STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
*Stat., Optim. Inf. Comput.*, Vol. **x**, Month 202x, pp 0–17.
Published online in International Academic Press (www.IAPress.org)

# Machine Learning in Education: An Overview of Applications and Datasets

Amal Azeroual [1,2,*], Naoual Attaoui [3], Mourad Azhari [1,2], Soufiane El Adraoui [4], Benayad Nsiri [5], Mohamed Essaaidi [6]

[1] *Center of Guidance and Planning, Rabat, Morocco*
[2] *faculty of Sciences, Mohammed V University Rabat, Morocco*
[3] *TIMS Team, Faculty of Sciences, Abdelmalek Essaadi University, Tetouan, Morocco*
[4] *ISD Laboratory, Faculty of Sciences, Abdelmalek Essaadi University, Tetouan, Morocco*
[5] *ENSAM, Mohammed V University,Rabat, Morocco*
[6] *ENSIAS, Mohammed V University,Rabat, Morocco*

**Abstract** Artificial intelligence (AI) is a field of research that has attracted a lot of attention in the past few years due to its rapid development and global reach. From integrating AI into everyday life, technology, and its use has drastically changed along with the understanding of the world surrounding us. Its application ranges from automating mundane tasks to making crucial AI-driven decisions in medicine, entertainment, finance, and even retail. One of the most positive and profound areas of development in AI is in education because if developed further, it can change the ways education is rendered, enhance the quality of education, and most importantly, tailor the learning process according to the needs of students all over the world. AI is not a fad when it comes to educational learning but a very essential shift regarding how information is taught and learned. The goal of this paper is to analyze several applications of Machine Learning (ML), a subfield of artificial intelligence, in the context of education. In particular, we want to focus on the variety of ML by defining supervised learning, unsupervised learning, reinforcement learning, and also deep learning, and their importance to the improved practice of education. These techniques will be further explored in terms of real-life application in teaching and learning processes through illustrative case studies. These case studies will illustrate how ML models aid in customizing learning, increasing student participation, and enhancing collaboration. We also take a closer look at the educational datasets most commonly cited in the literature, together with the evaluation metrics that researchers rely on when developing and validating ML-based learning systems. Beyond reviewing current applications,, this paper also considers the broader technical, pedagogical, ethical, and organizational questions that shape how ML can realistically be integrated into educational contexts. In addition, we draw attention to several emerging directions in the field, such as multimodal learning analytics, explainable AI, and hybrid approaches that bring together human and machine intelligence. By combining a synthesis of established ML methods with a discussion of their limitations and the avenues they open for future work, This paper aims to provide a realistic and balanced perspective on the actual contributions and limitations of machine learning in the evolution of education.

**Keywords** Education, Artificial Intelligence, Machine Learning, Deep Learning, Educational Datasets

**DOI:** 10.19139/soic-2310-5070-2568

## 1. Introduction

Currently, AI is changing how things are done in different industries by creating new ways to automate difficult tasks and enhance human capabilities through data analysis innovation. In the education field, students' use of educational platforms produces massive data sets [1]. This data can range from students' academic performance to how they engage with various learning technologies. As it stands now, AI is becoming increasingly important in processing such data, and other approaches via which refined solutions to meet the student's learning requirements

*Correspondence to: Amal Azeroual(Email: amal.azeroual@gmail.com). Center of Guidance and Planning, AV Ezaitoune, Errayad, BP 6222, Rabat, Morocco.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

can be devised effectively and efficiently.It also supports teachers in improving their pedagogical practices. AI have also the ability to automate administrative tasks, thereby fostering more efficient and effective educational environments [2].

ML, a key subfield of AI, is particularly well-suited to take advantage of these datasets. It uses algorithms qualified to learn from the data. This article aims to examine the potential of ML in the field of education across its different categories: supervised, unsupervised, reinforcement, and deep learning. In a non-exhaustive manner, we illustrate how these models are transforming the educational experience. The applications are diverse, ranging from predicting student success (supervised learning) to identifying similar learning styles (unsupervised learning), creating interactive tutoring systems that adapt to students (reinforcement learning), and analyzing educational content (deep learning). We also dive into popular datasets that support ML research in education, explaining their strengths and limitations, and looking at the metrics that help us measure how well these ML models perform in real-world settings.

Beyond presenting a survey of current applications, this article also takes a step back to consider the technical, pedagogical, ethical, and organizational challenges that influence how ML can realistically be integrated into educational environments. Questions related to data quality, algorithmic bias, model transparency, teacher training, and infrastructure capacity still represent major hurdles for an adoption that is both dependable and fair. We also highlight a number of emerging trends — such as multimodal learning analytics, explainable AI, and hybrid frameworks that bring together human and machine intelligence — which are beginning to reshape research priorities and anticipated directions in the field. The core contribution of this review is to provide a perspective that is both coherent and nuanced. It does so by bringing together: (1) an overview of ML techniques and the ways in which they are used in educational contexts; (2) a contextualized synthesis of the datasets and evaluation metrics commonly employed in ML-based research; (3) a critical examination of the challenges and limitations that the field continues to face; and (4) a reflection on the research paths that are now emerging for the years ahead. By doing so, this work aims to enrich current discussions on how ML can meaningfully contribute to transforming educational practices.

## 2. Machine Learning Models in Education

Based on their learning approach and the type of problems they aim to solve; we can generally classify ML algorithms into different categories. Supervised learning, unsupervised learning, reinforcement learning, and deep learning are some of the several types of ML algorithms that we use in this overview. In the ongoing section, we illustrate the applications of ML in education according to each category with studies found in the literature (see Figure 1).

### *2.1. Supervised learning*

Supervised learning algorithms are trained on a labelled dataset, where the input data matches the target output. These algorithms aim to learn a mapping function that can correctly predict the output for new data. For this type of algorithm, its applications in education are diverse. Student performance prediction, grading, and assessment are the areas that have attracted the most researchers.

#### *2.1.1. Student Performance Prediction*

Student success prediction is one of the most widespread applications of MLin the educational sector, as attested to by the volume of research studies available in scientific databases [3]. Instructors use predictive analytics to identify students with a high risk of failure or withdrawal. MLcan help identify students who are on the verge of having academic trouble and intervene promptly by analyzing patterns of attendance, assignment submission, and testing [4]. Besides improving the retention aspect of learning, this proactive method boosts overall academic performance, as the problems are addressed well in advance before things take a turn for the worse [5].

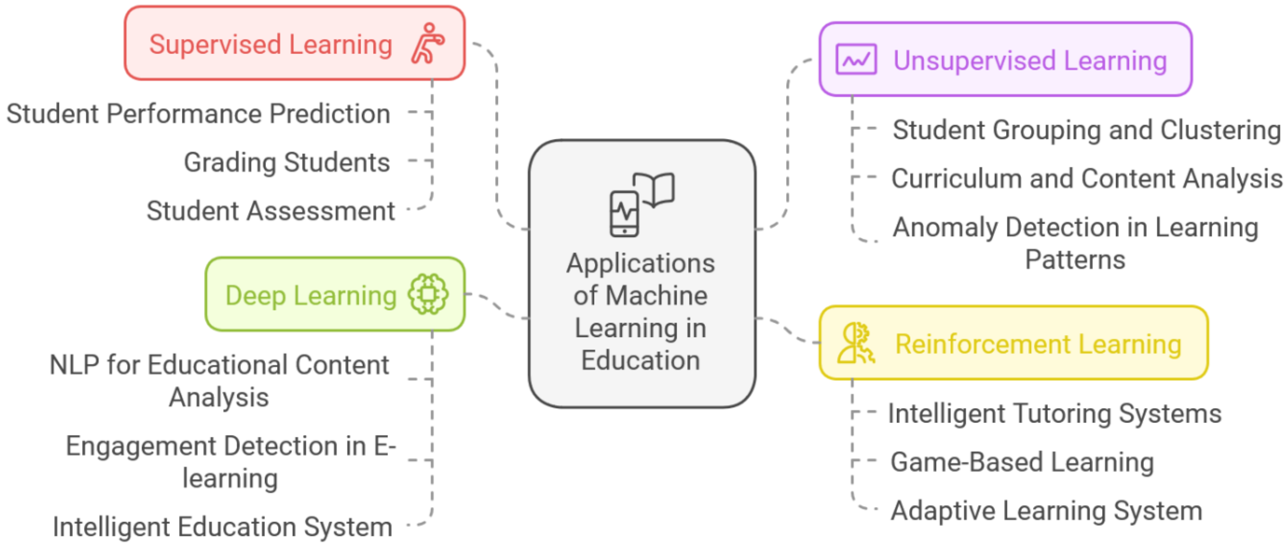


Figure 1. Applications of ML in Education.

