KMANPUB AI and Tech in Behavioral and Social Sciences Vol. 3 No. 4 (2025): 1-17

Article history:
Received 28 July 2025
Revised 18 October 2025
Accepted 25 October 2025
Published online 07 November 2025

Personalized Learning through Machine Teaching and Machine Learning: Enhancing Adaptive Educational Systems

Farzaneh Nasresfahani[1], Mohammad Reza Eslahchi[2*]

[1] Department of Mathematics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran, Tehran, Iran
[2] Department of Computer Science, Faculty of Mathematical Sciences, Tarbiat Modares University, P.O. Box 14115-134, Tehran. Iran

*** Corresponding author email address**: eslahchi@modares.ac.ir

Article Info

Article type:
Original Research

How to cite this article:
Nasresfahani, F. & Eslahchi, M. R., (2025). Personalized Learning through Machine Teaching and Machine Learning: Enhancing Adaptive Educational Systems. *AI and Tech in Behavioral and Social Sciences*, *3*(4), 1-17.
https://doi.org/10.61838/kman.aitech.4524

ABSTRACT

Machine teaching, as a new approach to artificial intelligence, focuses on the purposeful design and selection of educational data to optimize the learning process. When it comes to human education, the quality of educational data plays a fundamental role in identifying and enhancing those individual characteristics of students that have the greatest impact on their learning path. This article demonstrate the role of machine teaching in personalize learning and highlights the importance of selecting data that can reveal students' strengths and weaknesses in various skills. The results show that, in addition to accelerating his learning process, a strategic and optimal data design can provide personalized educational paths and help teachers make more effective educational decisions.

Keywords: *Machine teaching, Pedagogical content knowledge, Personalized learning, Machine learning*

1. Introduction

In the early 19th and 20th centuries, approaches such as monitorial systems and attempts to adapt instruction to each learner's pace and level first emerged (Idowu, 2024). Particularly in the 1960s and 1970s, individualized instructional models, such as the Keller Plan, and theories like Bloom's Mastery Learning gained prominence as some of the most significant developments of that era. Learner progress at an individual pace is typically emphasized by these models (Block & Burns, 1976). This focus on individualized pacing paved the conceptual way for integrating technology into the learning process. Thus, with the advent of computers and educational software in the 1960s and 1970s, the first examples of technology-enhanced personalization began to emerge. For instance, PLATO and TICCIT allowed students to engage with the material and advance at their own pace (Anastasio & Alderman, 1973). These early initiatives prepare the foundation for the modern concept of personalized learning, which is now defined as a transformative approach that tailors educational strategies and material to the unique needs, preferences, and goals of each learner (Chatti et al., 2013; Miliband, 2006).

(Gómez et al., 2014)Personalized learning is an educational strategy that customizes instruction, materials, and learning experiences to match the unique needs, abilities, interests, and pace of every student. Rather than employing a uniform curriculum, personalized learning enables students to advance according to their skills and preferred learning styles, frequently utilizing adaptive technologies, varied teaching methods, and ongoing assessments to inform instruction (Grant & Basye, 2014; Kaswan et al., 2024). Customizing the learning process can improve learning outcomes and general satisfaction, as well as motivation, engagement, and knowledge retention (Gómez et al., 2014). However, even with its potential, implementing personalized learning successfully in today's educational systems remains very challenging. To overcome these challenges and make scalable, data-driven personalization possible, emerging technologies, particularly artificial intelligence, are essential.

Artificial intelligence first emerged as a research field in 1956. During the 2010s, the discipline experienced rapid breakthroughs and reached unprecedented levels of growth and application (McCarthy et al., 2006). In the meantime, attending to the principles of pedagogy and pedagogical content knowledge has also gained special importance. Pedagogy, as the art and science of teaching, includes a set of methods and strategies that facilitate the effective transfer of knowledge. Amidst this, a concept which was first proposed by Shulman, emphasizes the integration of specialized content knowledge with teaching skills (Shulman, 1987). Nonetheless, despite comprehensive quantitative and qualitative investigations into individualized education, the constraints of human judgment and the limited precision linked to small-scale data are still apparent.

(Nguyen et al., 2019)In this context, machine learning has emerged as a potent artificial intelligence instrument that facilitates the analysis and interpretation of extensive educational data, hence enabling more precise predictions and data-driven insights into learning processes (Nguyen et al., 2019).

With the expansion of computer technologies and the development of machine learning algorithms, the need to use appropriate data to improve educational processes has become more important than ever. Machine learning, like any other computational problem, relies on the careful design of algorithms that transform raw inputs , which is called training data, into desired outputs, which is a model with optimal performance. In the simplest definition, an algorithm is a set of instructions that transforms input into output; such as a number sorting algorithm that arranges an unordered list in a desired order. In some cases, there are several efficient algorithms with different characteristics and the goal is to optimize time, memory, or both. However, in many real-world problems, there is no specific algorithm for converting input into output. For example, in education, identifying which students are at risk of academic failure in mathematics is a major challenge. The input includes data such as previous test scores, class attendance, study times, and online learning activities, and the output should provide a yes or no answer: Does this student need special educational intervention? There is no explicit algorithm for this problem, because academic failure is influenced by complex individual, social, and motivational factors that change over time. In these situations, sample data is used. By collecting behavioral and academic data from thousands of students, the patterns in the data can be used to learn models that predict academic failure. In this way, machine learning enables the analysis of big educational data and the prediction of future student behavior (Asnicar et al., 2024; Nafea et al., 2024; Srivastava, 2023). Today, using learning management platforms, huge data such as the number of study hours, the level of participation in class activities, and test results are recorded. If properly analyzed, this data can help design personalized educational programs. For instance, identifying students who experience challenges in learning basic mathematical concepts and providing them with specific reinforcement exercises. Although it seems impossible to fully identify the reasons for academic failure, analyzing the patterns in the data allows for relative prediction of this process. Increasing the amount of study or regular attendance in class is usually associated with improved grades. Also, identifying the best times to study based on recorded behavioral patterns can improve the quality of learning.

While machine learning has emerged as a promising methodology in education, the efficacy of these models is fundamentally contingent upon the quality and selection of training data (Angeioplastis et al., 2025; Esomonu, 2025). Inadequately selected or limited data can result in biased results, lower generalizability, and erroneous conclusions. To tackle this issue, the concept of machine teaching has been proposed.

In contrast to machine learning, which aims to uncover knowledge from pre-existing data, machine teaching focuses on the careful design and selection of training

AITBSS AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

examples to direct the learner toward a specific model (Zhu, 2015; Zhu, 2018). By thoughtfully assembling instructional data, machine teaching enhances both efficiency and precision while ensuring that computational learning methods meet educational goals, establishing it as a significant framework for promoting personalized and adaptive education.

This article seeks to explore how machine learning and machine teaching can enhance the educational process and their applications in learning environments, particularly focusing on personalized learning informed by a pedagogical viewpoint.

2. Literature Review

The concept of artificial intelligence can be traced back to the period of shaping the idea of thinking machines. One of the founding figures in this regard is Alan Turing, who introduced ideas in the 1940s about simulating human intelligence and learning through machines. By proposing the perceptron algorithm (Rosenblatt, 1958) in the early 1950s by Frank Rosenblatt, one of the first substantial contributions toward computational models of learning was provided. Though this model demonstrated potential at first, its inherent flaws, particularly in handling nonlinear pattern recognition, soon became evident.

Minsky and Papert conducted a critical examination of the perceptron during the 1960s in their esteemed book (Minsky & Papert, 1969). These points led to a decrease in worldwide interest in neural networks among researchers for over a decade. Consequently, the focus of research shifted again towards symbolic and rule-based methods to the point that the 1980s witnessed the creation of algorithms like linear regression, principal component analysis (PCA), and clustering techniques. This developement in reasearch n this fieled established crucial theoretical frameworks for the burgeoning field of machine learning, propelled by advancements in computational capabilities and the increasing impact of statistical approaches in data analysis.

