Applying machine learning to augment the design and assessment of immersive learning experience

Chih-Pu Dai

Department of Learning Design and Technology

University of Hawaiʻi at Mānoa, Honolulu, HI 96822, USA

cdai@hawaii.edu

Abstract[1] The use of machine learning has seen a remarkable rise in education research with extraordinary potential to enhance immersive learning experience. An immersive learning experience, in which learners participate in simulated virtual environments, can promote deep learning as learners actively explore and construct knowledge within the learning environments. Despite the growing interest and increasing applications, the ways in which machine learning can be used to augment the design and assessment of immersive learning experience remain an open area of exploration. Machine learning can be used to provide adaptive and personalized learning, increase interactivity and engagement, and track learning activities in immersive learning environments. In this chapter, the author describes the current state of research on machine learning in immersive learning environments, including adaptive and personalized learning, natural language processing and conversational artificial intelligence, and data and learning analytics. The author also outlines the potential future directions for the applications of machine learning in

[1] In M. S. Khine (ed.), *Machine Learning in Educational Sciences: Approaches, Applications and Advances*, https://doi.org/10.1007/978-981-99-9379-6_12

designing and assessing immersive learning experiences to inform educational sciences. This chapter serves as a useful reference for educational researchers, practitioners, and policy makers seeking to make informed decisions on the design and assessment of immersive learning experiences.

Keywords Machine learning; Natural language processing; Immersive learning environments; Artificial intelligence; Personalized learning

1. Introduction

Immersive learning experience afforded by learning technologies, for example, virtual reality, augmented reality, mixed reality, and simulation games has gained attention in education research due to its affordances to offer a learning experience that is not limited or constrained by time and space. Education researchers and practitioners have also been putting efforts into exploring how to provide more personally meaningful and adaptive immersive learning experiences (Dai & Ke, 2022; Liu et al., 2020). However, providing such learning experience for individual learners is not an easy task because each learner is unique in their diverse and dynamic learning backgrounds, learning states, and learning needs. Further, as learners actively interact with the immersive learning environments, they vary in terms of their learning trajectories and interactions leading to diverse performance outcomes; hence it is difficult to adapt to individual needs with machines or computer systems. For example, despite decades of extensive efforts, a previous meta-analysis found no significant difference between adaptive mechanism and non-adaptive mechanism in the context of educational games (Liu et al., 2020). There is a critical need to improve these technologies in sound ways so that adaptive and personalized learning can be realized. Machine learning—–a component of artificial intelligence, has experienced significant growth in recent years and emerged as a potential solution for designing and assessing personally meaningful and adaptive immersive learning experiences.

Jordan and Mitchell (2015) proposed that machine learning deals with two major questions: "How can one construct computer systems that automatically improve through experience? and What are the fundamental statistical-computational-information-theoretic laws that govern all learning systems, including computers, humans, and organizations?" (p. 255). Based on the suggestion (Joran & Mitchell, 2015), when designing and assessing immersive learning experience, machine learning is used as a holistic approach, involving the computers, learners, and learning settings, to learn automatically when learners have interacted in the systems and adapt to learners' interactions driven by learners' input data and algorithms. As such, designing and assessing immersive learning experience is an interdisciplinary effort that requires expertise and input from disciplines such as education, psychology, cognitive science, computer science, and data science. Machine learning should also be used purposefully to address problems. In particular, "a learning problem can be defined as the problem of improving some measure of performance when executing some tasks, through some type of training experience" (Jordan & Michell, 2015, p. 255). In designing and assessing immersive learning experience, specifically defined learning outcomes is of critical importance while the process-oriented interactions were collected and recorded as evidence and baseline conditions. The process-oriented interactions can be demonstrated in the forms of natural language and embodied movements. For example, learners can use text input and verbal chat to interact with the immersive learning environments. They can also perform actions to improve knowledge or skills.

In the following sections of this chapter, I introduce the current state of research on the integration of machine learning for the purpose of designing and assessing immersive learning environments. In particular, I explore how natural language processing, conversational artificial intelligence, multimodal learning analytics, and data and learning analytics have been used to enhance learning in immersive learning environments. The chapter will conclude with potential future directions for the applications of machine learning for designing and assessing immersive learning experiences, aiming to provide applicable insights for the field of educational sciences.

2. Machine learning approaches in educational sciences research

The applications of machine learning can be identified in many different industries, such as health, business, military, and education. According to Thormundsson (2022), the market of artificial intelligence is projected to grow to 126 billion U.S. dollars by 2025, machine learning is part of this tremendous market. In education, machine learning can be useful as it can be applied to provide personalized and adaptive learning as well as analyze large-scale datasets to inform educational practices (Chen et al., 2018).

2.1. What is machine learning?

Machine learning has grown in the past few years. There are several techniques or solutions in machine learning that can be used to tackle various problem domains, such as clustering problems, predictive problems, detecting problems, optimization problems, or association rules. With these solutions, machine learning can be applied in education to provide adaptive learning, personalized learning, and assessment. There are different applications of machine learning in education research, sometimes coupled with natural language processing.

One characteristic of immersive learning is that the interactions are multimodal. Therefore, machine learning has been applied in multiple ways to enhance the learning experiences.

Machine learning can be classified broadly as supervised learning, unsupervised learning, and reinforcement learning (Jordan & Mitchell, 2015). Jordan and Mitchell (2015) also pointed out that there are variant approaches in between this broad classification. For example, semi-supervised learning has been used to predict learning performance (Livieris et al., 2019). Semi-supervised learning is useful when a portion of data is labeled while a large portion of data is unlabeled (Livieris et al., 2019).

Supervised learning is when a machine learns to accurately identify labeled data with machine recognizable vector scores (LeCun et al., 2015). For example, if given a house, the machine will identify it as a house. To achieve this, a large amount of data will be collected and labeled with its category (e.g., house as a house), this dataset is a training dataset. During the training stage, machines will learn the patterns of the dataset. During the testing stage, when provided with an unseen dataset, the machine will use the training experience for problem-solving. That is, if given a house the machine is hopefully to identify it as a house with a high accuracy rate; if given a mouse, it will hopefully identify it as not a house. In addition to these pattern identification problems, supervised learning can also be used to address predictions or decision-making problems.

Unsupervised learning, on the other hand, refers to the type of machine learning that does not use labeled data, instead, raw data are provided, and the machine finds patterns or relations of the data on its own. Generally speaking, unsupervised learning is used for exploratory purposes. For example, Amershi and Conati (2009) collected log data from a learning environment; without labeling, they used unsupervised learning to cluster students' behaviors in the learning environments for the purpose of user model building to enhance the learning experience.

Reinforcement learning, which shares similar concepts to behaviorism in human learning, is when machines learn from the processes of rewards and punishments by an agent based on the actions taken. In other words, machines are engaged in pattern-finding processes with a policy and determined values. When the machines complete desired actions, they will get a reward, if the actions are undesired, they will be punished. For example, Bassen et al. (2020) used reinforcement learning to automatically give relevant assignments to the students in an online learning system. Based on the traced students' performance and their interactions with the course materials in the system, reinforcement learning agents update students' learning states and assign personalized tasks for the students.