A study [6] was conducted on the academic achievement grades of 1854 students who took the Turkish Language-I course at a public university in Turkey during the fall semester of 2019–2020. To forecast the final exam grades of undergraduate students, the study suggested the use of ML algorithms such as Random Forests (RF), nearest neighbor (NN), support vector machines (SVM), Logistic Regression (LR), Naïve Bayes (NB), and k-nearest neighbors (KNN). Three factors were used to compute and compare the algorithms' performances: faculty data, department data, and midterm test grades. According to the findings, the suggested model's classification accuracy ranged from 70 to 75%.

To predict success or failure in mathematics tests and to anticipate necessary improvements in learning, the authors in [7] developed predictive models based on five classifiers, such as LR, KNN, SVM, decision trees (DT), and RF. The study was applied to trends in international mathematics and science study (TIMSS) 2019 data from Morocco, taking into account a variety of student characteristics, such as academic and personal performance.

Another research [8] employed an LR model to predict students' academic performance and examine the factors influencing this performance. This research uses data from a survey conducted among secondary school students in math courses, consisting of a total of 33 indicators. The dataset used is publicly available on Kaggle under the name "Student Performance Dataset." The results indicate that the LR model is reliable and demonstrates accurate prediction (pass or fail), achieving 95.8% accuracy, 96.7% precision, 95.1% recall, and an F1 score of 95.8%. According to the authors, this research enables schools to identify students' learning difficulties and implement supportive measures to help them.

The performance of students who passed the high school entrance exam in 81 provinces of Turkey in 2019 was examined based on various socio-economic factors [9]. Support vector regression, RF, DT, and beta regression models were used to predict high school entrance examination success rates. According to the study, it is recommended to use both beta regression and RF models simultaneously to assess students' success rates.

### *2.1.2. Grading students*

ML also streamlines administrative tasks, reducing the burden on educators and allowing them to focus more on teaching. Automated grading systems, for example, can handle objective assessments efficiently, providing instant feedback to students and freeing up instructors to concentrate on more complex, subjective evaluations [10].

To optimize the overall performance of assessment methods and facilitate the grading process, the authors in [11] implemented various ML algorithms to automatically grade students' assignments. The research explored programming exercises and written responses. The ML algorithms used include NB, DT, RF, SVM, LR, KNN, and Ensemble Methods (EMs).

These pertinent features were extracted from the assignments, and models were trained on a varied set of examples. Research [12] in this same area proposed a very effective, fair, and intelligent grading system for evaluating English compositions, based on an improved version of the Adaboost algorithm. Results indicate that the new AdaBoost/CT algorithm offers a highly accurate grading of English compositions and allows pinpointing the weaknesses of each student.

#### *2.1.3. Student assessment*

Many studies have provided evidence supporting the possibility of automatic knowledge assessment. For example, a study [13] addressed the assessment of student performance via machine-learning approaches. Based on a dataset from an online learning environment, the study developed an evaluation model and identified the most relevant features for knowledge classification by using classifiers such as SVM, RF, DT, gradient boosting machine (GBM), linear regression, Gaussian Nave Bayes (GNB), and multilayer perceptron (MLP). The results indicate that the GBM classifier outperformed the classifiers with 98% accuracy, 99% prediction precision, and 97% recall.

### *2.2. Unsupervised Learning*

Unsupervised learning algorithms have significant applications in education, particularly for clustering. These methods enable educators to discover characteristics from unlabeled data without any prior knowledge of the expected output. We outline below the key applications of unsupervised learning in educational contexts.

#### *2.2.1. Student Grouping and Clustering*

An approach based on unsupervised clustering was proposed to help the Malaysian government reduce dropout rates among students from the lowest income class who have access to higher education institutions [14]. In this approach, three specific types of ML models (K-Means, Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN)) were used to analyze and classify the academic performance of students. The dataset used in this study contains information on $117,069$ undergraduate students, with $16$ attributes from $20$ public higher education institutions, including continuous variables (grades, accumulated credits) and categorical variables (discipline, program). The observed results showed that the optimized k-means, with K-Means++ initialization, achieved the best performance in terms of intra-cluster cohesion and inter-cluster separation, and produced five groups of students based on their academic performance.

In the context of health education, some researchers [15] used K-means and a Hierarchical Clustering Algorithm (HCA) to classify $310$ students into different categories according to the Felder-Silverman learning styles model. The experiments identified three significant clusters, validated using clustering quality indices such as the silhouette score and the Davies–Bouldin index. The findings indicate that students can be partitioned into a minimum of three groups in order to use educational methodologies that meet the needs of each cluster. According to an investigation [16], data clustering was used to assess student performance. In this research, the speed at which a user can process information was measured based on a comprehension text and a set of $9$ comprehension questions. The K-means algorithm was then applied to classify students into their respective groups. Intra-cluster distances were minimized using the standard K-Means cost function, and the classification allowed the identification of distinct performance profiles. In light of this classification, students' performance can be enhanced. An analysis [17] seeks to determine regions in Morocco where students are most vulnerable and in need of support.

The students were grouped based on their social vulnerability using an unsupervised competitive learning method based on K-Means algorithm and called the ”Centroid neural network”. This approach relies on an iterative update of the centroids following Hebb’s rule, while normalizing the input vectors. The goal is to ensure the convergence and stability of the clusters. The researchers used the PISA dataset, which provides detailed information about each student, including their parents’ background, socioeconomic status, and school conditions. The results show that social vulnerability negatively affects students’ cognitive development and ’Beni Mellal-Kh ´ enifra’ has the highest level of social vulnerability. ´

#### 2.2.2. *Curriculum and Content Analysis*

Analysis of curricula and educational content involves systematically examining programs, learning objectives, and teaching materials to identify trends, gaps, and opportunities for improvement. Today, ML enables the automation of this analysis at large scale and enhances understanding of the structure and themes of educational programs. Study [18] provides a concrete example of this approach. The authors reviewed 166 AI ethics programs from 105 universities worldwide. Their goal was to identify the major themes shaping these programs while also shedding light on disciplinary silos. To do this, they used an artificial intelligence technique called Latent Dirichlet Allocation (LDA), which can automatically detect recurring topics in course syllabi. Through this analysis, they were able to classify the content by discipline (computer science, humanities, law) and by cognitive level, based on Bloom’s taxonomy. They also compared the stated learning objectives with the teaching methods used. Their findings revealed a three-part model called BAG (Build, Assess, Govern): technological design (Build), impact and risk assessment (Assess), and governance and regulation (Govern). The study also highlighted notable differences across disciplines and regions, as well as frequent gaps between what courses aim to teach and how they are actually taught.