The introduction of the backpropagation algorithm by Rumelhart, Hinton, and Williams (Rumelhart et al., 1986) in the 1980s enabled multilayer neural networks to tackle more complex problems. This development represents a watershed that sparked a new interest in neural network research. As the 1990s rolled around, this field underwent an evolution with the emergence of foundational concepts and the publication of influential textbooks, including Machine Learning by Mitchell (Mitchell, 1997) and Pattern Recognition and Machine Learning by Bishop (Bishop, 2006). During this period, the development of algorithms such as ensemble methods, decision trees, and support vector machines (SVMs) during this time greatly expanded the practical uses of machine learning in domains like regression, prediction, and classification.

Machine teaching specifies the optimal training data to drive the learning algorithm to a target model with maximum efficiency (Zhu, 2015). Unlike traditional machine learning methods, where models must infer patterns from raw data, machine teaching focuses on intelligently crafting the data itself to enhance the learning process. Zhu and colleagues (Zhu, 2018), who introduced machine teaching as a complementary approach to machine learning, presented one of the early pioneering studies in this area.

In terms of education, machine teaching involves a teacher with a desired goal, whose objective is to find an optimal training sequence to steer a learner towards this goal. For instance, an educational goal for a teacher can be communication with a student via a set of demonstrations. The teacher manipulates the behavior of a machine learning system by maliciously modifying the training data via a machine teaching technique.

Machine teaching can be formulated as an optimization problem, where the objective function balances the size of the training dataset against the learner's prediction error. A well-known example of this approach involves Bayesian learners, where training data must be carefully designed to ensure that the probabilistic structure of the model is learned as accurately as possible (Liu & Zhu, 2017).

From a methodological perspective, machine teaching can be categorized into several subfields:

- Machine Teaching for Linear Models: This area focuses on identifying optimal datasets for learning models such as linear regression and support vector machines. Researchers aim to determine the minimal yet most informative set of examples that enable effective model training.
- Machine Teaching for Deep Learning: As model architectures have become more complex, designing optimal training data has also grown more challenging. Techniques in this area often involve selecting key examples from large datasets to maximize learning efficiency (Jia & Zhu, 2020).

AITBSS
AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

- Adaptive Machine Teaching: In this paradigm, the teacher can provide new data dynamically based on the learner’s performance during the training process, thereby enhancing learning effectiveness.

In these approaches, a set of data is presented to the learner to construct an optimal model. However, selecting the most appropriate dataset for the learning process remains a fundamental challenge in machine teaching. The authors of (Celikok et al., 2020) introduced an innovative approach that goes beyond static teaching by treating the learner as a dynamic agent. This methodology offers new opportunities to enhance the teaching process and can be regarded as a significant advancement in the development of machine learning techniques.

This approach has found applications across a variety of domains, including robot training, recommender system training, and personalized learning. In personalized learning, a human or machine teacher designs tailored training data for individual users to optimize learning outcomes.

The link between machine teaching and pedagogy is noteworthy. Pedagogy looks at how to teach people, while machine teaching aims to apply these ideas in artificial learning settings. This approach can improve human–machine interaction and raise the quality of AI learning. As a result, it can act as an intelligent teacher or, better to say, it can implement the idea of personalized learning in an artificial manner (Díaz & Nussbaum, 2024; Kayal, 2024; Pedro et al., 2019).

Pedagogy, which is both the science and art of teaching, comes from the Greek word paidagogos, meaning "guide of the child" (Smith, 2020). In ancient Greece, this term referred to slaves who accompanied and supervised children on their way to school or the learning process. Over time, the concept of pedagogy evolved into the systematic study and practice of teaching and learning.

During the Middle Ages, the Church mostly controlled education and teaching methods focused on memorizing and repeating religious texts (Goodson, 2017). When the Renaissance began, the focus moved toward humanism and critical thinking. This change resulted in a broader range of teaching and learning methods (Gutek, 1995). During this time, the teaching approach passed on knowledge to promote creative thinking and analytical skills.

In the 17th century, John Amos Comninus became one of the first theorists of modern teaching. He proposed ideas such as step-by-step instruction and using visual aids to improve the learning process (Morison, 1932). In the 18th and 19th centuries, the thoughts of Jean-Jacques Rousseau (Rousseau, 1762), Johann Pestalozzi (Pestalozzi, 1898), and Friedrich Fröbel (Froebel, 1886) greatly influenced educational theory. Rousseau, in his well-known work Émile, stressed the value of hands-on learning and natural education (Rousseau, 1762).

With the emergence of educational psychology in the early 20th century, educational theories based on cognitivism and behaviorism emerged. Among the influential researchers in this field are John Dewey and Jean Piaget, who played a significant role in the development of this field. Dewey introduced education as a social and participatory process, emphasizing learning through experience and critical thinking.

In recent decades, the rise of digital technologies and artificial intelligence has brought profound transformations to educational practices (Mayer, 2014). Within this context, concepts such as machine learning and machine teaching have emerged as integral components of digital pedagogy. These approaches support teachers and instructional designers in optimizing learning processes and providing more personalized experiences for students (Chen & Wang, 2021). Today, the field is recognized as a dynamic and multifaceted domain that integrates traditional, digital, and data-driven methods, while introducing novel concepts such as pedagogical content knowledge (PCK) and intelligent instructional models (Shulman, 1986).

3. Data and Machine Learning

Data represent a collection of discrete or continuous values that convey information, describing the quantity, quality, fact, statistics, or other basic units of meaning. This information carries no inherent meaning on its own. In fact, they require processing and analysis to be transformed into useful information. In the contemporary world, data are regarded as one of the most valuable resources for decision-making across domains. This is particularly true in fields such as machine learning and data science, where models and algorithms rely on data for training and continuous improvement of performance (Erickson & Nosanchuk, 1992; Yang & Li, 2018).

As discussed earlier, machine learning is a subfield of data science in which algorithms and models are trained using data. Simply explained, machine learning refers to algorithms that identify hidden patterns within data and

AITBSS
AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

leverage these patterns to make predictions or automate decision-making processes.

The terms feature and label are fundamental concepts that play a central role in data modeling and prediction. A feature refers to any measurable or observable variable used to describe, predict, or analyze a phenomenon within a dataset. In the context of machine learning, features are critical elements as they form the input to models and their quality directly affects model performance (Bishop, 2006). Conversely, a label is defined as the dependent or target variable in a dataset. It represents the value assigned to the input instances or features either by humans or a reliable system. Depending on the nature of the problem, labels may take various forms: numerical, categorical, binary, multi-class, or even multi-label (Saidabad et al., 2024; Yao et al., 2024). In certain artificial intelligence approaches, only features are utilized, while in others, both features and labels are employed. For instance, in supervised learning algorithms, as one of the means of a machine learning approach, the primary goal is to model the relationship between features and labels, enabling the prediction of correct labels for unseen data.