2.2. Application of Machine learning in Educational Sciences

While machine learning for immersive learning experience has begun to grow, machine learning has been utilized in various educational research contexts. Examining its applications in broad educational sciences contexts offers valuable insights and implications for the use of machine learning in immersive learning environments, particularly as the specific designs, development, and applications of machine learning in immersive learning are still in the early stages of innovation. By leveraging the experiences and successes of other educational contexts, educators can more effectively integrate machine learning into their immersive learning environments and drive advancements in the field as a whole.

To understand the types, contexts, disciplines, and targeted populations that machine learning has been used in educational sciences, Luan and Tsai (2021) conducted a review. They first mapped out *different categories* that machine learning has been applied in education.

Prediction donimates the included studies, with approximately 63% of the included studies aimed to provide prediction, followed by diagnosis or profiling (23%). 45% of the included studeis is in online settings, 25% in STEM and another 25% belongs to multiple domains. Another dominance of machine learning applications is that it has been mostly applied in higher education settings for university students (60%). The findings of the study underscored the limited presence of machine learning applications tailored specifically for K-12 students. Moreover, it emphasized the importance of careful deliberation when integrating machine learning into K-12 education, particularly in terms of addressing ethical concerns and ensuring developmental appropriateness. The study revealed that deploying machine learning in K-12 settings requires extensive efforts and considerations to navigate these crucial aspects effectively.

To further explore the applications, Shah et al. (2021) summarized the different ways machine learning can be applied in educational sciences. In general, prediction, clustering, and semi-supervised learning are the three most frequently used applications. Specifically, Shah et al. (2021) proposed “Academic Performance Predictions” (p. 6), “Data Mining to Find Out Hidden Patterns” (p. 7), “Improving Student Results on Basis of Past Experience” (p. 8), and “Semi-supervised Learning in Education” (p. 8). Generally speaking, predication leverages historic data or past students’ performance to generate future insights. Specifically, predicting *“at risk” students* in the learning environments is one of the prominent areas that can genuinely help learners and improve learning. Mining the data to understand the patterns of students’ learning is also a valuable way machine learning can be used to provide data-driven insights. For example, topic modeling is an unsupervised machine learning approach with a natural language processing technique to mine text-based data for educational use. In health education and promotion, topic modeling has been used with social media data (Valdez et al., 2021). Gencoglu et al. (2023) used topic modeling to analyze students' open-ended responses to teachers' teaching behaviors. By comparing machine learning-based topic models and human ratings, the authors finalized eight topics (i.e., *Topic 1*, Clear explanation, *Topic 2*, Student-centered supportive learning climate, *Topic 3*, Lesson Variety, *Topic 4*, Likable characteristics of the teacher, *Topic 5*, Evoking interest, *Topic 6*, Monitoring understanding, *Topic 7*, Inclusiveness and equity, and *Topic 8*, Lesson objectives and formative assessment).

3. The design and assessment of immersive learning experience

3.1. Immersive learning experience – How learning occurs?

Learning is a complex process, and it can be understood and studied from different learning perspectives, paradigms, and theories, such as behaviorism, cognitivism, and constructivism. Moreover, given its complex nature, learning is also being studied with its connections with cognition, development, technology, motivation, and neuroscience (Driscoll & Burner, 2021). As learning spaces and technologies evolved, connectivism (Goldie, 2016; Siemens, 2005) and learning principles related to *metaverse* (Hwang & Chien, 2022) have also emerged. Connectivism emphasized that learning is not an individualistic activity, instead, it connects nodes and knowledge sources (Goldie, 2016). In essence, knowledge development and management occur within *people* and *contexts* (Siemens, 2005). Learning and knowledge activities in massive open online courses (MOOCs) are usually explained with connectivism (Goldie, 2016). Similar to *connectivism*, in the context of artificial intelligence and metaverse, learning activities are "shared," "persistent," and "de-centralized" (Hwang & Chien, 2022, p. 1). In the metaverse, learners interact with multiusers (shared). Further, learning, working, and living are sustained activities (persistent), and unauthorized modifications to personal property and logs are not possible (de-centralized) (Hwang & Chien, 2022). These characteristics of metaverse call for new and updated learning theories and paradigms to explain the learning phenomena (Hwang & Chien, 2022).

While contemporary learning theories and paradigms are still in need of innovations, to explain the learning phenomena in immersive learning environments, *situativity theory* is often used as a perspective (Durning & Artino, 2011). Situativity theory suggests that “knowledge and thinking (cognition; i.e., situated cognition), as well as learning (i.e., situated learning), are situated in experience” (Durning & Artino, 2011, p. 188). Situativity theory is an extension of the theories of Vygotsky, Bandura , and Dewey; it contains multiple theoretical perspectives, including situated cognition, situated learning, distributed cognition, and embodied cognition (Durning & Artino, 2011). These theoretical perspectives emphasize different context interactions and learners’ experiences that are well-suited for explaining learning in immersive learning environments.

Another theoretical underpinning for the immersive learning experience is experiential learning (Kolb, 1984). According to Kolb, learning occurs through a four-stage cycle of concrete experience, reflective observation, abstract conceptualization, and active experimentation. This cycle emphasizes the importance of both concrete experiences and reflection in the learning process (Kolb, 1984). It involves engaging in concrete experiences, reflecting on those experiences, and applying the insights gained to new situations (Kolb, 1984). Experiential learning is often used to facilitate deep learning and promote learning transfer (including knowledge and skills). Experiential learning can take many forms, including internships, apprenticeships, simulations, and field trips. The goal of experiential learning is to create opportunities for learners to engage in authentic, real-world experiences that are relevant to their learning goals. By reflecting on these experiences and applying the insights gained to new situations, learners can deepen their understanding and develop practical skills that are transferable to different contexts.

Building upon situativity theory and experiential learning, immersive learning experience can be guided by different learning principles and theories depending on the desired achievements and outcomes. To elaborate, learning principles and theories should be intrinsically integrated and strongly aligned when considering the design and assessment mechanism in an immersive learning environment in conjunction with machine learning. In the subsequent sections, I will delve into the intricacies of the design and assessment of immersive learning environments with essential learning theories and principles to effectively serve educational purposes.

3.2. Augmenting the design of immersive learning experience with machine learning

"Immersive learning allows learners to freely explore, experience, interact with objects and characters, and try out new ideas and solutions in the virtual environment through experiential learning approach." (Ip et al., 2018, p. 505). To maximize the affordances of immersive learning environments, the design of an immersive learning experience can be augmented by machine learning as machine learning makes dynamic and adaptive interactions and personalization possible. To this end, learning designers can potentially assist learners in achieving and optimizing learning outcomes and experiences in technology-enhanced learning environments. These types of learning experiences have been theorized and promoted for decades. With the advancement of technology, the affordances of immersive learning environments elevated by machine learning can be naturally linked to sociocultural theory

(Vygotsky, 1978). Grounded in Vygotsky's (1978) sociocultural theory, the notion of the zone of proximal development provided the foundations for adaptive and personalized learning.