#### 2.2.3. *Anomaly Detection in Learning Patterns*

Anomaly detection is considered one of the interesting applications of ML in the field of education. For example, Demidova et al. [19] studied student behavior in an automated Python course using the Digital Teaching Assistant (DTA) system. This system records detailed information about each submission, such as accuracy, time spent, and strategies used. The researchers’ goal was to identify unusual behaviors and distinguish genuinely motivated students from those whose patterns might suggest cheating or a heavy reliance on external solutions. To analyze the programs, each solution was translated into a Markov chain, which represents the transitions between different code structures. The researchers then measured similarities between solutions using the Jensen–Shannon distance. Hierarchical agglomerative clustering with average linkage was applied to group similar programs. This process made it possible to distinguish between two profiles: motivated students, who experimented with different approaches, and suspicious students, whose work appeared repetitive or nearly identical. The study also examined submission timelines, revealing unusual behaviors such as mass submissions or long periods of inactivity. By combining these techniques—vectorization, similarity measurement, clustering, and temporal analysis—the authors developed a multidimensional view of student behavior. These findings provide instructors with tools to detect potential cheating early, deliver targeted support, and improve the design of programming exercises in automated learning environments.

### 2.3. ***Reinforcement Learning (LR)***

RL algorithms learn by interacting with an environment and receiving responses as rewards or punishments. RL approaches using Q-learning and its variants have demonstrated viability for many years.

#### 2.3.1. *Intelligent Tutoring Systems*

One study [20] examined how intelligent tutoring (ITS) systems might be modified to support autistic students who frequently encounter particular learning and communication issues. Instead of using rigid methods based on established rules, they have adopted learning by RL, an approach that allows the system to automatically modify its educational strategies through trial and error based on the learner's responses. To put this idea into practice, the researchers modeled the student using an artificial neural network (ANN). The system tested different learning scenarios employing both the "typical" learner and the autistic learner models. The ITS stored the student's past responses. This helped determine the student's current learning state. Then, it selected teaching actions using a modified version of Q-learning. When tested on these simulated students, the autistic model progressed more slowly and needed more practice, but it was still able to achieve performance levels similar to the typical model. The study demonstrates the promise of RL in making tutoring systems more flexible and personalized, especially for students with diverse learning needs. At the same time, the work is limited by the fact that it was tested only in simulations, with simplified variables and a narrow definition of success focused on immediate performance.

#### *2.3.2. Game-Based Learning*

Gamification is widely used to make learning more engaging, but traditional mechanisms (points, badges, or leaderboards) often focus on immediate rewards, without ensuring that learners actually develop lasting skills. One study [21] addresses this challenge by using RL to design rewards that not only motivate in the short term but also guide learners toward long-term educational goals. In their approach, learning is modeled as a Markov decision process. Each state represents the learner's current skill profile, each action corresponds to a possible learning activity, and rewards are dynamically adjusted based on their expected impact on acquiring target skills. This adaptive "reward shaping" turns rewards into real strategic guides. It encourages immediate motivation. At the same time, it directs learning toward coherent and sustainable skill development. Simulations conducted by the authors show that this method works better than simple gamification strategies, which often favor fun but poorly structured activities.

#### *2.3.3. Adaptive learning system*

Personalized learning is a common application of MLwithin education. Typical teaching systems mostly contend with the fact that everyone inside them learns in the same mold ignoring the specific needs of individual students. ML recognizes the data from performance metrics, level of engagement, and personal learning preference of a student and retrospects for the possible adaptation of behavioral and learning activities to individual students [22]. Adaptive learning systems like DreamBox and Knewton for example, apply MLalgorithms to change the difficulty of exercises during actual learning experiences so that students are continuously challenged yet capable of reaching their respective levels of mastery [23].

Analyzing the effect of RL-supported adaptive learning systems on learners' performance was the aim of another work [24]. In this way, the personalized adaptive knowledge extraction strategy (PAKES) was proposed, indicating personalized pedagogical sequences depending on the specific needs of each learner. It includes monitoring the knowledge states of the learners with a cognitive diagnostic model and conducting the Q-learning algorithm to optimize the learning path. Experimental results establish that PAKES attained a learning progress score of 62%, optimizing the balance between learner control and system control. The earlier cited study [25], under the guise of another research work, presented a reinforcement learning-based adaptive e-learning system utilizing multi-agent technology. The target goal was to design a system that would provide personalized recommendations about learning paths for students. The system considered, besides students' learning styles, their knowledge level, and possibly some disabilities, like hearing impairment, visual impairment, and dyslexia, to match the type of learning sequence with their needs. The method relies on a set of autonomous agents and uses the Q-learning algorithm to adjust the learning paths. The results show that this approach improves the system's performance by making the adaptation process more flexible.

### ***2.4. Deep learning***

Currently, deep learning is used in many fields and has the potential to revolutionize education.

#### *2.4.1. Natural Language Processing (NLP) for Educational Content Analysis*

The authors in [26] proposed an advanced multimodal neural network algorithm for efficient feature extraction in natural language processing, particularly for word segmentation in English. This approach is based on a deep neural network composed of independent subnetworks for each data modality, allowing textual, syntactic, and semantic information to be processed separately. It adopts a hybrid model combining bidirectional gated recurrent units (BI-GRU), which can efficiently capture long- distance dependencies through simplified state update and reset mechanisms, and a conditional random field (CRF) that models dependencies between successive labels to improve the consistency of sequential annotation. Compared to BI-LSTM, the BI-GRU has fewer parameters and faster computations, which contributes to the model's faster computation. The extracted representations are low-dimensional yet discriminative, facilitating classification, segmentation, and other linguistic tasks.

#### *2.4.2. Engage Detection in e-learning*

Bhardwaj et al. [27] propose a deep learning–based approach to assess student engagement in e- learning environments by analyzing real-time facial emotions (anger, joy, sadness, fear, disgust, surprise) and body postures. They use advanced CNNs, including Xception and NASNetMobile, as well as a hybrid model combining both, to extract relevant facial and behavioral features. Video data are complemented by surveys to calculate each student's Mean Engagement Score (MES). Model performance is evaluated using precision, recall, F1-score, and accuracy, reaching up to 89.3show that hybrid models effectively capture complex engagement cues, enabling automated systems for real-time monitoring and adaptive interventions in online courses.

In the context of e-learning, engagement detection can also be approached through the automatic analysis of students' textual feedback, allowing the inference of teaching quality and the perceived involvement of instructors. Li [28] proposes a model combining Particle Swarm Optimization (PSO), an attention mechanism, and a Long Short-Term Memory (LSTM) network to classify textual evaluations of online technical physical education courses. Student evaluations are first preprocessed and then fed into the LSTM to capture the temporal and contextual dependencies of sentences. The attention mechanism assigns weights to the most significant words or phrases for classification. Experimental results show an overall error rate of 10.1% for teachers by gender and 12.8% by teaching level, with an AUC of up to 0.821, demonstrating the high performance of this approach. Although the model does not directly measure engagement, the detailed analysis of textual feedback provides implicit indicators of teacher involvement and motivation, illustrating how deep learning methods can be used to detect or estimate engagement in online learning environments.

#### *2.4.3. Intelligent education system*

X. Zhang and Z. Cao [29] propose a deep learning–based framework for an intelligent education system in higher education. The system helps personalize learning and makes it more efficient by analyzing data from students' interactions, academic results, and behaviors on online platforms. It includes tools for collecting and analyzing data, predicting, and recommending resources. Deep neural networks (DNNs) model how students learn and help forecast their performance. Recurrent networks (RNNs, LSTMs) handle temporal sequences, such as learning progression. Recurrent networks (RNNs, LSTMs) track learning over time. The system can also recommend content tailored to each student. It predicts performance, identifies at-risk learners, and provides targeted interventions. Teachers can use it to spot common weaknesses and adapt courses. Overall, this approach makes learning more personalized, reduces teachers' workload through automated data analysis, and can be integrated with existing university systems for large-scale use.