Data can be categorized into different types, each playing a distinct role in data science and machine learning. Broadly, they are classified into two major groups: structured data and unstructured data. Structured data refers to information organized in well-defined formats, typically stored in tabular structures that can be easily managed in relational databases (e.g., SQL). Such data usually comes in numerical or textual form with clearly defined attributes, making it highly suitable for computational analysis and modeling (Azad et al., 2020; Eberendu, 2016). For instance, in a training dataset in a shape of a table with rows and columns designed to analyze student academic underperformance, multiple features are recorded, each potentially contributing to the prediction of a student's future academic outcomes. Some of these features are as follows

- Feature 1 (Class Attendance (hours)): The amount of time a student is physically present in class. This feature is a continuous numerical variable and may serve as an indicator of the student's active engagement in the learning process.
- Feature 2 (Homework Completion (hours)): The average number of hours a student spends completing assignments or exercises at home. This numerical feature reflects individual effort and diligence outside the classroom.
- Feature 3 (Parental Satisfaction (scale 1–5)): A qualitative assessment by parents of their child's academic performance within a given period. This discrete numerical variable can reflect the family's supportive or critical attitude toward the student's educational progress.
- Feature 4 (Extracurricular Activities (hours)): The amount of time a student devotes to activities outside the formal curriculum, such as sports, arts, and other non-academic pursuits. This numerical feature may indicate the balance or imbalance between academic and non-academic engagements.
- Feature 5 (Monthly Discipline Score (out of 20)): The school's formal assessment of a student's behavior and adherence to rules, recorded as a numerical value. This feature is generally associated with the student's organization, responsibility, and social interactions.
- Feature 6 (Number of Positive Points Received): Rewards awarded to the student by teachers or school staff for positive behaviors, such as participation, helping others, or academic improvement. This discrete numerical feature can serve as an indirect indicator of individual motivation and effort.
- Feature 7 (Previous Term GPA): The average of a student's grades from the previous term, expressed as a continuous numerical variable. This feature reflects the overall academic performance of the student up to that point.

In this example, the students are located in the rows of the data set, and the features can be found in the columns of the table. Similarly, the label in this example is academic decline or no decline, which is located in the last column. This column indicates whether a student has experienced academic decline. The label serves as the primary target for the machine learning model and is considered the dependent or output variable. Its values typically include "Academic Decline" or "No Decline." Table \ref{tab:students} presents an example of a hypothetical training dataset based on student features, aimed at predicting academic decline. In this table, the data label represents whether each student experiences academic decline or not, based on their features. As observed, the label is qualitative; however, it can be encoded numerically using values such as 0 and 1 for storage and processing in a computer system.

AITBSS
AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

Table 1. Sample training data for predicting academic decline in students

Label	Feature 7	Feature 6	Feature 5	Feature 4	Feature 3	Feature 2	Feature 1	Student
Academic Decline	14	3	17	5	4	8	12	Student 1
No Decline	17	2	18	4	5	10	15	Student 2
No Decline	12	5	15	6	3	4	8	Student 3
No Decline	18	1	19	2	5	15	20	Student 4
Academic Decline	13	4	16	7	4	6	10	Student 5
No Decline	16	2	17	3	5	12	18	Student 6
Academic Decline	17	3	11	6	2	15	14	Student 7
No Decline	19	4	15	4	4	11	16	Student 8
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.

4. Learning Paradigms

Machine learning, as one of the most important branches of AI, can be categorized in multiple ways. One of the most widely used classifications divides learning into four main types: supervised learning, semi-supervised learning, unsupervised learning, and reinforcement learning (Rajoub, 2020). This classification is primarily based on the availability of labeled data and the manner in which the model interacts with the data.

For instance, in a dataset aimed at predicting whether students will experience academic decline, supervised learning involves a dataset that includes both the relevant features for each student, such as class attendance, homework completion, parental satisfaction, and extracurricular activities, and the corresponding labels indicating the student's final status, which includes academic decline or no decline. A machine learning model uses these labeled data to identify patterns and relationships between the features and the label. Once trained, the model can predict the probable academic outcomes for new students based solely on their features, without requiring prior knowledge of their labels.

In contrast, unsupervised learning operates without any information regarding the labels of the samples. This approach relies solely on the observed features. In this setting, the learning algorithm seeks to identify underlying structures or hidden patterns in the data without reference to labels. For example, it may cluster students based on similarities in their behaviors and attributes. Such clustering can assist analysts in identifying groups of students with similar characteristics and designing targeted educational interventions for each group if necessary.

Each of these methods holds its own significance. Therefore, it is essential to evaluate both approaches in relation to a given dataset. However, their application depends not only on the available data, but also on the nature of the problem and the specific outcomes that need to be predicted. On the other hand, understanding these classifications is important not only in research contexts but also in practical applications. This helps us in selecting the appropriate learning approach, which can significantly impact the accuracy and efficiency of the resulting models (Love, 2002).

Although supervised, semi-supervised, unsupervised, and reinforcement learning are recognized as the four main categories in machine learning, this work focuses primarily on supervised and unsupervised learning. This choice is motivated by the widespread application and fundamental role of these two approaches in addressing diverse problems in data science and artificial intelligence. It is worth noting that the other learning paradigms also have significant applications in specific contexts. Semi-supervised learning, by leveraging both labeled and unlabeled data, seeks to combine the advantages of supervised and unsupervised approaches. Reinforcement learning, on the other hand, trains models to make optimal decisions in dynamic environments through feedback from the environment.

4.1. Supervised Learning

In {supervised learning}, the dataset consists of a collection of labeled samples $(\{(x_i, y_i)\}_{i=1}^{N})$. Each element (x_i) in this set is represented as a feature vector. A feature vector is an array in which each dimension

AITBSS AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

$(j = 1, \ldots, D)$ contains a value describing a specific aspect of the sample. This value is referred to as a {feature} and is denoted by $(x^{(j)})$ .For example, in the task of predicting whether a student will experience academic decline, if each sample (x) in the dataset represents an individual student, then the first feature $(x^{(1)})$ may correspond to class attendance (hours), the second feature $(x^{(2)})$ to homework completion (hours), the third feature $(x^{(3)})$ to parental satisfaction, and so on. Across all samples in the dataset, the feature at position (j) in the feature vector consistently represents the same type of information. This implies that if $(x_i^{(2)})$ in a sample (x_i) represents class attendance (hours), then $(x_k^{(2)})$ in every other sample (x_k) will also correspond to class attendance (hours), ensuring that this convention holds for all $(k = 1, \ldots, N)$. The label (y_i) can belong to a finite set of classes $(\{1, 2, \ldots, C\})$, a real number, or a more complex structure such as a vector, matrix, tree, or graph. In this study, it is assumed that y_i is either a member of a finite set of classes or a real number. For example, if the samples represent students and the task is to predict academic decline based on the features described above, there will be two classes (academic decline and no decline), which can be represented and stored in a computer as a binary variable (0 and 1) (Tshidi, 2022). The objective of a supervised learning algorithm is to construct a model from a dataset that receives a feature vector (x) as input and produces information that allows the inference of the corresponding label for that feature vector. For instance, a model trained on a dataset of students can take a feature vector. representing an individual student and predicting the probability of that student belonging to a particular class (academic decline or no decline) (Dridi, 2021; Nasteski, 2017).

4.2. Unsupervised Learning

Unsupervised learning is a branch of machine learning in which the model is trained using unlabeled data. In other words, in this approach, no prior information regarding the data labels or expected outputs for each input is available. The objective of the model is to discover hidden structures, patterns, and relationships in the data without any external supervision (Bishop, 2006). In this type of learning, algorithms aim to analyze the data to extract its intrinsic features and perform tasks such as grouping, dimensionality reduction, or identifying relationships among data points. This process can lead to the identification of categories or clusters of similar data, the reduction of the number of variables required to describe the data, or the detection of anomalies within the dataset (Barlow, 1989; Ghahramani, 2003).

Training data in machine learning consists of labeled examples which the model uses to learn patterns and make predictions. If the training data are biased, insufficient, or not representative of the real-world scenario, the model's predictions and decisions will reflect these shortcomings and, as a result, potentially lead to suboptimal outcomes. For instance, in personalized learning systems, where a machine learning model adapts its teaching strategies based on student performance, having diverse and representative data is essential for the model to accurately predict and respond to students' learning behaviors and needs (Naeem et al., 2023).