Different machine learning techniques have been applied to drive and augment the design of adaptive and personalized learning. Adaptive learning is thus arguably a key focus of machine-learning-driven immersive learning experience. Vaughan et al. (2016) discussed the principles of adaptive systems, defining an adaptive system as "a set of interacting entities that together are able to respond to changes" (p. 3). Liu et al. (2020) suggested that adaptive learning in simulation games considers *what to adapt* and *how to adapt*. *What to adapt* involves adaptation to learners and the instructional approaches; whereas *how to adapt* entails a competency-based approach using threshold and decision algorithms as well as a preference-based approach using classification machine learning techniques to tailor to students' interests. Similarly, Zahabi and Abdul Razak's (2020) review of adaptive virtual reality-based training suggested various variables that can be adaptive to the learners. Adaptivity can be implemented at different timing–before the training experience and during the training experience (Zahabi & Abdul Razak, 2020). In addition to adapting the content at different timing, providing adaptive feedback is another important variable in machine-learning-driven immersive learning experience design (Zahabi & Abdul Razak, 2020). With a similar emphasis as Zahabi and Abdul Razak (2020) on adaptive feedback, Vaughan et al. (2016) also pointed out that a key mechanism in an adaptive system is feedback loops. Citing Stenudd (2010), the machine learning feedback loop in the adaptive system utilized four categories——prediction, recognition, detection, and optimization. In these loops, the inputs are determined by the outputs from previous loops. Through constant update and response mechanisms driven by machine algorithms in the contexts, the functional adaptation to learners' behaviors and parameters adjustments is realized.

Recently, innovative machine learning approaches have been possible in addition to the traditional *a priori* competency-based approach, in which a competency model has been determined so that learners' data and performance will be examined accordingly. The recent machine learning models can be built with learners' in-situ generated data for more real-time and authentic designs (see Figure 1 for an example). To enhance students' learning experiences with machine learning, using a sound procedure is vital. Machine learning approaches follow a systematic procedure to be realized and implemented for the purpose of enhancing teaching and learning. As mentioned previously, machine learning is based on using large amounts of data to train or allow machines to learn, therefore, in real-world application of machine learning, the procedures have a focus on data wrangling and processing to solve different machine learning problems. For example, Wu et al. (2020) used a machine learning approach to classify text for social-media-based online discussion. In the context of an immersive learning environment, Rogers et al. (2021) proposed novel techniques integrating machine learning and virtual reality for surgical training. Both propositions from Wu et al. (2020, p. 68) and Rogers et al. (2021, p. 1251) suggested useful procedures that apply innovative machine learning approaches for human learning enhancement.

To be more specific, the first step is to gather data as learners engage in immersive learning and training experiences. These data can be automatically logged into the computer and can be very messy and complex. Next, data are to be processed and cleaned for feature extraction. Feature extraction can reduce the number of information used for machine learning and improve the relevance of the machine learning outcomes (Guyon & Elisseeff, 2006). Afterward, machine learning algorithms can be selected based on the purposes, problems, expected outcomes, and accuracy rate of the algorithms. Rogers et al. (2021) provided a few examples——K-nearest neighbor Gradient Boosting Logistic Regression Support Vector Machines Discriminant Analysis. In this process, the machine learning models and algorithms are trained and tested using the provided data, and the leanring outcomes are analyzed. Appropriate machine learning algorithms can be applied to learning contexts such as providing personalized, adaptive, and real-time feedback, classifying learners' interaction for personalized learning, and assessing learning (Dai, Ke, Pan, et al., 2023; Rogers et al., 2021).

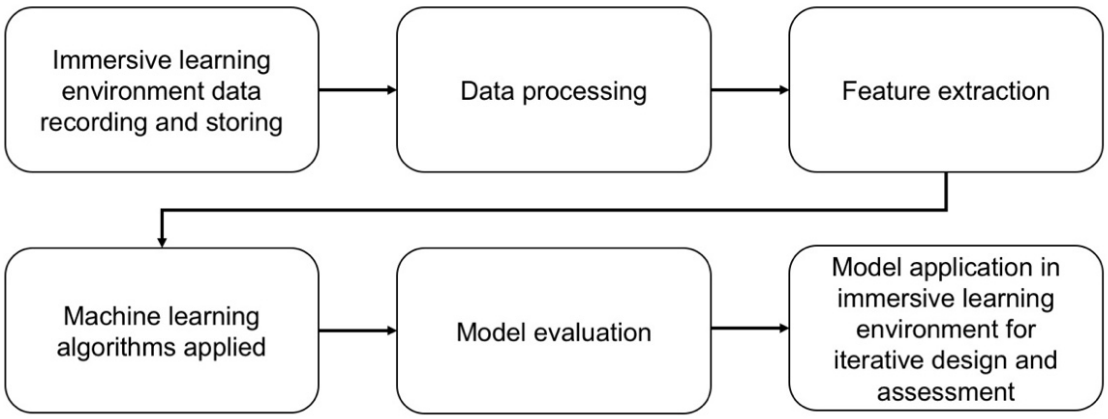

Figure 1. Machine learning approach for design and assessment in immersive learning environment.

There has been a growing interest in these applications among researchers in educational sciences. With a similar approach, Abouelenein and Nagy Elmaadaway (2023) used ANN to design neuro-computerized virtual learning environments for math preservice teachers to develop computational thinking. They followed a three-stage design process. The first stage involved exploratory factor analysis with input data and framework development. The second stage focused on process-oriented learning model building with preservice teachers to reinforce the development. The third stage aimed to generate adaptive neuro networks based on an input-output function that preservice teachers can use to fit into their learning progress.

Other innovations focused on using machine learning to guide learners or adapt to learners' profiles. For instance, Lee et al. (2014) created a director agent to support the learning experience in a narrative-centered immersive educational game. Supervised learning was used to create a director agent that can make adaptive decisions responding to learners' in-game actions. Dai, Ke, Pan, et al. (2023) used the Gaussian Mixture Model (GMM), an unsupervised machine learning model for soft clustering, to classify learners' interaction with learning supports in an immersive simulation game (called *E-Rebuild*, Ke et al., 2019). The machine learning technique coupled with natural language processing can be integrated into learning systems to inform the designs of immersive learning experience for data-driven design decisions. In assessing the accuracy of classification of different supervised machine learning algorithms (i.e., Naive Bayes, *k*-nearest neighbors, and support vector machines), Asbee et al. (2023) maintained that adaptive assessment can be best implemented with Naive Bayes in the context of virtual reality. In this context, support vector machines can also perform well in classifying learners, but not when there is missing data.

In Asbee et al.'s (2023) setting, tactile feedback was provided when learners were engaged in virtual reality. This highlighted another growing area of study, that is, machine learning-facilitated multimodal analytics. This area of study is unique in immersive learning environments and the findings can offer innovative insights into the design of immersive learning experience. For example, as the study of hand gestures has been a focus in this area, Bahceci et al. (2022) used supervised machine learning to understand and classify learners' hand gestures to improve the immersive learning experience. In a comprehensive review and case study, Philippe et al. (2020) highlighted the applications of multimodal interactions in virtual reality. They presented examples of practice training in the pharmaceutical industry and surgical field with hands-on activities in virtual reality. Despite emerging studies, the area is in its nascent stage and has been growing.