### ***2.5. Intelligent Comparison of ML Approaches in Education***

To understand the strengths and limits of the main ML approaches in education, the table below gives a clear comparison. It shows typical applications, key benefits, and common challenges for each method. This makes it easier to see where each approach works best.

Table 1. Comparison of Machine Learning Approaches in Education

| Approach | Typical Applications in Education | Strengths | Limitations | References |
|---|---|---|---|---|
| Supervised Learning | Student performance prediction, at-risk student detection | High accuracy in performance prediction | Strong reliance on labeled data, low generalizability, sensitive to class imbalance | [30], [31] |
| Unsupervised Learning | Student segmentation, learning profile identification, behavior analysis | Discovery of hidden structures, no need for labels | Difficult interpretation, challenging to extract actionable insights | [32], |
| Reinforcement Learning (RL) | Dynamic learning path adaptation, intelligent tutoring systems | Real-time personalization, engagement and motivation optimization | Requires large datasets often unavailable in schools, sophisticated infrastructure and algorithms needed, requires many iterations to converge | [33], |
| Deep Learning | Multimodal analysis (text, audio, video), emotion recognition, complex assessments | High modeling power, strong performance on rich datasets | High data and computation requirements, interpretability issues | [34] |

This comparison shows that no single ML approach can cover all educational needs. Each method has its strengths, but also context-specific limitations. That's why hybrid approaches are particularly promising [35] . They combine the best of each method: supervised learning for accurate assessments, unsupervised learning to understand and group learners, reinforcement learning to increase engagement in adaptive tutoring systems, and deep learning to process multimodal data like voice, text, or images. By blending these techniques, we can create more personalized learning experiences. Evaluations become more reliable, and the learning experience is both engaging and adaptable over time.

## 3. Educational Datasets for ML Research

Large-scale ML applications often benefit from available appropriate training data. However, as data increases in scale and importance, so do the risks of harm. These concerns arise in part because educational datasets typically contain the remnants of data-generating processes where complex human decisions were previously applied and do not capture all aspects of world complexity. Recent studies argue that several of these datasets capture offensive content, perpetuate stereotypical biases, and encroach on the privacy of public persons [1]. We argue that many educational datasets in natural language processing and, in particular, reading comprehension, event recognition, and dialogue are derived from routing processes. Moreover, commercial processes that define licenses may prohibit redistribution of the resulting trained models. The aim of this paper is to highlight several of these concerns and initiate dialogue on the development of best practices for dataset development processes. In particular, we hope to work towards guiding principles for structuring licensing terms and scientific processes behind dataset creation to encourage progress towards these principles. We believe that this topic is an essential interdisciplinary discussion necessary to foster research involving MLon the scientific process. In this paper, we address the role of educational datasets and datasets derived from commercially licensed texts in enabling research on MLfor natural language processing and language-related applications. We explore potential negative risks that exist when academic research does not systematically update, modify, or augment some of these educational datasets. Our perspective largely sits within an educational and scientific context and is inspired by ongoing discussions that have arisen from a variety of fields, including ethical quandaries that have arisen when utilizing educational data within the classroom, privacy concerns that affect educational data stored in various repositories, and respect for copyrighted material amidst an expanding appetite for the development of research using these data. Despite these existing discussions, we argue that the scientific process behind educational data generation has not been systematically assessed, formalized, or documented [3].

### *3.1. Significance of Educational Datasets in Machine Learning Research*

Educational datasets facilitate quicker transfer of MLsolutions to applications in education by enabling state-of-the-art experimental validation. They are more amenable to integrate into online courses as interactive resources, and ultimately, they may be used to yield insights and guide actions that improve educational outcomes. The community needs to develop new resources—from previously used datasets to new types of datasets—in deeper alignment with the communities most likely to benefit from them. Educational datasets could be used to generate data that can inform individual teachers and students, guide more data-driven pedagogical environments, and evaluate educational technology. We need to focus on producing datasets that, in some way, make it easier to conduct interesting research. For example, datasets of learners in specific neighborhoods could bring data to assess how learners interact with a certain reality [6].

Insights and novel methods and methodologies produced by research with these types of datasets can be integrated into materials, research agendas, and pedagogical projects of educators, researchers, and developers from very different sub-areas. This includes integrating artificial intelligence teaching materials in K-12 schools and cross-disciplinary dialogues with public audiences around the societal implications of technologically mediated processes. Educational datasets will be motivated by research questions that heavily involve knowledge or experiences from teachers and students, rather than impersonal statistical learning questions. Educational datasets should be reflective and supportive of risks associated with potentially sensitive social science research

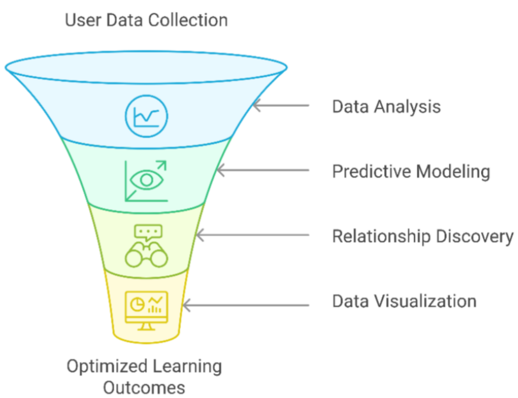


Figure 2. Optimizing Learning Outcome through Data

### 3.2. *Example of Educational Datasets*

Such datasets have contributed immensely to the evolution not just of MLin general, but also the development of its applications in the field of education. They are intended to solve the real problems of the education system, such as making personalized intelligent tutoring systems, improving learning, and anticipating student performance. However, the following are some of the important examples of such education datasets along with their applications:

Table 2. Educational Datasets:description and applications

| **Educational Datasets** | **Description** | **Applications** | **Reference** |
|---|---|---|---|
| ASSISTments Dataset | A dataset from the ASSISTments platform, which is designed for assessing students' learning progress in math | Skill prediction, student performance modeling, and adaptive learning systems | [8], [9] |
| EdNet Dataset | A large-scale dataset covering over 130 million interactions between students and an AI tutoring system | Knowledge tracing, recommendation systems, and content personalization | [11], [12] |
| National Education Longitudinal Study (NELS) | A longitudinal dataset that tracks students' educational progress and background | Educational inequality studies, student outcome prediction, and social determinants of learning | [13], [14] |
| MOOC Datasets (e.g., KDD Cup 2015) | Data collected from online courses, including students' clickstreams, quiz performances, and forum interactions | Dropout prediction, learning behavior analysis, and content effectiveness | [15], [16] |
| Open University Learning Analytics Dataset (OULAD) | Data from distance learning students at the Open University, including demographics, assessment scores, and course activities | Learning analytics, predicting student success, and intervention strategies | [18], [17] |
| PISA Dataset (Programme for International Student Assessment) | An international survey dataset evaluating 15-year-old students' skills in math, science, and reading | Cross-country educational comparisons, policy impact analysis, and skill profiling | [19], [20] |
| Educoder Dataset | Focused on programming education, containing detailed interactions of students solving coding tasks | Automated feedback systems, error analysis, and adaptive coding tutorials | [21],[24] |
| TEACH Dataset | Designed for understanding teacher-student interactions, this dataset includes video annotations and lesson transcripts | Classroom behavior analysis, teacher evaluation, and educational robotics | [25] |

By leveraging these datasets, ML researchers can design systems that enhance personalized education, provide predictive analytics for educators, and enable more inclusive and equitable learning environments.