An important aspect of training data in educational settings is their alignment with learning objectives. The data should represent diverse educational experiences and teaching methods so that the model can effectively simulate appropriate instructional strategies and performance assessments. For instance, if a machine learning model is designed to evaluate students' reading skills in English, the training dataset should include samples of various texts and activities that assess students' abilities across different linguistic dimensions, such as grammar, reading comprehension, and vocabulary. If the training data are limited to a specific type of text or a narrow range of student proficiency, the model will be unable to make accurate predictions for students with different characteristics (Rahman et al., 2024; Zeller & Riese, 2025). In this type of learning, the available data may be unlabeled, or only a portion of the data may contain labels , while the remaining data includes only students' behavioral features. In other words, the status of academic decline for these students is unknown. The primary objective in this type of learning is to uncover hidden structures or natural patterns within the data. For instance, students can be grouped into distinct clusters based on similarities in features. These clusters can reflect different learning

AITBSS AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

patterns, levels of engagement, or motivation among students using clustering methods.

As an illustration, one cluster might consist of students with high attendance, substantial homework completion, and high parental satisfaction, likely corresponding to students with strong academic performance. Identifying such groups, even without knowing the final label, can assist educational administrators and counselors in designing targeted interventions to improve student outcomes.

Ultimately, how well a model performs depends on the training data used in teaching and educational applications. This shows how important it is to gather and choose the appropriate data . It can help both teachers and students have better educational experiences when paired with suitable datasets.

5. Data Selection for Enhancing the Learning Process with a Pedagogical Approach

In the field of machine learning, numerous methods have been proposed to enhance the model training process with a focus on selecting or designing effective training data. These approaches primarily aim to accelerate, improve, and reduce the cost of learning by utilizing more informative data. Notable methods in this context include passive learning and active learning. In passive learning, the model simply receives a pre-prepared dataset and is trained on it without any targeted intervention. Although this approach is straightforward and widely used, it often requires large volumes of data and can be inefficient in many cases.

On the other hand, in active learning, the learner dynamically queries a predictive model for labels on specific data points that are considered most uncertain or informative. While this approach is relatively more efficient, it still requires repeated and potentially costly interactions with the predictive model, which may not be feasible in real-world scenarios with limited resources or time.

Among these approaches, however, machine teaching holds a distinct and advantageous position. Rather than relying on random selection or query-based data acquisition, machine teaching organizes the learning process in a deliberate and targeted manner, guided by an informed teacher.

Figure 1 provides a conceptual comparison among four key approaches to interacting with training data (Liu et al., 2021). In passive learning, the model merely receives an existing dataset and is trained in a passive manner. Active learning, by introducing the ability to query a predictive model, attempts to identify more informative data points; however, it still relies on numerous and potentially costly interactions.

In contrast, machine teaching leverages the knowledge of a teacher to design an optimized training dataset, which can ensure effective learning even with a minimal number of examples. This approach interacts with the learner only once, yet provides precise, compact, and targeted information.

Moreover, an advanced variant known as Iterative Machine Teaching has been proposed, in which the interaction between the teacher and the learner occurs iteratively over multiple stages. This process allows continuous refinement of the training and gradual improvement of the model's performance. Nevertheless, in both cases, the key advantage of machine teaching over other methods lies in its approach: the training process is designed not only to be data-driven but also knowledge-driven and structured. This characteristic makes it an efficient approach for educational scenarios with limited resources.

AITBSS
AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

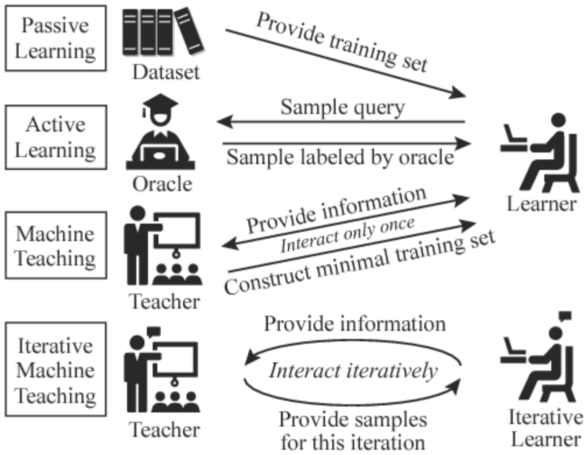

Figure 1. A Conceptual Comparison of Different Approaches to Interacting with Training Data in Machine Learning

6. Machine Teaching

Machine teaching is a process in which experts select optimized and targeted training data for training machine learning models. In this approach, rather than relying on a large volume of raw data, only key and informative data points are chosen to accelerate and enhance learning (Zhu, 2018). This process can be mathematically formulated as follows:

$$\min_{D \subset X} \{ \mathrm{L}(f_D, \theta) + \lambda \cdot \mathrm{Effort}(D) \}, \qquad (1.1)$$

where,

- D denotes the selected dataset drawn from the full data space X.
- f_D represents the model trained using the dataset D.
- $\mathrm{L}(f_D, \theta)$ is the loss function that measures the discrepancy between

the model's predictions and the true labels, where θ denotes the model parameters.

- $\mathrm{Effort}(D)$ quantifies the human and computational effort required to select

and process D.

- λ is a regularization parameter that balances model accuracy against the effort

involved in selecting the data.

This formulation essentially describes the primary objective of machine teaching, which is to enhance the learner's performance while minimizing the effort involved in selecting the training data. The main focus is on choosing key and informative data.

points rather than relying on a large volume of raw data, enabling the model to learn optimally. Machine teaching, therefore, aims for more efficient learning with limited resources both human and computational by selecting data that accelerates the learning process and improves model performance, while simultaneously reducing the cost and time required for data preparation.

The parameter λ in this equation, which balances model accuracy against the effort required for data selection, specifically emphasizes the importance of this trade-off. In other words, if the $\mathrm{Effort}(D)$ becomes excessively large, model accuracy may be compromised; conversely, if the effort is insufficient, there may not be enough data to effectively train the model. Ultimately, this formulation clearly illustrates the objective of machine teaching as an optimization process, where the goal is to achieve the best possible performance with minimal cost and resource expenditure.

6.1. Pedagogy and Pedagogical Content Knowledge (PCK)

Pedagogy, as an academic discipline, requires not only deep content knowledge but also teaching skills, an accurate understanding of learners' needs, and the ability to

AITBSS AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

adapt instructional methods to the specific characteristics of students (Freire, 2020; Kos, 2025). Shulman (Shulman, 1987), in his definition of pedagogy, introduced the concept of Pedagogical Content Knowledge (PCK), which comprises a combination of the following four key components:

- Content Knowledge (CK): Refers to possessing extensive and deep mastery of the subject matter being taught or studied. This knowledge goes beyond superficial information and includes an understanding of fundamental concepts, theories, and the relationships among them.
- Pedagogical Knowledge (PK): Encompasses the ability to employ appropriate teaching methods and to present content in various ways that are comprehensible to learners. This includes the use of examples, diagrams, analogies, and other instructional tools to effectively convey concepts.
- Understanding of Students' Conceptions of the Subject and Learning: Refers to recognizing how students perceive concepts and the preconceptions they hold about learning. This understanding helps instructors tailor their teaching strategies to meet learners' needs.
- Knowledge of Educational Contexts: Involves awareness of the environments, conditions, and factors that influence the teaching and learning process. This includes understanding educational systems, policies, school culture, and local community contexts.
- Knowledge of the Purpose of Education: Pertains to understanding the philosophy, long-term goals, and expected outcomes of the educational process. This includes recognizing the role of education in individual and social development, promoting values, and preparing individuals for life in a complex world.

These four components enable teachers to select the most appropriate teaching methods tailored to the characteristics of their students and the subject matter. In other words, PCK allows teachers to deliver instruction in a personalized and effective manner and optimize their teaching strategies based on careful analysis of students' ability, characteristics, and needs (Jiménez Sierra et al., 2023; Van Driel & Berry, 2010).