3.3. Natural language processing and conversational artificial intelligence

Immersive learning experience can be enhanced with the integration of virtual agents using natural language processing (NLP). Virtual agents can be represented as chatbots, virtual humans, conversational agents, pedagogical agents, or virtual beings in the literature (e.g., Dai & Ke, 2022; Ke et al., 2020; Sinatra et al., 2021; So et al., 2023). Virtual agents can provide interactive experience through dialogic learning (Graesser & McNamara, 2010) as well as coaching and guiding the learners (Psotka, 1995). NLP in virtual agents makes natural interactions possible (Psotka, 1995). There are several different machine learning approaches behind the virtual agents that drive natural language interactions with the learners——Bayesian network (Johnson, 2003), logistic regression and long short-term memory (LSTM) (Nye et al., 2021), latent semantic analysis, and large language models (e.g., Dai & Ke, 2022; Bhowmik et al., 2022). As one of the examples, the Bayesian network allows researchers and learning designers in educational sciences to explore ways to implement natural language interactions with learners in computer-based systems. Bayesian network is a predictive machine learning approach modeling a set of variables that represent learners' characteristics such as motivation, knowledge, and other attributes. As the model learns from learners' data and actions, the model updates its predictive mechanisms for the purpose of recommending adaptive and personalized content that fit into learners' zone of proximal development (Vygotsky, 1978). In a more complex systems design for the purpose of training for early career teachers' classroom management skills, Delamarre et al. (2021) adopted MASCARET framework (Querrec et al., 2004), a type of multi-agent system that utilized Unified Modeling Language (UML) with semantic approach and different modeling architectures, to build a 3D immersive simulation-based learning environment.

Designing tutoring experience in intelligent tutoring systems that immerse learners in conversational knowledge exchange, Graesser et al. (2014) adopted expectation-and misconception-tailored (EMT) dialog structure to foster students' learning and scientific reasoning. Using semantic-pattern-completion algorithms, Graesser et al. (2014) crafted conversational agents that engage in conversational patterns with expected answers, misconceptions, or alternative answers to constantly negotiate meaning and facilitate reasoning. *Natural language understanding* is another way to realize human-machine conversations. Reviewing the literature on chatbots with conversational artificial intelligence, Wollny et al. (2021) indicated that chatbots used large-scale data crawled from the Internet to enable text- and/or voice-based interactions with learners. Their review found that chatbots in education have been predominantly used for skill improvement (32%), followed by efficiency of education (25%) and motivation (13%). They also revealed that chatbots in education assume the pedagogical role of learning the most (49%), followed by assisting (20%) and mentoring (15%). In simulation-based learning settings, Dai and Ke (2022)'s review revealed that conversational agents can assume different roles with different AI technologies, for example, providing guidance, acting as teachers, peers, or teachable agents.

There are four distinct types of conversational artificial intelligence to date, ranked by their level of interactivity. These include scripted AI, rule-based AI, module-based AI, and natural language processing/machine learning AI (Dai, Ke, Pan et al., 2023). Scripted AI is the least interactive, while natural language processing/machine learning AI is the most interactive. Despite the perception that natural language processing or conversational agents are a recent development, more than fifty years ago, ELIZA was designed to converse in natural language in a manner resembling psychotherapists (Weizenbaum, 1966). In some previous systems, human learners/users have reported perplexity, inappropriateness, irrelevance, and bias as a result of interactions with virtual agents (So et al., 2023). In recent times, the development of large language models with big data and billions of parameters has been prolific in creating improved natural language understanding. Large language models have been integrated into different contexts for applications. For example, So et al. (2023) used a large language model to design virtual partners in an online conferencing tool in public health settings. In educational sciences, large language models have become a transformative tool in conversational artificial intelligence for learning. Large language models' capabilities to carry out and maintain conversations with learners are ideal for educational purposes (Dai et al., 2024; Dai & Ke, 2022).

For the training of teaching in virtual reality-supported simulation-based learning (Dai et al., 2023; Ke et al., 2021), a large language model (e.g., Generative Predictive Transformers 2, GPT-2) has been used to design and develop virtual humans in *OpenSim* (Bhowmik et al., 2022; Dai et al., 2021; Ke et al., 2021). Ke et al. (2021) created *Evelyn*, virtual student agent with conversational artificial intelligence, to assist preservice teachers in practicing ambitious science teaching enactment. Generative artificial intelligence affords real-time authentic dialogs between human learners and virtual agents whereas virtual reality environments (i.e., OpenSim) can afford immersive and authentic sense of presence with suitable scenario designs for in-situ practices (Dai et al., 2023). According to So et al. (2023), Ke et al. (2021), and Dai (2023), one way to apply large language models to a local context is by training them on local datasets and integrating them into targeted computer systems. Figure 2 depicts a generic architecture for using large language models in education within a local context. This approach can enable more efficient and accurate natural language processing which could lead to improved educational outcomes.

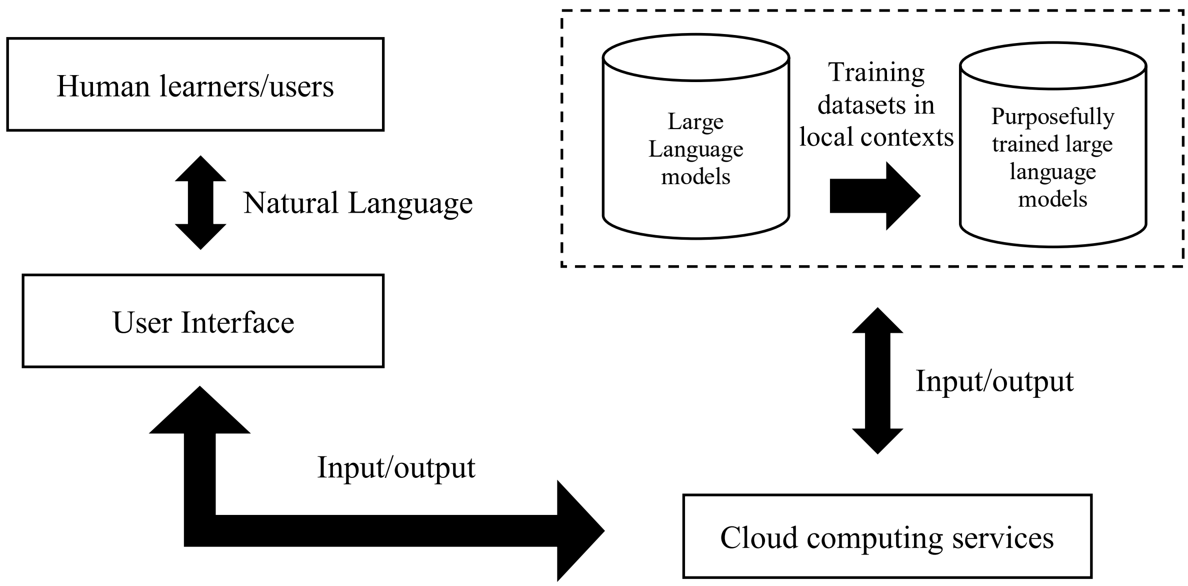

Figure 2. *Conversational artificial intelligence architecture with large language models application*

The designs of the teacher learning system with large language models demonstrated one alternative way to create dialog systems grounded in theories of situativity (Durning & Artino, 2011) and experiential learning (Kolb, 1984) that help learners with artificial intelligence. By localizing and contextualizing the large language models in educational sciences (Dai & Ke, 2022), the system is equipped with capabilities to carry on domain-generic conversations as well as engage in domain-specific interactions. A growing number of studies have suggested that the conversational agents alone provide practice and experiential learning opportunities for inquiry-based learning while scaffolding or learning support integrated into such learning settings is the most beneficial; in other words, combining conversational agents with additional leanring support enhances the effectiveness of the learning experience (de Jong et al., 2023).