ML in education only works well if the data is good. But school data often has problems like missing values, confusing labels, or mixed results from different teaching styles and varied assessment practices. To fix this, the data needs to be cleaned, normalized, and kept fair for all student groups. It's also important to clearly explain where the data came from, follow strong protocols for handling sensitive information, and work closely with teachers to make sure the data truly reflects meaningful learning indicators. Addressing these challenges strengthens the validity of ML models and supports their responsible use in educational settings.

## 4. Evaluation of Model Performance

Metrics related to education for MLwill be effective tools for evaluating predictive models and clustering quality. The selection of these metrics has much to do with the educational goal, whether it be predicting at-risk students, examining grading reliability, or clustering students according to performance levels or learning preferences. Such metrics thus provide an in-depth assessment of educational models (see Figure 3).

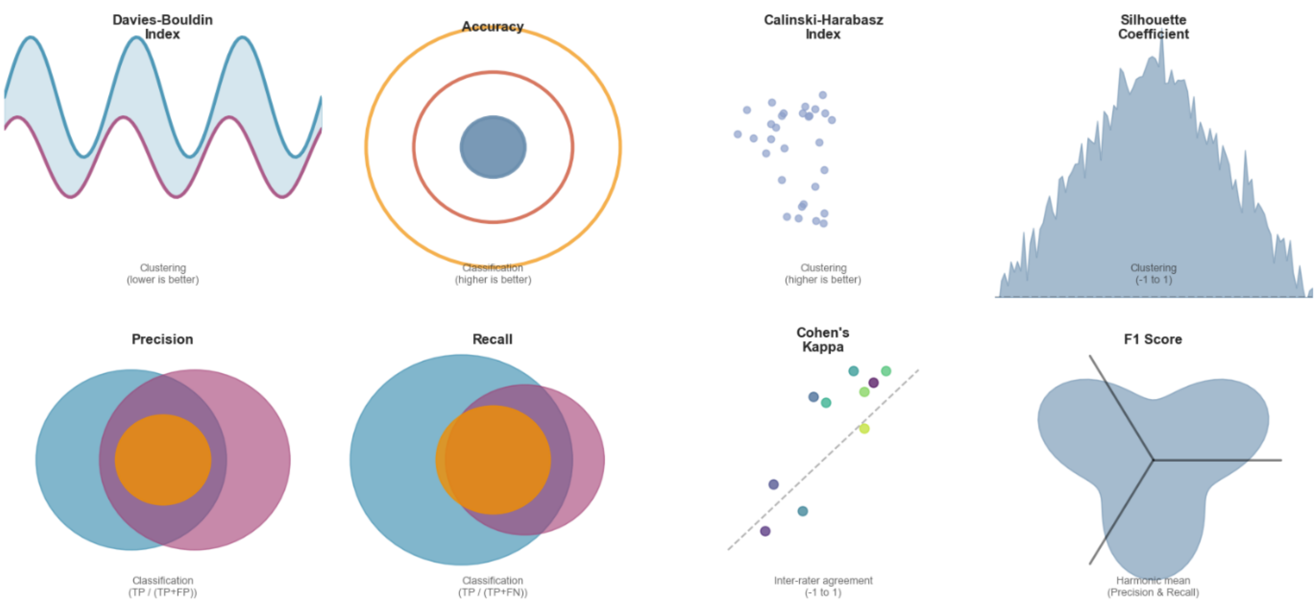


Figure 3. Educational Model Evaluation Metrics

### *4.1. Accuracy*

Accuracy expresses the ratio of values correctly predicted to the total predicted number of values [36]. Accuracy It manifests as one among the common measures that may be used to evaluate a classification model's overall performance. However, it is said that the figure can be misleading when applied to imbalanced datasets, which often include a majority of students as passers in an educational scenario [37]. Accuracy is measured as follows:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}} \tag{1}$$

### 4.2. Precision

Precision is an evaluation metric in ML for classification models [38]. Practically, it compares correct positive predictions to all positive predictions made by the model in the following formula:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{2}$$

Precision is a metric that is closely related to educational applications because there should be really few false positives caught. One application could well be detecting whether a student is failing or dropping out. Very few should be misses for not-at-risk students by using very high precision.

### 4.3. Recall

Recall is one of the ways to quantify the number of actual positives that are correctly identified as true positives [38]. It is measured by the following equation:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{3}$$

In the application of ML in the field of education, recall refers to the proportion of at-risk students identified correctly against all the students truly at-risk. It is an important performance measure among the early warning systems made to reduce the number of missed cases characterized as false negatives, such as future school dropouts. However, high recall tends to increase the number of false positives, as it overlooks fewer at-risk students: Something that is crucial to enabling targeted, effective interventions to improve outcomes in education.

### 4.4. F1 Score

The F1 Score [38] is calculated as:

$$F1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

The F1 score is an important metric in the education domain for evaluating classification models, especially when class distributions are not so evenly split. It presents a balanced view of how the model performs by mixing precision (correctly identifying positive predictions) with recall (detecting positive instances). This makes it applicable in education, wherein the minority, the ones being at risk or wrong answers, are most valued. This permits one to procure a well-balanced minimization of false positives and false negatives; hence, such scores become applicable for identifying students at risk of failing, automating grading, and enabling recommendation systems.

### 4.5. Cohen's Kappa

Cohen's Kappa Coefficient $\beta_k$ [38] is a statistical measure that determines the degree of agreement between two raters/classifications, given the agreement could have happened by chance. Kappa is calculated as:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{5}$$

where $P_o$ is the observed agreement and $P_e$ is the expected agreement by chance.

In education, the metric is best to use for assessing how reliable human-based assessments would be regarding the grading of student responses or visual-deep perspective grading of assignment quality. For instance, in automated or semi-automated exam grading assessments, Cohen's Kappa measures how far two teachers or a teacher and another unknown algorithm would agree on any evaluation criteria. A value close to 1 indicates a very strong agreement, while a value near 0 indicates that the observed agreement was due to chance. Hence, in education, this indicator provides reliability and validity about how the assessment and decision-making on student outcomes stands.

### 4.6. Silhouette Coefficient (SC)

The Silhouette Coefficient (SC) [39] is a metric used for the comparison of clustering quality applied to datasets. It ranges from $-1to1$, with large positive values indicating better-defined clusters. For a point $i$, its Coefficient will be calculated as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{6}$$

where:

$$a(i) = \text{average intra-cluster distance (cohesion)}$$
$$b(i) = \text{average inter-cluster distance to nearest cluster (separation)}$$

In education, clustering is used to ascertain and validate clusters of students based on behaviors, performance, or other common characteristics (like learning styles or examination results). Coherence in clusters constitutes a SC with two major aspects: the proximity of points within the same cluster versus the distance of that cluster from others. A SC value near to 1 suggests strongly defined clusters, while lower SC values, including negative ones, indicate that the groups overlap or are poorly formed. This measurement can help determine the intentions behind any cluster formed in the personalization of a learning pathway or the identification of groups needing some specific form of intervention in education.

### 4.7. Calinski–Harabasz Index

The Calinski–Harabasz index [14] is a measure used to evaluate the quality of clustering. Its formula is:

$$CH = \frac{\text{tr}(B_k)}{\text{tr}(W_k)} \times \frac{n - k}{k - 1} \tag{7}$$

where:

- $tr(B_k)$=between-cluster dispersion
- $tr(W_k)$= within-cluster dispersion
- $n$= number of data points
- $k$= number of clusters

In education, the meaningfulness of clusters can be popularly applied to study features such as student behaviours, academic performance, or learning preferences. The index compares the other form of dispersion, called inter-cluster, against one another. When the index value is high, it indicates that clusters are compacted together and well separated from each other; it implies meaningful grouping. This index could be used, for example, to validate the separation of learners in adaptive learning systems, where the creation of homogeneous groups is vital to offer personalised educational content for each profile.