In recent years, combining machine teaching systems with PCK has become a new way to improve teaching processes (Bengsch, 2024; Su et al., 2024; Tang et al., 2023; Zeroual et al., 2024). The main aim of this combination is to use human teaching insights along with the accuracy of data patterns to enhance instruction. In this setup, teachers use their PCK to help choose educational data. This means that selecting instructional data considers not just general traits and set algorithms but also the specific needs of students and the content's features. In other words, machine teaching algorithms can use this data to tailor instruction and improve both the choice of learning materials and the way content is delivered to meet individual student needs. For instance, in a personalized learning system where machine learning models employ sophisticated algorithms to analyze students' learning behaviors and performance, instructors can leverage this information to guide the selection of appropriate content and adjust teaching strategies. This process, optimized through machine teaching and pedagogical knowledge, ultimately fosters an interactive, efficient, and more effective learning experience. Unlike conventional machine learning, where the learner autonomously identifies patterns from the input dataset, machine teaching defines an active role for the teacher. The teacher not only collects data but also purposefully and intelligently determines which information should be presented to maximize the learner's effectiveness.

To illustrate this concept, consider the previously mentioned example of predicting student academic decline. Suppose a teacher has access to a dataset containing students' behavioral and academic information, including class attendance, number of homework assignments completed, parental satisfaction, extracurricular activities, discipline scores, and so on. In a traditional machine learning system, these data would be fed into the model without significant human intervention to identify patterns. However, in machine teaching, the instructor may analyze the data and determine that, for example, parental satisfaction has little impact on predicting academic decline. Consequently, the teacher may choose to exclude this feature from the training dataset while emphasizing more influential features such as class attendance or the number of rewards received.

This process of selecting, removing, or emphasizing specific features constitutes an optimized design of the educational dataset, leading to improved model efficiency, reduced training costs, and faster, more accurate learning. In essence, machine teaching seeks to transfer maximum knowledge to the learner using minimal data.

E-ISSN: 3041-9433

7. The Role of Machine Teaching in Pedagogy

In this section, we examine the applications of machine teaching in enhancing the teaching process from a pedagogical perspective.

7.1. Human-Guided Data Selection

In machine teaching, the role of a teacher as a selector of training data is highly significant. In these models, the teacher or expert selects key and critical data points intelligently and goal-oriented from the datasets for training the system. These data points typically contain important and informative content that the teacher recognizes this importance and can effectively help the system learn more quickly and accurately. In other words, the human expert leverages their experience and knowledge to choose data that supports the learning algorithm (Zeroual et al., 2024). By introducing an additional constraint, it is ensured that the selected data aligns with pedagogical principles. This optimal data can be formulated as follows:

$$D_{\text{optimal}} = \arg\min_{D\subset X} \{\mathrm{L}(f_D,\theta) + \lambda \cdot \mathrm{Effort}(D)\} \quad \text{s.t.} \quad PCK(D) \geq \tau, \tag{1.2}$$

where PCK(D) is a function that measures the degree to which the selected data comply with the established pedagogical principles, and τ is a threshold that ensures the quality and relevance of the data. In this equation:

- (D_{optimal}) is the optimal dataset selected by the machine teaching algorithm.
- $(\mathrm{L}(f_D,\theta))$ is the model's prediction error function, representing the quality of data selection from the learning algorithm's perspective.
- $(\mathrm{Effort}(D))$ denotes the amount of effort exerted by the teacher in selecting the data.
- $\left(PCK(D)\right)$ is a function that measures the alignment of the selected data with pedagogical principles.
- (τ) is a threshold that ensures the compliance and quality of the data.

The main goal of this approach is to ensure that the data selection process is directed not only by machine learning criteria but also by pedagogical principles. This means that the data chosen for training the model should align with learners' Requirements and specific pedagogical characteristics, which should be detected by the teacher. To achieve this, an additional constraint is incorporated into the process to guarantee that the selected data are consistent with pedagogical principles.

In (1.1), the objective is to select a set of data that minimizes both the prediction error and the teacher's effort in data selection, while ensuring that the data also satisfy pedagogical requirements.

This process specifically enables teachers to select data based on their teaching knowledge and experience, tailoring it to the learning demands of individual students or groups of students. For instance, a teacher might select data that specifically focuses on solving complex problems related to fundamental concepts in a math class. This allows students who are struggling with those concepts to make progress through accurate and targeted data. Additionally, with this personalized learning system, other students can dedicate their time to focusing on what they need to do.

For instance, in the process of instructing students on fractions by a teacher, the main challenge is that students are at varying levels of understanding. It means that some may require an initial visual introduction, while others achieve a deeper understanding through solving more complex problems. In traditional methods, all students receive the same set of data, and the teacher may not necessarily take into account students' abilities, levels of understanding, or other important factors. However, in a machine teaching approach, the teacher can purposefully select training data based on the individual states of each student. For instance, if the teacher identifies that a student has a weak understanding of fractions, machine teaching can provide data that includes visual representations or other possible support tools to strengthen their intuition. Conversely, for a student who has a basic grasp of fractions but struggles with converting them to decimals, the system can offer exercises specifically targeting these conversions.

In this way, machine teaching, by selecting training data tailored to individual needs, not only facilitates the learning process but also ensures that each learner receives instruction aligned with their level of understanding, enabling them to make meaningful progress. One of the primary applications of this approach is in personalized learning systems, where data are dynamically selected and presented based on each student's individual requirements. For instance, in an AI-based learning system, the algorithm can choose training data by evaluating students' past performance through the knowledge and information of the

E-ISSN: 3041-9433

teacher. However, it is crucial that this selection also considers pedagogical characteristics. This may include selecting data designed to align with different learning styles, such as visual, auditory, or kinesthetic learning.

In this context, emerging technologies can provide helpful tools that enable teachers to offer a more effective and tailored learning experience for students and make the learning process more effective (Zhu, 2013, 2018).

8. Dynamic Learner Training with Internal States

In the field of machine teaching, selecting an optimal dataset for learning remains a fundamental challenge. However, a recent study entitled "Teaching to Learn: Sequential Teaching of Agents with Inner States" (Celikok et al., 2020) introduces a novel approach that goes beyond static teaching by considering the learner as a dynamic agent with internal states.

learner's internal state evolves throughout the interaction with the teacher. Unlike traditional methods that assume learners have fixed biases, in this new model, the internal states of the learner can change and subsequently influence their future learning. This feature enables the teaching of learners who are capable of generalizing to new datasets, which is not possible with traditional approaches.

One of the key contributions of this study is the distinction between manipulative learning and general learning.

In manipulative learning, the teacher can steer the learner toward a specific model by withholding certain data. This approach, which raises both ethical and practical concerns, may lead to excessive learner dependence on the teacher. In contrast, general learning aims to improve the learner in a way that allows them to acquire appropriate models even in the teacher's absence. Using an optimal control approach, the authors provide a framework to model these two scenarios. This framework conceptualizes the teaching problem as a binary game between two agents (the teacher and the learner). In this model, the learner's internal state, which includes bias settings and model hyperparameters evolves over time. These changes result from the teacher's actions and can lead to improved learning outcomes in the future. The results demonstrate that the teacher can enhance the learner's internal state by optimizing interactions, thereby facilitating more efficient learning (Albrecht et al., 2024). This framework can serve as a theoretical foundation for developing new numerical methods for solving learning models. In particular, leveraging the idea of evolving internal states in the design of numerical optimization techniques may improve the performance of proposed models. This aspect of the research provides a promising direction for future studies (Persson, 2024).

One of the aspects of fundamental challenges in education lies in the diversity of students' prior knowledge, learning styles and individual needs. Machine teaching can address this issue by analyzing previous data and students' performance to select a set of targeted and personalized learning materials (Kaswan et al., 2024). For instance, in mathematics courses, the system can provide supplementary exercises tailored to each student's requirement through a teacher, based on the common mistakes they make when solving problems. This personalization enhances students' motivation and engagement, while also improving their long-term performance. The process is directly aligned with the principles of individualized instruction and learning based on each student's specific needs. In fact, personalized learning is one of the fundamental principles of an effective education. In this regard, a key capability of machine teaching lies in its ability to dynamically adapt instructional content to the learner's requirements. Such models can identify and deliver appropriate learning materials based on each student's past performance data, knowledge level, and learning style (Zohuri & Mossavar-Rahmani, 2024).

