, *4*, Article 100129. https://doi.org/10.1016/j.caeai.2023.100129

Dowell, N., Lin, Y., Godfrey, A., & Brooks, C. (2019). Promoting inclusivity through time-dynamic discourse analysis in digitally-mediated collaborative learning. *International Conference on Artificial Intelligence in Education*, 207–219. https://doi.org/10.1007/978-3-030-23204-7_18

Gough, D., Thomas, J., & Oliver, S. (2017). *An introduction to systematic reviews*. SAGE Publications Ltd. 1-352. http://digital.casalini.it/9781473968219

Hans, V. B., & Crasta, S. J. (2019). Digitalization in the 21st century: Impact on learning and doing. *Journal of Global Economy*, *15*(1), 12–23.

Hsu, T.-C., Abelson, H., Lao, N., Tseng, Y.-H., & Lin, Y.-T. (2021). Behavioral-pattern exploration and development of an instructional tool for young children to learn AI. *Computers and Education: Artificial Intelligence*, *2*, Article 100012. https://doi.org/10.1016/j.caeai.2021.100012

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *1*, 1–5. https://doi.org/10.1016/j.caeai.2020.100001

Jabeur, S. B. en, Gharib, C., Mefteh-Wali, S., & Arfi, W. Ben. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166*, Article 120658. https://doi.org/10.1016/j.techfore.2021.120658

Kaelbling, L. P., & Moore, A. W. (1996). Reinforcement learning: A survey. *The Journal of Artificial Intelligence Research*, *4*, 237–285.

Kahraman, H. T., Sagiroglu, S., & Colak, I. (2010). Development of adaptive and intelligent web-based educational systems. *2010 4th International Conference on Application of Information and Communication Technologies*, 1–5.

Kaya, I. E. (2019). Artificial neural networks as a decision support tool in curriculum development. *International Journal on Artificial Intelligence Tools*, *28*(04), Article 1940004. https://doi.org/10.1142/S0218213019400049

Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology*, *45*(3), 298–311. https://doi.org/10.1080/17439884.2020.1754236

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M. D., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, *13*(18), Article 10424. https://doi.org/10.3390/su131810424

Lamb, R., Neumann, K., & Linder, K. A. (2022). Real-time prediction of science student learning outcomes using machine learning classification of hemodynamics during virtual reality and online learning sessions. *Computers and Education: Artificial Intelligence*, *3*, Article 100078. https://doi.org/10.1016/j.caeai.2022.100078

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, *50*(6), 2824–2838. https://doi.org/10.1111/bjet.12861

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in intelligence unleashed.*

Martínez Torres, J., Pastor Pérez, J., Sancho Val, J., McNabola, A., Martínez Comesaña, M., & Gallagher, J. (2020). A functional data analysis approach for the detection of air pollution episodes and outliers: A case study in Dublin, Ireland. *Mathematics*, *8*(2) https://doi.org/10.3390/math8020225

Martínez-Comesaña, M., Eguía-Oller, P., Martínez-Torres, J., Febrero-Garrido, L., & Granada-Álvarez, E. (2022). Optimisation of thermal comfort and indoor air quality estimations applied to in-use buildings combining NSGA-III and XGBoost. *Sustainable Cities and Society*, *80*, Article 103723. https://doi.org/10.1016/j.scs.2022.103723

Martínez-Comesaña, M., Ogando-Martínez, A., Troncoso-Pastoriza, F., López-Gómez, J., Febrero-Garrido, L., & Granada-Álvarez, E. (2021). Use of optimised MLP neural networks for spatiotemporal estimation of indoor environmental conditions of existing buildings. *Building and Environment*, *205*, Article 108243. https://doi.org/10.1016/j.buildenv.2021.108243

Martínez-Tenor, Á., Cruz-Martín, A., & Fernández-Madrigal, J. A. (2019). Teaching machine learning in robotics interactively: The case of reinforcement learning with Lego ® Mindstorms. *Interactive Learning Environments*, *27*(3), 293–306. https://doi.org/10.1080/10494820.2018.1525411

Merceron, A., & Yacef, K. (2004). Mining student data captured from a web-based tutoring tool: Initial exploration and results. *Journal of Interactive Learning Research*, *15*(4), 319–346.

Mostow, J., & Beck, J. (2006). Some useful tactics to modify, map and mine data from intelligent tutors. *Natural Language Engineering*, *12*(2), 195–208. https://doi.org/10.1017/S1351324906004153

Okewu, E., Adewole, P., Misra, S., Maskeliunas, R., & Damasevicius, R. (2021). Artificial neural networks for educational data mining in higher education: A systematic literature review. *Applied Artificial Intelligence*, *35*(13), 983–1021. https://doi.org/10.1080/08839514.2021.1922847

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021a). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología (English Edition)*, *74*(9), 790–799. https://doi.org/10.1016/j.rec.2021.07.010

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021b). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, Article 105906. https://doi.org/10.1136/bmj.n71

Pivetti, M., Di Battista, S., Agatolio, F., Simaku, B., Moro, M., & Menegatti, E. (2020). Educational robotics for children with neurodevelopmental disorders: A systematic review. *Heliyon*, *6*(10), e05160. https://doi.org/10.1016/j.heliyon.2020.e05160

Richardson, M., & Clesham, R. (2021). Rise of the machines? The evolving role of Artificial Intelligence (AI) technologies in high stakes assessment. *London Review of Education*, *19*(1), 1–13.