### 4.8. Davies–Bouldin Index

The Davies-Bouldin index [14] increasingly describes the quality of clustering in that it examines both the internal validity and external separation of clusters. Within education, the application of this index could be in comparing the differences in performance, behaviors, or learning styles regarding their groupings. The equation is:

$$DB = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} \left( \frac{S_i + S_j}{d_{i,j}} \right) \tag{8}$$

Where:

- $S_i$ is the intra-cluster dispersion of cluster $i$ (average distance between each point in cluster $i$ and its centroid)
- $S_j$ is the intra-cluster dispersion of cluster $j$ (average distance between each point in cluster $j$ and its centroid)

- $d_{i,j}$ is the inter-cluster distance between centroids of clusters $i$ and $j$
- $k$ is the number of clusters

A lower DB index indicates:

- Better cluster compactness (smaller $S_i$ values)
- Better separation between clusters (larger $d_{i,j}$ values)

In educational contexts, this index helps validate learner segmentation in adaptive learning systems, ensuring meaningful groupings for personalized pedagogical interventions.

## 5. Critical Analysis of ML in Education

ML is rapidly making its place in education. It promises personalized learning paths, automated assessments, and stronger support for teachers and institutions. But along with these opportunities come important challenges—technical, pedagogical, ethical, and organizational. Addressing these challenges openly is essential for developing strategies that use ML responsibly and effectively in schools and universities.

### *5.1. Technical Limitations*

Technical challenges in educational ML mainly come from the robustness of models and the quality and availability of data. In many institutions, data are often limited, fragmented, or heterogeneous [40]. They can also differ widely from one context to another, making them hard to use reliably. As a result, models may not fit local realities very well [41, 42, 43]. These limitations affect generalizability: a model trained on a "general" dataset can quickly lose effectiveness when applied in a different institution [43, 44]. Furthermore, many models, like deep neural networks or random forests, also act as "black boxes." Their decisions are hard for teachers and decision-makers to understand. This lack of transparency can reduce trust and make it harder to adopt these decision-support systems in education [45].

### *5.2. Pedagogical Barriers*

The main pedagogical challenges of using ML in education concern teachers' skills, how well these tools fit into daily teaching, and creating classroom dynamics that uses ML without losing sight of learning goals. Studies show that most teachers don't get enough training in AI and ML. This limits their ability to use these tools effectively or responsibly.

Training teachers is key to using ML effectively in schools. With the right support, teachers can learn how to read data, understand what the technology is showing, and use it in ways that actually help students learn. Good training helps teachers feel more comfortable with new tools, make smart and fair decisions, and use ML to improve teaching—instead of making things more confusing or unfair.

Teachers may also resist change, often because they already face heavy workloads [46], [47]. Additionally, adaptive platforms sometimes put too much importance on the technology itself, which can take attention away from core learning goals [48]. ML–based interventions also face algorithmic and methodological limits. Generative models can capture students' attention, but they sometimes give inaccurate or shallow answers. This can cause confusion and misinterpretation [49]. They also show weaknesses to support critical thinking and deeper analysis. At the same time, it's hard to measure their true educational value because we lack good indicators, especially for complex skills and knowledge transfer [50]. Together, these issues highlight that it's not enough for models to simply work. We also need solid ways to evaluate and optimize their real impact on learning.

### *5.3. Ethical and Social Issues*

Using ML in education also raises important ethical and social questions. Some of the biggest concerns are around privacy, bias, fairness, and transparency. For example, algorithmic bias is a major issue. If training data reflect social inequalities or lack representativeness, ML systems can create unfair outcomes that put marginalized groups at a disadvantage [48]. Another concern is the heavy dependence on large-scale data collection. While this makes powerful applications possible, it also brings serious risks around how students' sensitive information is stored, protected, and used [46, 51]. The lack of clear guidelines and consistent standards only adds to the problem, highlighting the need for strong frameworks that ensure ML in education is used responsibly and transparently [52].

### 5.4. *Institutional and Organizational Constraints*

Bringing ML into education isn't just about having the technology. It also requires solid infrastructure, clear policies, good coordination, and skilled people to manage it. In many contexts, these conditions are still incomplete or missing. Studies often point to weak infrastructure, poor governance, and limited institutional commitment—all of which slow down the sustainable use of ML in schools [46]. On top of that, material issues like a lack of equipment or poor internet access make it harder to deploy these systems and create unfair gaps in student access [46]. From a regulatory standpoint, frameworks are still incomplete, especially when it comes to certifying models and holding institutions accountable, which delays large-scale adoption [51]. Success also depends on having qualified staff and stable funding to keep systems updated, maintained, and evaluated [47]. But even then, problems like data silos and missing technical standards remain big hurdles, making it hard to connect with learning management systems (LMS) and to share resources securely [53].

One big challenge for schools using ML is the cost. It's not just about having the right staff and infrastructure, implementing ML also requires investments in hardware, secure data storage, software licenses, and regular maintenance. These costs can be significant, especially for schools with smaller budgets, and may make the gap between rich and poor schools even wider. That's why it's important to have smart funding plans, like help from the government, partnerships with companies, or using free open-source tools. Without steady funding, even great ML projects may struggle to be deployed or maintained over time.

### 5.5. *Strategies for Overcoming Challenges*

Adopting ML in education requires rigorous practices to make sure it's reliable, safe, and truly useful for teaching and learning. To get the best results and to build trust, there are several actions we can take that will help these systems be more effective and easier to accept in the classroom:

- Diversify datasets by collecting multicentric and representative information across different school levels, languages, and socio-economic contexts to reduce bias and strengthen model robustness.
- Conduct longitudinal studies to monitor the long-term impact of ML on student learning and adaptation, while testing models on external datasets to verify their generalizability.
- Involve teachers and students from the early stages of system design, particularly in defining needs and what success looks like, to align with educational goals and encourage its uptake [54].
- Develop continuous training programs that give educators not just the basics of ML, but also the skills needed to understand the ethical and practical implications of these technologies [55].
- Make model limitations transparent by documenting uncertainties and conditions of applicability to prevent inappropriate use.
- Ensure data security and compliance through proper encryption and clear policies for consent and access management for teachers, students, and parents.

## 6. Emerging Trends and Future Research Directions

This section examines emerging trends and outlines promising research directions that could further enhance the application of machine learning in educational contexts.

### *6.1. Emerging Trends*

Recent studies highlight new trends that are reshaping how ML is applied in education. The focus is moving toward approaches that combine multiple types of data, generate content dynamically, provide clear explanations, and encourage collaboration —always grounded in ethics and human-centered design. Multimodal approaches, for example, go beyond simple metrics like grades or login times. They analyze natural language, speech, and even facial expressions to capture student engagement and emotions with over 85% accuracy. This offers deeper insight into both cognitive and emotional needs [56].

Generative models are also on the rise. They can quickly create learning materials, help with writing, and give instant feedback. Large language models (LLMs), such as GPT-based systems, can personalize exercises and resources while fostering student creativity.This points toward an Education 5.0 era, where humans and machines work together to shape learning experiences [57, 58].

At the same time, interest in explainable and ethical AI is growing. Methods such as SHAP and LIME help make algorithms more transparent and fairer by giving teachers and students clear reasons behind recommendations. This kind of transparency builds trust and supports shared responsibility [59].

Finally, collaboration across disciplines is growing. Researchers, computer scientists, and policymakers are joining forces to create solutions that truly fit the needs and realities of education. All of these trends are coming together to create an educational ecosystem where technology adjusts to each individual's needs, while keeping teachers and students at the heart of education, guided by principles of ethics, transparency, and collaboration.