Let P_{learner} represent the distribution of student characteristics, such as past knowledge, learning preferences and typical personal difficulties. A teacher's primary objective is to select a set of learning resources, or better to say data, which are suitable for each student and best suit their individual characteristics. In order to accomplish this, intelligent data selection is used in machine teaching algorithms to introduce personalized educational content. The following formula can be used to formalize this idea:

$$\min_D \mathrm{E}_{x\sim P_{\text{learner}}}\left[\mathrm{L}(f_D(x), y)\right] + \lambda \cdot \mathrm{Effort}(D), \qquad (1.3)$$

In this formulation, D denotes the set of instructional data to be individually tailored for each learner. Also, the loss term $\mathrm{L}(f_D(x), y)$ measures the error between the model's prediction and the actual output y. Furthermore, the distribution P_{learner} characterizes the learner's individual profile. For the instructional effort involved in

AITBSS
AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

selecting appropriate materials and maintaining consistency with pedagogical principles, the regularization term $\lambda \cdot \text{Effort}(D)$ accounts .

This approach makes sure that the selected materials are not only sound from a pedagogical perspective but also well matched to the learner's individual profile. It also allows the system to keep updating the instructional content in a flexible way as the learner progresses (Laak & Aru, 2024).

8.1. Advantages of Personalizing the Learning Process

Personalization of the learning process has significant advantages, especially compared to traditional education. In traditional education, all students are taught in a similar way, even if this method does not match the specific needs of each individual. However, in personalized methods, each student receives educational content specifically tailored to his or her abilities and needs. This process can increase students' motivation to learn and improve their academic performance.

One of the practical examples in this regard is in language teaching. In this course, students may have different challenges in different skills such as grammar, vocabulary, and pronunciation. Personalized systems can accurately identify each student's common problems and provide grammar or text exercises tailored to their level (Dandachi, 2024). Furthermore, using machine teaching to personalize the learning process can effectively leverage automated capabilities to deliver educational content at the right times. For instance, when a student makes a mistake in solving a particular problem, the system can automatically send them additional exercises tailored to that mistake (Jung, 2024).

8.2. Enabling Timely and Adaptive Feedback Delivery

Timely and accurate feedback is one of the key factors in improving learning. Machine teaching, using active learning algorithms, can quickly identify students' mistakes and weaknesses and provide teachers with suggestions for correction and improvement. In fact, a machine learning-based system can guide students in the right direction by providing suggestions and additional explanations while solving a problem. This method makes the learning process dynamic and adaptive, and the educational impact is greatly increased. Therefore, one of the outstanding advantages of combining machine learning with PCK is the ability to provide immediate and consistent feedback. In this setting, a reinforcement learning-based model can adjust the data selection process in real time:

$$R(s,a) = \Delta \mathrm{L}(f_D(s), y) - \gamma \cdot \text{Effort}(D), \quad (1.4)$$

where $R(s,a)$ is the reward signal for state s and action a and γ balances the importance of improving accuracy against reducing effort. This gradual feedback loop helps to continuously improve both the training data and the learning model.

8.3. Practical applications during the courses

To better understand the role of machine teaching in pedagogy, we present some practical and creative examples that demonstrate the potential of this technology in transforming the education process:

- Mathematics Classes: Machine teaching systems can be employed to create dynamic, personalized learning pathways. By analyzing students' real-time behavior and performance, such systems can recommend problems and challenges tailored to each learner's level and style. For instance, an AI-based system might identify a student's weaknesses in specific concepts (e.g., fractions) and generate interactive games or visual puzzles that gradually increase in complexity, thereby enhancing motivation and engagement.
- Technical and Vocational Training: Virtual reality (VR) simulators combined with machine teaching can create interactive and intelligent work environments, allowing students to engage with complex real-world scenarios. For instance, in automotive repair training, a machine teaching system can analyze a student's performance within the simulation to provide immediate feedback, identify mistakes, and offer personalized guidance, such as suggesting the appropriate tools or novel repair techniques.
- Language teaching: Using natural language processing (NLP) models to create intelligent, interactive conversations with students. These systems can analyze a student's accent, grammar, and vocabulary and design creative speaking or writing exercises, such as group storytelling or simulations of real-life situations (such as job

interviews), that reinforce language skills in an engaging and practical way.

Each of these applications demonstrates the potential of machine learning to improve the learning experience from a pedagogical perspective and adapt to the diverse needs of students.

9. Implementation Challenges and Considerations

Despite the many benefits that come with using machine learning techniques in pedagogy, this approach also faces various challenges that need to be managed effectively. In this regard, it is essential to pay attention to the following:

- Quality of Educational Data: One of the main challenges in applying machine teaching methods in pedagogy is ensuring the accuracy, diversity, and impartiality of the educational data. Inaccurate or biased data can directly undermine the quality of student learning and hinder knowledge transfer. To ensure reliable outcomes, it is essential to carefully evaluate and correct the data sources used to train the models.
- Teacher Training: Another key challenge is the need to enhance teachers' knowledge and skills in using machine teaching tools. Many educators, due to limited training or unfamiliarity with these technologies, are unable to fully leverage the potential of machine teaching. Therefore, investing in professional development courses and providing supportive resources can significantly improve the adoption of these technologies.
- Integration with Traditional Approaches: Integrating modern technologies into educational settings that have traditionally relied on direct instruction and human interaction can be complex. Achieving effective integration requires close collaboration among IT specialists, educational researchers, and teachers to ensure that innovative approaches complement and optimize traditional teaching methods.

10. Conclusion

Machine teaching, as an advanced branch of machine learning, holds significant potential for enhancing educational and pedagogical processes. This approach leverages intelligent algorithms and predictive models to enable teachers to dynamically and individually tailor instructional content for each student. Rather than merely providing data to the system, teachers in machine teaching actively design educational materials in a way that supports each student's learning. Through this process, teachers can use data and advanced analytics to anticipate students' learning trajectories and adjust instructional content according to their individual needs. For example, if the system detects that a student is struggling with a particular concept, machine teaching algorithms can automatically modify the learning materials to help the student grasp the concepts more efficiently and accurately. This approach allows teachers to actively and intelligently personalize the learning process for each student. Consequently, educational systems can continuously monitor student progress and provide more precise feedback.

Overall, machine teaching not only contributes to improving instructional quality but also empowers teachers to leverage analytical data to optimize their teaching, making learning a more personalized and effective experience. Machine teaching, as an advanced branch of machine learning, holds significant potential for enhancing educational and pedagogical processes. This approach leverages intelligent algorithms and predictive models to enable teachers to dynamically and individually tailor instructional content for each student. Rather than merely providing data to the system, teachers in machine teaching actively design educational materials in a way that supports each student's learning.

Authors' Contributions

Farzaneh Nasresfahani conceived the study idea. MohammadReza Eslahchi and Farzaneh Nasresfahani developed the theoretical framework, and verified the analytical methods. They also drafted the initial version of the manuscript. Mohammadreza Eslahchi provided guidance and support in writing and took the lead in revising the manuscript. All authors critically reviewed the manuscript for important intellectual content, approved the final version to be published, and agree to be accountable for all aspects of the work.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

E-ISSN: 3041-9433

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

We would like to thank the research and technology council of Farhangian University for financial support through this investigation under post doctorate project No 50000.17473.120.

Declaration of Interest

The authors report no conflict of interest.

Funding

This research was financially supported by the Research and Technology Council of Farhangian University under Postdoctoral Project No. 50000.17473.120.