The virtual student agent system with artificial intelligence *and* integrated support has been found to be effective (Dai, 2023). In an experimental design study, Dai (2023) examined the effects of learning support on preservice teachers' teaching knowledge. The study also investigated its impacts on preservice teachers' self-efficacy when interacting with virtual artificial intelligence student agents. The results revealed that in comparison to the no-treatment control group, preservice teachers in the learning support group with agents and the agent-only group performed significantly better on teaching knowledge and skills. Although no significant results were found between groups on teaching self-efficacy, the learning support group's teaching self-efficacy improved significantly pre- to post-intervention.

3.4. Assessment for immersive learning experience

Assessment is a critical component of any form of learning. Scholars have been discussing the integral roles of assessment in promoting and encouraging learning (Shepard, 2000; Shute et al., 2021). Assessment serves multiple roles for learning (Shepard, 2000; Shute et al., 2021). It can be used to determine the learning outcomes (e.g., knowledge, skills, and other traits and capabilities), provide feedback for learners to continuously improve, and maintain and sustain motivation to learn. Essentially, the relations between feedback and assessment have been studied extensively (Hattie & Timperley, 2007; Shute, 2008). Hattie and Timperley (2007) maintained that feedback received by teachers or learners is the result of formative assessment. The author suggested that assessment is devised to understand the gap between the current learning state and the determined learning goals, at three levels: “about tasks, about the processes or strategies to understand the tasks, and about the regulation, engagement, and confidence to become more committed to learn” (p. 101). Shute (2008) further elucidated that feedback and assessment are mutually reinforcing for diagnosis and contribute to learning and performance. In line with a focus on learning and personalization, Shepard (2021) emphasized the need to pursue equitable assessment practices that are coupled with ambitious teaching practices. The integration of ambitious teaching practices and equitable assessment urges teachers and designers to gain a thorough and deep understanding of each student, including their learning, emotional, social, and cultural backgrounds as well as the development of disciplinary practice. That is, situated in sociocultural theory, considerations for assessment in immersive learning experience include equitable approach, disciplinary practices, and shared goals.

Assessing immersive learning outcomes with machine learning and artificial intelligence can not only inform learning designers what learners do in the immersive learning environments, but learners themselves can also benefit (Shepard et al., 2018) from unobtrusive assessments designed and developed in the immersive learning environments. There are many different ways to design and blend sound assessment with learning. In computer-based learning environments, stealth assessment is one way to achieve such goals (Shute, 2011). “Stealth assessment refers to ECD-based assessments that are woven directly and invisibly into the fabric of the learning environment” (Shute & Ke, 2012, pp. 52-53). Based on the notion of stealth assessment, Ke and Shute (2015) presented assessment design heuristics in immersive games suggesting that machine learning approaches, including, for example, data mining and learning analytics are promising and the integration of assessment and learning task design should be carefully planned in the early stages of the immersive learning environment design and development with iterative testing and refinement. An important driving mechanism behind stealth assessment is evidence-centered design (ECD) (Mislevy et al., 2003). For the stealth assessment to be pedagogically sound and personally meaningful in immersive learning environments, competency and protocols are important, this is called the competency model in stealth assessment (Mislevy et al., 2003; Shute, 2011). Competency models include the knowledge, skills, and other capabilities and aspects of learning to be assessed. Similar to the assessment in other contexts, stealth assessment in immersive learning environments based on competency models ensures that the shared goals of the learning experience can be achieved (Shepard et al., 2018).

Aligning with the framework of ECD (Mislevy et al., 2003), other applications of machine learning in assessment for students' learning have used threshold values, baseline, and indices to perform assessment tasks in immersive learning environments. For instance, in the context of medical education, Winkler-Schwartz et al. (2019) used novel processes with machine learning algorithms to generate and extract practice-based evidence and metrics (e.g., with four categories in the operation of surgery: movement, force, bleeding, or tissue) for assessing surgical task in a high-fidelity simulation. The results of the machine learning classification suggested that among four machine learning algorithms (i.e., k-nearest neighbor, Naive Bayes, discriminant analysis, support vector machine), k-nearest neighbor can use the least number of performance metrics (i.e., 6) to classify performances from different expertise groups. Nevertheless, Boulet and Durning (2019) have advocated for more psychometric studies to ensure the validity of assessments done by machine learning algorithms.

Understanding the mechanisms of machine learning driving the assessment is crucial. The ability to provide equal, just, and inclusive assessments in immersive learning environments largely depends on the machine learning models and the data used to devise these assessments. The efforts to provide deep and meaningful learning for *all*, gave rise to the considerations that the models and data used should include underrepresented groups of learners to capture their interests and promote in-depth and rigorous learning via formative feedback and assessment (Dai & Ke, 2022; Ke & Shute, 2011).

More recently, using unsupervised machine learning for data-driven assessment is growing in the context of digital simulation-based learning environments. Using topic modeling, Littenberg-Tobias et al. (2021) found that the structural topic model algorithms (Roberts et al., 2019) can properly recognize the natural language text responses by participants that indicates equity practices and mindsets across four simulation modules (i.e., Jeremy's Journal, Coach Wright, Roster Justice, and Layers, see Littenberg-Tobias et al., 2021, p. 5 for details).

Aligning with the recent development of machine learning with data-driven assessment, one prominent area of machine learning for immersive learning experience assessment is the application of multimodal learning analytics that afford innovative and unobtrusive assessment and formative feedback (Blikstein, 2013; Ouhaichi et al., 2023). Ouhaichi et al. (2023) provided a comprehensive review on the ways multimodal learning analytics can be applied in immersive learning technologies such as virtual reality and mixed reality. Their findings revealed that virtual reality is one key aspect of multimodal learning analytics (Ouhaichi et al., 2023). Lorenzo et al. (2013) used multiple modalities and sources of data to create a comprehensive assessment framework including body movement, voice, eye movement, attention, and empathy. Lee-Cultura et al. (2022) used multimodal data capturing young children's interaction with a simulation game (i.e., Marvy learns: Motion-Based Learning Technologies) to inform learning design. They used principal component analysis (PCA) and exploratory factor analysis as fundamental machine learning analyses to identify the relationships between human coded analysis and multimodal learning analytics on young children's problem-solving. They emphasized that the results can be used for providing feedback on learners' cognition and affection. Similarly, using PCA, Kroeze et al. (2021) developed an automated assessment tool to identify students' weaknesses in concept maps during inquiry-based learning in virtual immersive laboratories (i.e., with functions that allow students to manipulate variables in science learning). The assessment tool was applied to provide students with feedback on their weaknesses in science conceptualization.

These studies have focused on multiple dimensions of learning and these aspects involve the study of learners' movement, affection, learning products, and learning states. Indeed, the application of multimodal learning analytics to provide formative feedback and in-situ assessment for immersive learning experience requires interdisciplinary expertise such as learning sciences, affective computing, and human-computer interaction (Cukurova et al., 2020).