Rigueira, X., Araújo, M., Martínez, J., García-Nieto, P. J., & Ocarranza, I. (2022). Functional data analysis for the detection of outliers and study of the effects of the COVID-19 pandemic on air quality: A case study in Gijón, Spain. *Mathematics*, *10*(14), 2374. https://doi.org/10.3390/math10142374

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, *40*(6), 601–618. https://doi.org/10.1109/TSMCC.2010.2053532

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers and Education*, *51*(1), 368–384. https://doi.org/10.1016/j.compedu.2007.05.016

Romero, C., Zafra, A., Luna, J. M., & Ventura, S. (2013). Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data. *Expert Systems*, *30*(2), 162–172. https://doi.org/10.1111/j.1468-0394.2012.00627.x

Samoilescu, R., Dascalu, M., Sirbu, M., Trausan-Matu, S., & Crossley, S. A. (2019). Modeling collaboration in online conversations using time series analysis and dialogism. *International Conference on Artificial Intelligence in Education*, 458–468. https://doi.org/10.1007/978-3-030-23204-7_38

Santos, O. C., & Boticario, J. G. (2014). Involving users to improve the collaborative logical framework. *The Scientific World Journal*, *2014*, Article 893525. https://doi.org/10.1155/2014/893525

Sarker, I. H. (2022). AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, *3*(2), 1–20. https://doi.org/10.1007/s42979-022-01043-x

Sarker, I. H. (2021a). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, *2*(6), 1–20. https://doi.org/10.1007/s42979-021-00815-1

Sarker, I. H. (2021b). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, *2*(3), 1–21. https://doi.org/10.1007/s42979-021-00592-x

Thanh, N. H., & Tuan, B. A. (2021). We are using artificial intelligence in assessing students' achievement at high schools: A case study in mathematics. *Proceedings of the 5th Asia Pacific International Modern Sciences Congress*, 139–149.

Thomas, C., Puneeth Sarma, K. A. V., Swaroop Gajula, S., & Jayagopi, D. B. (2022). Automatic prediction of presentation style and student engagement from videos. *Computers and Education: Artificial Intelligence*, *3*, Article 100079. https://doi.org/10.1016/j.caeai.2022.100079

Tlili, A., Chang, M., Moon, J., Liu, Z., Burgos, D., Chen, N. S., & Kinshuk. (2021). A systematic literature review of empirical studies on learning analytics in educational games. *International Journal of Interactive Multimedia and Artificial Intelligence*, *7*(2), 250–261. https://doi.org/10.9781/ijimai.2021.03.003

Troncoso-Pastoriza, F., Martínez-Comesaña, M., Ogando-Martínez, A., López-Gómez, J., Eguía-Oller, P., & Febrero-Garrido, L. (2022). IoT-based platform for automated IEQ spatio-temporal analysis in buildings using machine learning techniques. *Automation in Construction*, *139*, Article 104261.

Tuyishimire, E., Mabuto, W., Gatabazi, P., & Bayisingize, S. (2022). Detecting learning patterns in tertiary education using K-means clustering. *Information*, *13*(2), 94. https://doi.org/10.3390/info13020094

Van Laar, E., Van Deursen, A. J. A. M., Van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, *72*, 577–588. https://doi.org/10.1016/j.chb.2017.03.010

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

Ventura, M. Della. (2018). Twitter as a music education tool to enhance the learning process: Conversation analysis. In L. Deng, W. W. K. Ma, & C. W. R. Fong (Eds.), *New media for educational change* (pp. 81–88). Singapore: Springer.

Wiley, J., Hastings, P., Blaum, D., Jaeger, A. J., Hughes, S., Wallace, P., Griffin, T. D., & Britt, M. A. (2017). Different approaches to assessing the quality of explanations following a multiple-document inquiry activity in science. *International Journal of Artificial Intelligence in Education*, *27*(4), 758–790. https://doi.org/10.1007/s40593-017-0138-z

Xu, X., Wang, J., Peng, H., & Wu, R. (2019). Prediction of academic performance associated with internet usage behaviors using machine learning algorithms. *Computers in Human Behavior*, *98*, 166–173. https://doi.org/10.1016/j.chb.2019.04.015

Yunanto, A. A., Herumurti, D., Rochimah, S., & Kuswardayan, I. (2019). English education game using non-player character based on natural language processing. *Procedia Computer Science*, *161*, 502–508. https://doi.org/10.1016/j.procs.2019.11.158

Zadeh, L. A. (2008). Is there a need for fuzzy logic? *Information Sciences*, *178*(13), 2751–2779. https://doi.org/10.1016/j.ins.2008.02.012
Zafari, M., Sadeghi-Niaraki, A., Choi, S.-M., & Esmaeily, A. (2021). A practical model for the evaluation of high school student performance based on machine learning. *Applied Sciences*, *11*(23). https://doi.org/10.3390/app112311534
Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025
Zhang, Z., Yilmaz, L., & Liu, B. (2021). *A critical review of inductive logic programming techniques for explainable AI*. https://doi.org/10.48550/ARXIV.2112.15319. 1-17, arXiv
Zhang, R., & Zou, D. (2020). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 1–47. https://doi.org/10.1080/09588221.2020.1744666