Ethics Considerations

This article does not contain any studies involving human participants or animals performed by any of the authors. Therefore, ethical approval was not required for this study

References

Albrecht, S. V., Christianos, F., & Schäfer, L. (2024). *Multi-agent reinforcement learning: Foundations and modern approaches*. MIT Press. https://www.marl-book.com/

Anastasio, E. J., & Alderman, D. L. (1973). Evaluation of the Educational Effectiveness of PLATO and TICCIT. https://eric.ed.gov/?id=ED088934

Angeioplastis, A., Aliprantis, J., Konstantakis, M., & Tsimpiris, A. (2025). Predicting Student Performance and Enhancing Learning Outcomes: A Data -Driven Approach Using Educational Data Mining Techniques. *Computers*, *14*(3), 83.

Asnicar, F., Thomas, A. M., Passerini, A., Waldron, L., & Segata, N. (2024). Machine learning for microbiologists. *Nature Reviews Microbiology*, *22*(4), 191-205. https://www.nature.com/articles/s41579-023-00984-1

Azad, P., Navimipour, N. J., Rahmani, A. M., & Sharifi, A. (2020). The role of structured and unstructured data managing mechanisms in the Internet of things. *Cluster computing*, *23*, 1185-1198. https://link.springer.com/article/10.1007/s10586-019-02986-2

Barlow, H. B. (1989). Unsupervised learning. *Neural computation*, *1*(3), 295-311. https://psycnet.apa.org/record/1991-03394-001

Bengsch, G. (2024). *Redefining traditional pedagogy: The integration of machine learning in the contemporary language education classroom*. IGI Global Scientific Publishing. https://www.researchgate.net/publication/379156463_Redefining_Traditional_Pedagogy_The_Integration_of_Machine_Learning_in_the_Contemporary_Language_Education_Classroom

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer. https://books.google.com/books/about/Pattern_Recognition_and_Machine_Learning.html?id=kTNoQgAACAAJ

Block, J. H., & Burns, R. B. (1976). Mastery learning. *Review of research in education*, *4*, 3-49. https://journals.sagepub.com/doi/10.3102/0091732X004001003

Celikok, A., Murena, P. A., & Kaski, S. (2020). Teaching to Learn: Sequential Teaching of Agents with Inner States. *Journal Name*, *Volume*(Issue). https://arxiv.org/abs/2009.06227

Chatti, M. A., Dakova, S., Thüs, H., & Schroeder, U. (2013). Tag-based collaborative filtering recommendation in personal learning environments. *IEEE Transactions on learning technologies*, *6*(4), 337-349. https://www.researchgate.net/publication/260722681_Tag-Based_Collaborative_Filtering_Recommendation_in_Personal_Learning_Environments

Chen, X., & Wang, Y. (2021). Artificial Intelligence in Education: A Review of Pedagogical AI Applications. *Journal of Educational Technology*.

Dandachi, I. E. (2024). AI-powered personalized learning: Toward sustainable education. In *Navigating the intersection of business, sustainability and technology* (pp. 109-118). Springer. https://link.springer.com/chapter/10.1007/978-981-99-8572-2_5

Díaz, B., & Nussbaum, M. (2024). Artificial intelligence for teaching and learning in schools: The need for pedagogical intelligence. *Computers & Education*, 105-171. https://www.sciencedirect.com/science/article/abs/pii/S036013152400085X

Dridi, S. (2021). Supervised learning-a systematic literature review. *preprint*, *Dec*. https://www.researchgate.net/publication/381415563_Supervised_Learning_-_A_Systematic_Literature_Review

Eberendu, A. C. (2016). Unstructured Data: an overview of the data of Big Data. *International Journal of Computer Trends and Technology*, *38*(1), 46-50. https://www.researchgate.net/publication/309393428_Unstructured_Data_an_overview_of_the_data_of_Big_Data

Erickson, B., & Nosanchuk, T. (1992). *Understanding data*. McGraw-Hill Education (UK). https://books.google.com/books/about/Understanding_Data.html?id=kS-2AAAAIAAJ

Esomonu, N. P.-M. (2025). Utilizing AI and Big Data for Predictive Insights on Institutional Per formance and Student Success: A Data-Driven Approach to Quality Assurance. *AI and Ethics, Academic Integrity and the Future of Quality Assurance in Higher Education*, 29.

Freire, P. (2020). Pedagogy of the oppressed. In *Toward a sociology of education* (pp. 374-386). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780429339530-34/pedagogy-oppressed-paulo-freire

Froebel, F. (1886). *The education of man* (Vol. 5). A. Lovell. https://books.google.com/books/about/The_Education_of_Man.html?id=ADkXAAAAYAAJ

Ghahramani, Z. (2003). Unsupervised learning. In *Summer school on machine learning* (pp. 72-112). Springer. https://link.springer.com/chapter/10.1007/978-3-540-28650-9_5

Gómez, S., Zervas, P., Sampson, D. G., & Fabregat, R. (2014). Context-aware adaptive and personalized mobile learning delivery supported by UoLmP. *Journal of King Saud*

University-Computer and Information Sciences, *26*(1), 47-61. https://www.researchgate.net/publication/259157646_Context-aware_adaptive_and_personalized_mobile_learning_delivery_supported_by_UoLmP

Goodson, I. F. (2017). *Learning, Curriculum, and Pedagogy*. Routledge.

Grant, P., & Basye, D. (2014). *Personalized learning: A guide for engaging students with technology*. International Society for Technology in Education. https://books.google.com/books/about/Personalized_Learning.html?id=96apCgAAQBAJ

Gutek, G. L. (1995). *Renaissance education and educational thought*. Springer.

Idowu, E. (2024). Personalized Learning: Tailoring Instruction to Individual Student Needs. https://www.researchgate.net/publication/385773808_Personalized_Learning_Tailoring_Instruction_to_Individual_Student_Needs

Jia, H., & Zhu, X. (2020). Efficient Data Selection for Machine Teaching with Deep Learners. *Proceedings of ICML*.

Jiménez Sierra, Á. A., Ortega Iglesias, J. M., Cabero-Almenara, J., & Palacios-Rodríguez, A. (2023). Development of the teacher's technological pedagogical content knowledge (TPACK) from the Lesson Study: A systematic review. https://www.researchgate.net/publication/368680867_Development_of_the_teacher%27s_technological_pedagogical_content_knowledge_TPACK_from_the_Lesson_Study_A_systematic_review

Jung, I. (2024). Personalized education for all: the future of open universities. *Open Praxis*, *16*(1), 24-36. https://www.aacu.org/oer/research/personalized-education-for-all-the-future-of-open-universities

Kaswan, K. S., Dhatterwal, J. S., & Ojha, R. P. (2024). AI in personalized learning. In *Advances in technological innovations in higher education* (pp. 103-117). CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003376699-9/ai-personalized-learning-kuldeep-singh-kaswan-jagjit-singh-dhatterwal-rudra-pratap-ojha

Kayal, A. (2024). Transformative Pedagogy: A Comprehensive Framework for AI Integration in Education. In *Explainable AI for Education: Recent Trends and Challenges* (pp. 247-270). Springer. https://www.researchgate.net/publication/386190559_Transformative_Pedagogy_A_Comprehensive_Framework_for_AI_Integration_in_Education

Kos, Ž. (2025). Critical pedagogy. In *Encyclopedia of Diversity* (pp. 1-13). Springer. https://www.researchgate.net/publication/387753896_Critical_Pedagogy

Laak, K.-J., & Aru, J. (2024). AI and personalized learning: bridging the gap with modern educational goals. *arXiv preprint arXiv:2404.02798*. https://arxiv.org/abs/2404.02798

Liu, J., & Zhu, X. (2017). Optimal Teaching for Bayesian Learners in the Exponential Family. *Conference on Neural Information Processing Systems (NeurIPS)*.