The importance of evaluating the effectiveness of teaching practices when preservice teachers are engaged in immersive teacher learning cannot be overemphasized. While traditional methods of evaluation, such as observation and self-reporting, have their limitations, recent advancements in natural language processing and machine learning offer a promising alternative. In a recent project (Bhowmik et al., 2022; Dai, 2023; Ke et al., 2021), researchers developed an algorithm that uses natural language processing and deep learning to assess the teaching practices of preservice teachers. By analyzing the language used by these teachers, the algorithm is able to detect and classify the type of teaching practice being employed. To develop the algorithm, the researchers collected natural language data from preservice teachers and used human judgment to identify the specific teaching practices being used, such as lecturing, Socratic questioning, or ambitious teaching practices (e.g., the orchestration of students' ideas or resources). The data was then used to train the algorithm, resulting in an impressive level of accuracy. This innovative approach has several merits and the potential to improve the way teaching practices are supported in immersive teacher learning. By providing preservice teachers with targeted feedback and support for their teaching practices, can better prepare them for their careers as educators and contribute to improved educational outcomes for their students. Sharing similar notions of helping teachers with their professional practices, in simulation games, Westera et al. (2018) pursued an automatic essay scoring mechanism to reduce teachers' workload in the context of online training. They proposed that "a multilayer perceptron network with two hidden neurons within the hidden layer" (p. 220) was the most appropriate for the assessment purpose after performing the cross-validation with human scoring and excluding biases.

4. Discussion and future directions

In this chapter, I have delineated the design and assessment of immersive learning environments augmented by machine learning. Immersive learning environments offer learners valuable opportunities for *in-situ* practices and knowledge construction. When learners are engaged in these learning environments, they generate a vast amount of data that can be leveraged for machine learning. Drawing on constructivist theories, immersive learning environments supported by machine learning are well-suited for adaptive and personalized learning. However, there is a need for more applicable theories to explain the increasingly complex learning with these learning environments and technologies.

I highlighted several key aspects of immersive learning experiences with machine learning. First, for the design of immersive learning experiences, different machine learning techniques can be integrated. For example, topic modeling (e.g., Gencoglu et al., 2023), GMM (e.g., Dai et al., 2023), and ANN (e.g., Abouelenein & Nagy Elmaadaway, 2023). Crucially, a multi-stage approach is essential for producing pedagogically sound designs and validated results. The pipeline of modern machine learning applications is consistent and rigorous in the literature within various contexts (see Figure 1). Authentic data collection and storage, pre-processing, training, and evaluation are essential steps to ensure the quality of the results.

Second, the design of immersive learning environments can also be enhanced by using conversational agents that are driven by artificial intelligence. Large language models are excellent tools to facilitate human-computer interactions for deep learning. The integration of conversational agents in immersive learning environments also observed creative applications such as in public health settings (So et al., 2023) and teacher education (Dai, 2023; Ke et al., 2021). Training the large language models with localized authentic learning data can ensure the educational values of such learning technologies (see Figure 2).

Third, in addition to augmenting the designs of immersive learning environments, assessment is another salient aspect of machine learning applications in immersive learning environments. Integrating unobtrusive assessments in immersive learning environments helps to provide feedback and understand the learning outcomes. Machine learning is an ideal tool to accomplish these goals with a data-driven approach that tailors to individual differences, considering *all* learners. That said, to create equal, inclusive, and just assessment in immersive learning environments, the diversity and inclusiveness of the training data is critical. There were different approaches to assessment in immersive learning environments. ECD was one of the pioneering approaches. Novel processes with machine learning algorithms generate and extract practice-based evidence and metrics but not necessarily with an *a priori* competency model (Winkler-Schwartz et al., 2019). Finally, considering the characteristics of immersive learning environments, multimodal learning analytics are prominent future directions. Multimodal learning analytics use tracking devices to provide information about learners' actions and behaviors and therefore automatic feedback can be provided to individual students based on their performance.

The future of machine learning-integrated immersive learning environments is promising. While existing learning principles and theories from a constructivist perspective can explain machine learning integration in immersive learning well, as the learning landscape grows increasingly complex (e.g., in metaverse), it is necessary to develop novel learning principles and theories to explain learning in such intricate settings. Moreover, the design and assessment of multimodal learning in immersive learning environments are emerging areas that require further development and exploration. Future research is needed to advance these areas. Finally, machine learning-integrated immersive learning environments should be designed with learners' datasets that are inclusive, diverse, and just.

5. Conclusion

By introducing modern approaches of machine learning integration into the design and assessment for immersive learning environments, this chapter provides useful heuristics and applications for learning designers, researchers, and policy and decision makers to consider when using these technologies. The integration of modern machine learning approaches, along with conversational artificial intelligence, into the design and assessment of immersive learning environments holds great promise for educational sciences. By harnessing the power of data-driven techniques, immersive learning experiences can be enhanced in ways that foster more engaging, personalized, and inclusive learning experiences for *all* learners to be successful.

References

Abouelenein, Y. A. M., & Nagy Elmaadaway, M. A. (2023). Impact of Teaching a Neuro-Computerized Course Through VLE to Develop Computational Thinking Among Mathematics Pre-service Teachers. *Journal of Educational Computing Research*, Advance online publication. https://doi.org/10.1177/07356331231165099

Amershi, S., & Conati, C. (2009). Combining unsupervised and supervised classification to build user models for exploratory learning environments. *Journal of educational data mining, 1*(1), 18-71. https://doi.org/10.5281/zenodo.3554659

Asbee, J., Kelly, K., McMahan, T., & Parsons, T. D. (2023). Machine learning classification analysis for an adaptive virtual reality Stroop task. *Virtual Reality*, 1-17. https://doi.org/10.1007/s10055-022-00744-1

Bahceci, O., Pena-Rios, A., Buckingham, G., & Conway, A. (2022). Supervised Machine Learning Hand Gesture Classification in VR for Immersive Training. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 748-749). IEEE. https://doi.org/10.1109/VRW55335.2022.00225

Bassen, J., Balaji, B., Schaarschmidt, M., Thille, C., Painter, J., Zimmaro, D., ... & Mitchell, J. C. (2020). Reinforcement learning for the adaptive scheduling of educational activities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-12). http://dx.doi.org/10.1145/3313831.3376518

Bhowmik, S., Barrett, A., Ke, F., Yuan, X., Southerland, S., Dai, C-P., West, L., & Dai, Z. (2022). Simulating students: An AI chatbot for teacher training. In Chinn, C., Tan, E., Chan, C., & Kali, Y. (Eds.). *Proceedings of the 16th International Conference of the Learning*

Sciences (ICLS) 2022 (pp. 1972-1973). Hiroshima, Japan: International Society of the Learning Sciences. https://repository.isls.org/bitstream/1/8669/1/ICLS2022_1972 -1973.pdf

Blikstein, P. (2013). Multimodal learning analytics. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 102-106). https://doi.org/10.1145/2460296.2460316

Boulet, J. R., & Durning, S. J. (2019). What we measure… and what we should measure in medical education. *Medical Education, 53*(1), 86-94. https://doi.org/10.1111/medu.13652