### *6.2. Future Research Directions*

Future research on ML in education should aim for solutions that bring different elements together while being sustainable and effective over the long term:

- A priority is to conduct longitudinal studies to assess the real and lasting effects of ML on learning, beyond immediate gains [60].
- Looking into hybrid models that combine algorithmic insights with human expertise is a promising way to support teachers' autonomy while boosting the educational value of these tools [61].
- Establishing classroom labs where teachers, students, education researchers, and ML engineers co-develop and iterate prototypes ensures alignment between pedagogical requirements and technical constraints [62].
- Increasing attention to linguistic, cultural, and socio-economic diversity is essential to ensure that ML supports truly inclusive education adapted to local realities.

Following these directions will strengthen the technical capabilities of ML in education. More importantly, they will ensure that its innovations align with core values like fairness, transparency, and accountability.

## 7. Conclusion

Integrating AI, particularly ML, into the educational area represents an education delivery and experience undergoing a fundamental transformation. Through the utilization of power elements ML categories include supervised, unsupervised, reinforcement, and deep learning. Educational institutions have gained the capability to deliver tailored learning experiences, predict student needs, and optimize administrative processes. Using the right techniques, along with good datasets and appropriate metrics, can help us better understand how students learn and create teaching strategies that work more effectively. But for ML to really work in education, we have to deal with challenges like data quality, bias in algorithms, making models easy to understand, training teachers, and having the right technology in place. To use ML responsibly, we need strong governance, teamwork between different experts, and a sustained commitment to equity and inclusion so that technological progress genuinely benefits all learners. Meanwhile, several emergent trends, including multimodal learning analytics, explainable AI, and hybrid systems that blend human and machine intelligence, signal the emergence of the next generation of intelligent educational technologies. Such directions will be explored in future research through longer-term studies, more representative datasets, and by closer partnerships among educators, data scientists, and policymakers.

This paper aims to give a balanced and thoughtful view on the role of ML in transforming educational environments, melding the overview of current ML applications with the critical examination of the limitations of the current uses and future prospects. The continued development of ML in education promises adaptive, equitable, and efficient learning systems capable of supporting a wide range of learners in an increasingly digital world.

## REFERENCES


1. A. Algarni, *Data mining in education*, International Journal of Advanced Computer Science and Applications, vol. 7, no. 6, pp. 456–461, 2016.
2. J. Smith and L. Johnson, *The Role of Machine Learning in Modern Education*, Educational Review Quarterly, vol. 33, no. 2, pp. 145–160, 2022.
3. D. Kučak, V. Juričić, and G. Dambić, *Machine learning in education - a survey of current research trends*, Annals of DAAAM & Proceedings, vol. 29, 2018.
4. M. Garcia and P. Martinez, *Predictive Analytics for Student Success: A Machine Learning Approach*, Journal of Educational Analytics, vol. 19, no. 4, pp. 301–320, 2022.
5. Q. Nguyen, T. Brown, and P. O'Connor, *Early Intervention Strategies Through Predictive Machine Learning Models*, Educational Intervention Journal, vol. 18, no. 3, pp. 210–225, 2023.
6. M. Yağcı, *Educational data mining: prediction of students' academic performance using machine learning algorithms*, Smart Learning Environments, vol. 9, no. 1, p. 11, 2022.
7. A. Elouafi, I. Tammouch, S. Eddarouich, and R. Touahni, *Evaluating various machine learning methods for predicting students' math performance in the 2019 TIMSS*, Indonesian Journal of Electrical Engineering and Computer Science, vol. 34, no. 1, pp. 565–574, 2024.
8. Y. Li, *Data analysis of student academic performance and prediction of student academic performance based on machine learning algorithms*, Communications in Humanities Research, vol. 32, pp. 65–71, 2024.
9. T. Koç and P. Akın, *Estimation of high school entrance examination success rates using machine learning and beta regression models*, Journal of Intelligent Systems: Theory and Applications, vol. 5, no. 1, pp. 9–15, 2022.
10. P. O'Connor and Y. Wang, *Automated Grading Systems: Enhancing Efficiency in Education*, Journal of Educational Technology Systems, vol. 49, no. 2, pp. 198–215, 2021.
11. D. Chen and F. Xu, *Design and implementation of machine learning algorithms in automatic grading of students' assignments*, Journal of Electrical Systems, vol. 20, no. 3s, pp. 899–919, 2024.
12. L. Huang and L. Ma, *Research on the Application of Intelligent Grading Method based on Improved ML Algorithm in Sustainable English Education*, Scalable Computing: Practice and Experience, vol. 25, no. 1, pp. 451–463, 2024.
13. N. Alruwais and M. Zakariah, *Evaluating student knowledge assessment using machine learning techniques*, Sustainability, vol. 15, no. 7, p. 6229, 2023.
14. A. F. Mohamed Nafuri, N. S. Sani, N. F. A. Zainudin, A. H. A. Rahman, and M. Aliff, *Clustering analysis for classifying student academic performance in higher education*, Applied Sciences, vol. 12, no. 19, p. 9467, 2022.
15. N. Tzenios, *Clustering students for personalized health education based on learning styles*, Sage Science Review of Educational Technology, vol. 3, no. 1, pp. 22–36, 2020.
16. R. Vankayalapati, K. B. Ghutugade, R. Vannapuram, and B. P. S. Prasanna, *K-means algorithm for clustering of learners performance levels using machine learning techniques*, Revue d'Intelligence Artificielle, vol. 35, no. 1, pp. 99–104, 2021.
17. I. Tammouch, A. Elouafi, S. Eddarouich, and R. Touahni, *Centroid competitive learning approach for clustering and mapping the social vulnerability in Morocco*, International Journal of Advanced and Applied Sciences, vol. 9, no. 9, pp. 70–77, 2022.
18. R. T. Javed et al., *Get out of the BAG! Silos in AI ethics education: unsupervised topic modeling analysis of global AI curricula*, Journal of Artificial Intelligence Research, vol. 73, pp. 933–965, 2022.
19. L. A. Demidova, P. N. Sovietov, E. G. Andrianova, and A. A. Demidova, *Anomaly detection in student activity in solving unique programming exercises: Motivated students against suspicious ones*, Data, vol. 8, no. 8, p. 129, 2023.
20. A. Venkatesh, T. Gonsalves, A. Monk, K. Buckner, B. S. Sarma, and B. Ravindran, *Intelligent tutoring systems using reinforcement learning to teach autistic students*, in *Home Informatics and Telematics: ICT for The Next Billion*, Springer, pp. 65–78, 2007.
21. R. Pauly, L. Heindrich, V. Amo, and F. Lieder, *What to learn next? Aligning gamification rewards to long-term goals using reinforcement learning*, in *Proceedings of the Multi-disciplinary Conference on Reinforcement Learning and Decision Making*, vol. 2022, pp. 8–11, 2022.
22. A. Brown, B. Clark, and S. Davis, *Adaptive Learning Technologies in Modern Education*, Educational Technology Journal, vol. 45, no. 3, pp. 234–250, 2021.
23. J. Lee and M. Park, *Real-Time Adaptation in E-Learning Platforms Using Machine Learning*, International Journal of E-Learning & Teaching Technologies, vol. 20, no. 1, pp. 77–90, 2023.
24. M. Z. Islam, R. Ali, A. Haider, M. Z. Islam, and H. S. Kim, *Pakes: a reinforcement learning-based personalized adaptability knowledge extraction strategy for adaptive learning systems*, IEEE Access, vol. 9, pp. 155123–155137, 2021.
25. H. El Fazazi, M. Elgarej, M. Qbadou, and K. Mansouri, *Design of an adaptive e-learning system based on multi-agent approach and reinforcement learning*, Engineering, Technology & Applied Science Research, vol. 11, no. 1, pp. 6637–6644, 2021.
26. D. Wang, J. Su, and H. Yu, *Feature extraction and analysis of natural language processing for deep learning English language*, IEEE Access, vol. 8, pp. 46335–46345, 2020.
27. P. Bhardwaj, P. Gupta, H. Panwar, M. K. Siddiqui, R. Morales-Menendez, and A. Bhaik, *Application of deep learning on student engagement in e-learning environments*, Computers & Electrical Engineering, vol. 93, p. 107277, 2021.