Liu, W., Liu, Z., Wang, H., Paull, L., Schölkopf, B., & Weller, A. (2021). Iterative teaching by label synthesis. *Advances in Neural Information Processing Systems*, *34*, 21681-21695. https://arxiv.org/abs/2110.14432

Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic bulletin & review*, *9*(4), 829-835. https://link.springer.com/article/10.3758/BF03196342

Mayer, R. E. (2014). *Multimedia Learning: Principles and Applications*. Cambridge University Press. https://books.google.com/books/about/Multimedia_Learning.html?id=5g0AM1CHysgC

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, *27*(4), 12-12. https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904

Miliband, D. (2006). Choice and voice in personalised learning. *Schooling for tomorrow: Personalising education*, 21-30. https://www.semanticscholar.org/paper/Choice-and-Voice-in-Personalised-Learning-Miliband/694f1f228c5f8611a60cbb90f2d4546211c0a90c

Minsky, M., & Papert, S. (1969). *Perceptrons*. MIT Press. https://books.google.com/books/about/Perceptrons.html?id=U-O9BHlTvJIC

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill. https://books.google.com/books/about/Machine_Learning.html?id=ifdcswEACAAJ

Morison, S. (1932). *John Amos Comenius and the beginnings of modern education*. Oxford University Press. https://books.google.com/books/about/Comenius_and_the_Beginnings_of_Education.html?id=cg07siz_sm0C

Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An unsupervised machine learning algorithms: Comprehensive review. *International Journal of Computing and Digital Systems.* https://www.researchgate.net/publication/368983958_An_Unsupervised_Machine_Learning_Algorithms_Comprehensive_Review

Nafea, A. A., Alameri, S. A., Majeed, R. R., Khalaf, M. A., & Al-Ani, M. M. (2024). A short review on supervised machine learning and deep learning techniques in computer vision. *Babylonian Journal of Machine Learning*, 48-55. https://mesopotamian.press/journals/index.php/BJML/article/view/261

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons. b*, *4*(51-62), 56. https://www.researchgate.net/publication/328146111_An_overview_of_the_supervised_machine_learning_methods

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., Lopez Garcia, A., Heredia, I., Malík, P., & Hluchý, L. (2019). Machine learning and deep learning frameworks and libraries for large- scale data mining: a survey. *Artificial Intelligence Review*, *52*(1), 77-124. https://link.springer.com/article/10.1007/s10462-018-09679-z

Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. https://unesdoc.unesco.org/ark:/48223/pf0000366994

Persson, H. (2024). Deep Reinforcement Learning for Multi-Agent Path Planning in 2D Cost Map Environments: using Unity Machine Learning Agents toolkit. In.

Pestalozzi, J. H. (1898). *How Gertrude teaches her children: An attempt to help mothers to teach their own children and an account of the method*. CW Bardeen. https://books.google.com/books/about/How_Gertrude_Teaches_Her_Children.html?id=9aOcAAAAMAAJ

Rahman, S. I., Ahmed, S., Fariha, T. A., Mohammad, A., Haque, M. N. M., Chellappan, S., & Noor, J. (2024). Unsupervised machine learning approach for tailoring educational content to individual student weaknesses. *High-Confidence Computing*, *4*(4), 100-228. https://www.sciencedirect.com/science/article/pii/S266729522400031X?dgcid=rss_sd_all

AITBSS
AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433

Rajoub, B. (2020). Supervised and unsupervised learning. In *Biomedical signal processing and artificial intelligence in healthcare* (pp. 51-89). Elsevier. https://dokumen.pub/biomedical-signal-processing-and-artificial-intelligence-in-healthcare-developments-in-biomedical-engineering-and-bioelectronics-1nbsped-0128189460-9780128189467.html

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386-408. https://psycnet.apa.org/record/1959-09865-001

Rousseau, J. J. (1762). *Emile: or, On Education*. Basic Books. https://books.google.com/books/about/Emile.html?id=nWj_ZbVcwgIC

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533-536. https://www.nature.com/articles/323533a0

Saidabad, M. Y., Hassanzadeh, H., Ebrahimi, S. H. S., Khezri, E., Rahimi, M. R., & Trik, M. (2024). An efficient approach for multi-label classification based on Advanced Kernel-Based Learning System. *Intelligent Systems with Applications*, *21*, 200-332. https://www.sciencedirect.com/science/article/pii/S2667305324000085

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*. https://www.wcu.edu/webfiles/pdfs/shulman.pdf

Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, *57*(1), 1-22. https://people.ucsc.edu/~ktellez/shulman.pdf

Smith, J. (2020). *The history of pedagogy*. Educational Review Press.

Srivastava, M. (2023). Addressing spurious correlations in machine learning models: a compreh ensive review. *OSF Prepr*. https://www.researchgate.net/publication/370210419_Addressing_Spurious_Correlations_in_Machine_Learning_Models_A_Comprehensive_Review

Su, N., Wang, Y., Ye, C., & Liu, Y. (2024). Artificial Intelligence in General Context (AIGC) Technology's Impact on Higher Education Pedagogical Content Knowledge (PCK) Teaching. *Advances in Education, Humanities and Social Science Research*, *12*(1), 831-841. https://www.researchgate.net/publication/387366905_Artificial_Intelligence_in_General_Context_AIGC_Technology%27s_Impact_on_Higher_Education_Pedagogical_Content_Knowledge_PCK_Teaching

Tang, J., Zhou, X., Wan, X., Daley, M., & Bai, Z. (2023). ML4STEM professional development program: Enriching K-12 STEM teaching with machine learning. *International journal of artificial intelligence in education*, *33*(1), 185-224. https://link.springer.com/article/10.1007/s40593-022-00292-4

Tshidi, M. (2022). Supervised learning algorithm's role in flagging programming concepts that call for Information Technology teachers' attention. https://www.researchgate.net/publication/377627448_Supervised_learning_algorithm%27s_role_in_flagging_programming_concepts_that_call_for_Information_Technology_teachers%27_attention

Van Driel, J. H., & Berry, A. (2010). Pedagogical content knowledge. In *International encyclopedia of education* (pp. 656-661). Elsevier. https://www.researchgate.net/publication/271429393_Pedagogical_Content_Knowledge

Yang, F., & Li, F. W. B. (2018). Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. *Computers & Education*, *123*, 97-108. https://www.sciencedirect.com/science/article/abs/pii/S0360131518300861

Yao, R., Tian, M., Lei, C.-U., & Chiu, D. K. W. (2024). Assigning multiple labels of sustainable development goals to open educational resources for sustainability education. *Education and Information Technologies*, *29*(14), 18477-18499. https://link.springer.com/article/10.1007/s10639-024-12566-6

Zeller, J., & Riese, J. (2025). Competency Profiles of PCK Using Unsupervised Learning: What Implications for the Structures of pPCK Emerge From Non-Hierarchical Analyses? *Journal of Research in Science Teaching*. https://onlinelibrary.wiley.com/doi/full/10.1002/tea.70001

Zeroual, N., Lamia, M., & Hafidi, M. (2024). A pedagogical orientation system based on blockchain technology and machine learning. *Education and Information Technologies*, *29*(3), 2905-2930. https://link.springer.com/article/10.1007/s10639-023-11941-z

Zhu, X. (2013). Machine teaching for bayesian learners in the exponential family. *Advances in Neural Information Processing Systems*, *26*. https://arxiv.org/abs/1306.4947

Zhu, X. (2015). Machine teaching: An inverse problem to machine learning and an approach toward optimal education. https://ojs.aaai.org/index.php/AAAI/article/view/9761

Zhu, X. (2018). Machine Teaching: An Inverse Problem in Machine Learning and a Path Toward Optimal Teaching. https://www.researchgate.net/publication/363003952_Machine_Teaching_An_Inverse_Problem_to_Machine_Learning_and_an_Approach_Toward_Optimal_Education

Zohuri, B., & Mossavar-Rahmani, F. (2024). Revolutionizing education: The dynamic synergy of personalized learning and artificial intelligence. *International Journal of Advanced Engineering and Management Research*, *9*(1), 143-153. https://www.researchgate.net/publication/380216282_Revolutionizing_Education_The_Dynamic_Synergy_of_Personalized_Learning_and_Artificial_Intelligence

AITBSS AI and Tech in Behavioral and Social Sciences
E-ISSN: 3041-9433