Chen, G., Yang, J., Hauff, C., & Houben, G. J. (2018). LearningQ: a large-scale dataset for educational question generation. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 12, No. 1). 481-490. https://doi.org/10.1609/icwsm.v12i1.14987

Cukurova, M., Giannakos, M., & Martinez-Maldonado, R. (2020). The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology, 51*(5), 1441-1449. https://doi.org/10.1111/bjet.13015

Dai, C-P. & Ke, F. (2022). Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review. *Computers & Education: Artificial Intelligence, 3*, 100087. https://doi.org/10.1016/j.caeai.2022.100087

Dai, C-P., Ke, F., Dai, Z., & Pachman, M. (2023). Improving teaching practices via virtual reality-supported simulation-based learning: Scenario design and the duration of implementation. *British Journal of Educational Technology*. Advance online publication. https://doi.org/10.1111/bjet.13296

Dai, C-P., Ke, F., Dai, Z., West, L., Bhowmik, S., & Yuan, X. (2021). Designing artificial intelligence (AI) in virtual humans for simulation-based training with graduate teaching assistants. In de Vries, E., Hod, Y., & Ahn J. (Eds.). *Proceedings of the 15th International Conference of the Learning Sciences - ICLS 2021* (pp. 1101-1102). Bochum, Germany: International Society of the Learning Sciences. https://repository.isls.org/bitstream/1/7418/1/1101-1102.pdf

Dai, C-P., Ke, F., Pan, Y., & Liu, Y. (2023). Exploring students' learning support use in digital game-based math learning: A mixed-methods approach using machine learning and multi-cases study. *Computers & Education, 194*, 104698. https://doi.org/10.1016/j.compedu.2022.104698

Dai, C-P., Ke, F., Pan, Y., Moon, J., & Liu, Z. (2023, April). *A meta-analysis on the effects of using artificial intelligence-powered virtual agents in simulation-based learning*. Paper Session presented at the 2023 AERA Annual Meeting. Chicago, IL.

Dai, C-P., Ke, F., Zhang, N., Barrett, A., West, L., Bhowmik, S., Southerland, S. A., & Yuan, X. (2024). Designing Conversational Agents to Support Student Teacher Learning in Virtual Reality Simulation: A Case Study. In *Proceedings of The ACM CHI conference on Human Factors in Computing Systems* (ACM CHI '24). Honolulu, HI. https://doi.org/10.1145/3613905.3637145

Dai, C.-P. (2023). *Enhancing Learning Achievements and Self-Efficacy for Preservice Teachers Using Model-Based Support in Simulation-Based Learning with Artificial Intelligence-Powered Virtual Agents*. Doctoral dissertation. Florida State University.

de Jong, T., Lazonder, A. W., Chinn, C. A., Fischer, F., Gobert, J., Hmelo-Silver, C. E., ... & Zacharia, Z. C. (2023). Let's talk evidence–The case for combining inquiry-based and direct instruction. *Educational Research Review*, 100536. https://doi.org/10.1016/j.edurev.2023.100536

Delamarre, A., Shernoff, E., Buche, C., Frazier, S., Gabbard, J., & Lisetti, C. (2021). The interactive virtual training for teachers (IVT-T) to practice classroom behavior management. *International Journal of Human-Computer Studies, 152*, 102646. https://doi.org/10.1016/j.ijhcs.2021.102646

Driscoll, M. P., & Burner, K. J. (2021). *Psychology of learning for instruction* (4th edition). Pearson.

Durning, S. J., & Artino, A. R. (2011). Situativity theory: a perspective on how participants and the environment can interact: AMEE Guide no. 52. *Medical teacher, 33*(3), 188-199. https://doi.org/10.3109/0142159X.2011.550965

Gencoglu, B., Helms-Lorenz, M., Maulana, R., Jansen, E. P., & Gencoglu, O. (2023). Machine and expert judgments of student perceptions of teaching behavior in secondary education: Added value of topic modeling with big data. *Computers & Education, 193*, 104682. https://doi.org/10.1016/j.compedu.2022.104682

Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age?. *Medical teacher, 38*(10), 1064-1069. https://doi.org/10.3109/0142159X.2016.1173661

Graesser, A. C., Li, H., & Forsyth, C. (2014). Learning by communicating in natural language with conversational agents. *Current Directions in Psychological Science, 23*(5), 374-380. https://doi.org/10.1177/0963721414540680

Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist, 45*(4), 234–244. https://doi.org/10.1080/00461520.2010.515933.

Guyon, I., & Elisseeff, A. (2006). An introduction to feature extraction. *Feature extraction: foundations and applications*, 1-25.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hwang, G.-J., & Chien, S.-Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. *Computers and Education: Artificial Intelligence, 3*, 100082. https://doi.org/10.1016/j.caeai.2022.100082

Ip, H. H. S., Li, C., Leoni, S., Chen, Y., Ma, K. F., Wong, C. H. T., & Li, Q. (2018). Design and evaluate immersive learning experience for massive open online courses (MOOCs). *IEEE Transactions on Learning Technologies, 12*(4), 503-515. http://doi.org/10.1109/TLT.2018.2878700

Johnson, W. L. (2003). Using agent technology to improve the quality of web-based education. In *Web intelligence* (pp. 77–101). Springer.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255-260. https://doi.org/10.1126/science.aaa8415

Ke, F, Shute, V, Clark, K., & Erlebacher, G. (2019). *Interdisciplinary design of the game-based learning platform: A phenomenological examination of the integrative design of game, learning, and assessment*. Springer.

Ke, F. F., & Shute, V. J. (2015). Design of game-based stealth assessment and learning support. In C. Loh, Y. Sheng, & D. Ifenthaler (Eds.), *Serious games analytics* (pp. 301-318). Springer.

Ke, F., Dai, Z., Dai, C-P., Pachman, M., Chaulagain, R., & Yuan, X. (2020). Designing virtual agents for simulation-based learning in virtual reality. In R. Zheng (Ed.), *Cognitive and Affective Perspectives on Immersive Technology in Education* (pp. 151-170). IGI Global. https://doi.org/10.4018/978-1-7998-3250-8.ch008

Ke, F., Yuan, X., Southerland, S. (2021). Teaching Practices with Multiplayer Mixed Reality Simulations and Virtual Students. *National Science Foundation (Award Abstract # 2110777)*. Retrieved from: https://www.nsf.gov/awardsearch/showAward?AWD_ID=2110777

Kolb, D. A. (1984). *Experiential learning experience as a source of learning and development*. Prentice Hall.