28. F. Li, *Creation of Deep Learning Scenarios in the Network Teaching of Physical Education Technical Courses*, Scalable Computing: Practice and Experience, vol. 25, no. 1, pp. 271–284, 2024.
29. X. Zhang and Z. Cao, *A framework of an intelligent education system for higher education based on deep learning*, International Journal of Emerging Technologies in Learning, vol. 16, no. 7, p. 233, 2021.
30. C. Romero and S. Ventura, *Educational Data Mining: A Review of the State of the Art*, IEEE Trans. Syst., Man, Cybern. C, 2010.
31. S. Papamitsiou and A. Economides, *Learning Analytics and Educational Data Mining in Practice*, Educ. Technol. Soc., 2014.
32. J. Peña-Ayala, *Educational Data Mining: A Survey*, Expert Syst. Appl., 2014.
33. M. A. Md Zuki, N. Mohamad Ali, and J. K. Chaw, *Reinforcement learning: methods and recent applications*, J. Inf. Syst. Technol. Manag., vol. 9, no. 36, pp. 67–89, 2024. Int. J. Appl. Res., 2018.
34. K. He, X. Zhang, S. Ren, and J. Sun, *Deep Residual Learning for Image Recognition*, CVPR, 2016.
35. G. Al-Tameemi, J. Xue, I. Hadi, and S. Ajit, *A Hybrid Machine Learning Approach for Predicting Student Performance Using Multi-class Educational Datasets*, Procedia Computer Science, vol. 238, pp. 888–895, 2024.
36. S. N. Brohi, T. R. Pillai, S. Kaur, H. Kaur, S. Sukumaran, and D. Asirvatham, *Accuracy Comparison of Machine Learning Algorithms for Predictive Analytics in Higher Education*, in *Advances in Intelligent Systems and Computing*, Springer, pp. 254–261, 2019.
37. A. F. Rozi, A. Wibowo, and B. Warsito, *Resampling Technique for Imbalanced Class Handling on Educational Dataset*, Jurnal Informatika: Juita, vol. 11, no. 1, p. 77, 2023.
38. Ž. Vujović, *Classification model evaluation metrics*, International Journal of Advanced Computer Science and Applications, vol. 12, no. 6, pp. 599–606, 2021.
39. P. J. Rousseeuw, *Silhouettes: a graphical aid to the interpretation and validation of cluster analysis*, Journal of Computational and Applied Mathematics, vol. 20, pp. 53–65, 1987.
40. S. M. Levin, *Review of Machine Learning Models for Application in Adaptive Learning for Higher Education Student*, International Journal For Multidisciplinary Research, vol. 6, no. 2, 2024.
41. S. M. Razaulla, M. Pasha, and M. U. Farooq, *Integration of machine learning in education: challenges, issues and trends*, in Machine learning and internet of things for societal issues, pp. 23–34, Springer Nature Singapore, Singapore, 2022.
42. N. A. Rahman, S. L. Wang, T. F. Ng, and A. S. Ghoneim, *Artificial Intelligence in Education A systematic review of machine learning for predicting student performance*, Journal of Advanced Research in Applied Sciences and Engineering Technology, 2024.
43. S. Esakkiammal and K. S. Kasturi, *Advancing Educational Outcomes with Artificial Intelligence Challenges, Opportunities, And Future Directions*, International journal of computational and experimental science and engineering, 2024.
44. J. Gardner, Y. Yang, R. Baker, and C. Brooks, *Enabling End-To-End Machine Learning Replicability A Case Study in Educational Data Mining*, arXiv, 2018.
45. T. Räz, *ML interpretability: Simple isn't easy*, Studies in history and philosophy of science, vol. 103, pp. 159–167, 2024.
46. U. R. G. Inca et al., *Desafíos de la Educación para la Implementación de la Inteligencia Artificial*, Ciencia latina, 2024.
47. W. Forero-Corba and F. N. Bennasar, *Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación una revisión sistemática*, RIED: Revista Iberoamericana de Educación a Distancia, 2023.
48. E. S. Animashaun, B. T. Familoni, and N. C. Onyebuchi, *Advanced machine learning techniques for personalising technology education*, Computer science & IT research journal, 2024.
49. E.Ortega-Ochoa, J. Sabaté, M. Arguedas, J.Conesa, T. Daradoumis, and S. Caballe, *Exploring the utilization and deficiencies of Generative Artificial Intelligence in students' cognitive and emotional needs*, Frontiers in artificial intelligence, vol. 7, 2024.
50. B. Memarian and T. Doleck, *A review of assessment for learning with artificial intelligence*, Computers in Human Behavior: Artificial Humans, vol. 2, no. 1, 2024.
51. X. V. Huong, *The implications of artificial intelligence for educational systems challenges, opportunities, and transformative potential*, The American journal of social science and education innovations, 2024.
52. P. K. Goel et al., *AI and Machine Learning in Smart Education*, in Advances in web technologies and engineering book series, 2024.
53. A. Kumar et al., *A Comprehensive Survey on AI in Learning Management System*, preprint, 2025.
54. Mei Tan, H. Lee, D. Wang, and H. Subramonyam, *Is a Seat at the Table Enough? Engaging Teachers and Students in Dataset Specification for ML in Education*, ACM, New York, NY, USA, 2023.
55. L. H. Luan, P. Geczy, H. Lai, J. Gobert, S. J. H. Yang, H. Ogata, J. Baltes, R. Guerra, P. Li, and C.-C. Tsai, *Challenges and Future Directions of Big Data and Artificial Intelligence in Education*, Front. Psychol., vol. 11, p. 580820, 2020.
56. A. Imran, J. Li, and A. Alshammari, *AI-driven educational transformation in ICT: Improving adaptability, sentiment, and academic performance with advanced machine learning*, PLoS One, vol. 20, no. 5, e0317519, 2025.
57. X. Tu, J. Zou, W. J. Su, and L. Zhang, *What should data science education do with large language models?*, arXiv preprint arXiv:2307.02792, 2023.
58. S. Chakraborty, *Generative AI in modern education society*, arXiv preprint arXiv:2412.08666, 2024.
59. R. S. Akhmadieva, N. A. Kalmazova, T. Belova, A. Prokopyev, N. M. Molodozhnikova, and V. Y. Spichak, *Research trends in the use of artificial intelligence in higher education*, Frontiers in Education, vol. 9, 2024.
60. H. Hammoudi Halat, S. A. Al-Kuwari, S. A. Elhassan, A. Al-Raeesi, M. A. Mohammed, and F. A. Al-Kuwari, *Use of machine learning to assess factors affecting progression, retention, and graduation in first-year health professions students in Qatar: a longitudinal study*, BMC Medical Education, vol. 23, no. 1, p. 634, 2023.
61. T. Thomas, J. P. Nunes, B. Schmitz, and A. Holzinger, *Improving Student Learning with Hybrid Human-AI Tutoring: A Three-Study Quasi-Experimental Investigation*, arXiv preprint arXiv:2312.11274, 2023.
62. M. Cukurova, *The interplay of learning, analytics and artificial intelligence in education: A vision for hybrid intelligence*, British Journal of Educational Technology, vol. 55, no. 2, pp. 344–358, 2024.