Kroeze, K. A., Van Den Berg, S. M., Veldkamp, B. P., & De Jong, T. (2021). Automated Assessment of and Feedback on Concept Maps during Inquiry Learning. *IEEE transactions on learning technologies, 14*(4), 460-473. https://doi.org/10.1109/TLT.2021.3103331

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lee-Cultura, S., Sharma, K., & Giannakos, M. (2022). Children's play and problem-solving in motion-based learning technologies using a multi-modal mixed methods approach. *International Journal of Child-Computer Interaction, 31*, 100355. https://doi.org/10.1016/j.ijcci.2021.100355

Lee, S. Y., Rowe, J. P., Mott, B. W., & Lester, J. C. (2013). A supervised learning framework for modeling director agent strategies in educational interactive narrative. *IEEE Transactions on Computational Intelligence and AI in Games, 6*(2), 203-215. https://doi.org/10.1109/TCIAIG.2013.2292010

Littenberg-Tobias, J., Borneman, E., & Reich, J. (2021). Measuring equity-promoting behaviors in digital teaching simulations: A topic modeling approach. *AERA Open, 7*, 1-19. https://doi.org/10.1177/23328584211045685

Liu, Z., Moon, J., Kim. B., & Dai, C-P. (2020). Integrating adaptivity in educational games: a combined bibliometric analysis and meta-analysis review. *Educational Technology Research & Development, 68*(4), 1931-1959. https://doi.org/10.1007/s11423-020-09791

Livieris, I. E., Drakopoulou, K., Tampakas, V. T., Mikropoulos, T. A., & Pintelas, P. (2019). Predicting secondary school students' performance utilizing a semi-supervised learning approach. *Journal of educational computing research, 57*(2), 448-470. http://doi.org/10.1177/0735633117752614

Lorenzo, G., Pomares, J., & Lledó, A. (2013). Inclusion of immersive virtual learning environments and visual control systems to support the learning of students with Asperger syndrome. *Computers & Education, 62*, 88–101. https://doi.org/10.1016/j.compedu.2012.10.028

Luan, H., & Tsai, C. C. (2021). A review of using machine learning approaches for precision education. *Educational Technology & Society, 24*(1), 250-266. https://www.jstor.org/stable/26977871

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary research and perspectives, 1*(1), 3-62. https://doi.org/10.1207/S15366359MEA0101_02

Nye, B. D., Davis, D. M., Rizvi, S. Z., Carr, K., Swartout, W., Thacker, R., & Shaw, K. (2021). Feasibility and usability of MentorPal, a framework for rapid development of virtual mentors. *Journal of Research on Technology in Education, 53*(1), 21–43. https://doi.org/10.1080/15391523.2020.1771640.

Ouhaichi, H., Spikol, D., & Vogel, B. (2023). Research trends in multimodal learning analytics: A systematic mapping study. *Computers and Education: Artificial Intelligence, 4.* 100136. https://doi.org/10.1016/j.caeai.2023.100136

Philippe, S., Souchet, A. D., Lameras, P., Petridis, P., Caporal, J., Coldeboeuf, G., & Duzan, H. (2020). Multimodal teaching, learning and training in virtual reality: a review and case study. *Virtual Reality & Intelligent Hardware, 2*(5), 421-442. https://doi.org/10.1016/j.vrih.2020.07.008

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional science, 23*(5-6), 405-431. https://doi.org/10.1007/BF00896880.

Querrec, R., Buche, C., Maffre, E., & Chevaillier, P. (2004). Multiagents systems for virtual environment for training. application to fire-fighting. *International Journal of Computers and Applications, 1* (1), 25–34.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v091.i02

Rogers, M. P., DeSantis, A. J., Janjua, H., Barry, T. M., & Kuo, P. C. (2021). The future surgical training paradigm: Virtual reality and machine learning in surgical education. *Surgery, 169*(5), 1250-1252. https://doi.org/10.1016/j.surg.2020.09.040

Shah, D., Patel, D., Adesara, J., Hingu, P., & Shah, M. (2021). Exploiting the capabilities of blockchain and machine learning in education. *Augmented Human Research, 6*, 1-14. https://doi.org/10.1007/s41133-020-00039-7

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational researcher, 29*(7), 4-14. https://doi.org/10.3102/0013189X029007004

Shepard, L. A. (2021). Ambitious Teaching and Equitable Assessment: A Vision for Prioritizing Learning, Not Testing. *American Educator, 45*(3), 28-37. Retrieved from https://files.eric.ed.gov/fulltext/EJ1321974.pdf

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational measurement: issues and practice, 37*(1), 21-34. https://doi.org/10.1111/emip.12189.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research, 78*(1), 153-189. https://doi.org/10.3102/0034654307313795

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503-524). Information Age Publishers.

Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In D. Ifenthaler, D. Eseryel, & Ge, X. (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 43-58). Springer.

Shute, V., Rahimi, S., Smith, G., Ke, F., Almond, R., Dai, C. P., ... & Sun, C. (2021). Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games. *Journal of Computer Assisted Learning, 37*(1), 127-141. https://doi.org/10.1111/jcal.12473

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning, 2*(1). Retrieved from http://www.itdl.org/

Sinatra, A. M., Pollard, K. A., Files, B. T., Oiknine, A. H., Ericson, M., & Khooshabeh, P. (2021). Social fidelity in virtual agents: Impacts on presence and learning. *Computers in Human Behavior, 114*, 106562. https://doi.org/10.1016/j.chb.2020.106562

So, C., Khvan, A., & Choi, W. (2023). Natural conversations with a virtual being: How user experience with a current conversational AI model compares to expectations. *Computer Animation and Virtual Worlds, e2149*. https://doi.org/10.1002/cav.2149

Stenudd, S. (2010). *Using machine learning in the adaptive control of a smart environment*. Master's thesis. The VTT Technical Research Centre of Finland in the Software Architectures and Platforms Knowledge Centre.

Thormundsson, B. (2022, June 28). *Machine learning - Statistics & Facts*. Statista. https://www.statista.com/topics/9583/machine-learning/#topicOverview

Valdez, D., Picket, A. C., Young, B. R., & Golden, S. (2021). On mining words: the utility of topic models in health education research and practice. *Health Promotion Practice, 22*(3), 309-312. https://doi.org/10.1177/1524839921999050

Vaughan, N., Gabrys, B., & Dubey, V. N. (2016). An overview of self-adaptive technologies within virtual reality training. *Computer Science Review, 22*, 65-87. https://doi.org/10.1016/j.cosrev.2016.09.001.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological process*. Harvard University Press.

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36-45. https://dl.acm.org/doi/pdf/10.1145/365153.365168

Westera, W., Dascalu, M., Kurvers, H., Ruseti, S., & Trausan-Matu, S. (2018). Automated essay scoring in applied games: Reducing the teacher bandwidth problem in online training. *Computers & Education, 123*, 212–224. https://doi.org/10.1016/j.compedu.2018.05.010

Winkler-Schwartz, A., Yilmaz, R., Mirchi, N., Bissonnette, V., Ledwos, N., Siyar, S., ... & Del Maestro, R. (2019). Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. *JAMA network open, 2*(8), e198363-e198363. https://10.1001/jamanetworkopen.2019.8363

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet?-A systematic literature review on chatbots in education. *Frontiers in artificial intelligence, 4*, 654924. https://doi.org/10.3389/frai.2021.654924

Wu, J. Y., Hsiao, Y. C., & Nian, M. W. (2020). Using supervised machine learning on large-scale online forums to classify course-related Facebook messages in predicting learning achievement within the personal learning environment. *Interactive Learning Environments, 28*(1), 65-80. https://doi.org/10.1080/10494820.2018.1515085

Zahabi, M., & Abdul Razak, A. M. (2020). Adaptive virtual reality-based training: a systematic literature review and framework. *Virtual Reality, 24*, 725-752. https://doi.org/10.1016/j.cosrev.2016.09.001