RAND CORPORATION

Perspective

Expert insights on a timely policy issue

January 2019

Artificial Intelligence Applications to Support K–12 Teachers and Teaching

A Review of Promising Applications, Opportunities, and Challenges

Robert F. Murphy

The successes of recent applications of artificial intelligence (AI) in performing complex tasks in health care, financial markets, manufacturing, and transportation logistics have been well documented in the academic literature and popular media. The increasing availability of large digital data sets, refined statistical techniques, and advances in machine-learning algorithms and data processing hardware, coupled with large sustained corporate investments, have led to dramatic gains in speech, image, and object recognition. In turn, these gains have enabled transformational advances in technologies impacting everyday lives; such advances include autonomous driving capabilities, intelligent virtual assistants (e.g., Apple's Siri, Amazon's Alexa, Google Assistant), medical imaging diagnostics, text-to-text language translation, and speech-to-text applications.[1] In comparison, the influence of AI applications in the education sphere, which first appeared almost four decades ago, has been limited. However, increasingly, product developers, AI researchers, education technology advocates, and venture capitalists are turning their attention to education and speculating about ways that advanced AI techniques, such as "deep" machine learning, may dramatically shape the future of kindergarten through grade 12 (K–12) education, including classroom instruction, the role of the teacher, and how students learn.[2]

Although various AI advocates are currently touting a myriad of new applications for K–12 education, there is little evidence yet to support the usefulness of these applications to districts, schools, and teachers. In light of this evidence gap, in this paper, I identify several existing applications of AI that have shown promise in helping teachers address important challenges in the classroom and that may point the way to how advances in AI techniques might be used to provide value to teachers by augmenting their capacity.[3] In addition to highlighting the promise of these technologies, I discuss some of the key technical challenges that need to be addressed in order to realize the full potential of AI to support teachers.

The scope of this paper is limited to AI applications designed to address any one of three core challenges of teaching: (1) providing differentiated instruction and feedback in mixed-ability classrooms, (2) providing students with timely feedback on their writing products, or (3) identifying students who may be struggling to learn and make progress toward graduation.[4] In each case, I discuss the aspects of the challenge that make it particularly suitable for an AI-based solution and the conditions necessary for the application of advanced machine-learning AI applications.

In the first section of this paper, I define AI. In the second section, I review applications of AI in schools to support instruction and learning. And in the third section, I offer recommendations that product developers, publishers, district and school administrators, and researchers will need to carefully consider as they contemplate the application of advanced AI methods to support K–12 teachers.

What Is Artificial Intelligence?

Throughout this paper, the term *artificial intelligence* refers to applications of software algorithms and techniques that allow computers and machines to simulate human perception and decisionmaking processes to successfully complete tasks. AI has been applied to simpler tasks, such as sending automated phone calls and texts from banks when an unusual transaction appears on someone's account, and more-complex tasks, such as allowing an automobile with advanced driver-assistance systems to automatically stay in its lane and keep a safe distance from the vehicle immediately in front of it.[5] The systems take in data relevant to the task, typically from sensors in the environment or from a prepopulated database; process the data through the system's statistical algorithms; generate a prediction or decision; and then, in some applications, convert that prediction into a recommendation for the user or an action for a piece of machinery (e.g., an automobile or robot). AI-based applications are currently being used to classify and recognize images on the internet; compose original media content, including music and news articles; and predict the likelihood of outcomes, such as next week's weather, a customer's emotional state, the likelihood that a particular student will graduate from college on time, and the next movie a person might want to rent from Netflix.

The current AI applications used in education, as well as in other fields, are examples of what the AI community calls *narrow* or *weak* AI. This term refers to AI applications that use software code (or algorithms) to perform a single, specific function, such as a chat bot responding to a customer's question or a driverless vehicle distinguishing between a stop sign and a yield sign. In addition, narrow AI includes home-based virtual assistants, such as Siri and Alexa, as well as IBM's Watson, which is one of the most sophisticated narrow applications of AI and is currently deployed in a variety of commercial applications.[6]

Abbreviations

AI	artificial intelligence
AES	automated essay scoring
ITS	intelligent tutoring system
K–12	kindergarten through grade 12

As long as these applications are used within the narrow contexts in which they were designed to operate and learn, their performance is quite accurate and reliable. However, once the applications are used in a new or different context, they can become prone to error and have limited utility. A good example of this is the inability of most AI-based chat bots used in customer service applications to respond to users' novel questions or commands. In these cases, when a user asks a question or gives a verbal or text command that is outside the set of questions and commands on which an application was trained, the application will opt to connect the customer with a human service agent.

As long as these applications are used within the narrow contexts in which they were designed to operate and learn, their performance is quite accurate and reliable. However, once the applications are used in a new or different context, they can become prone to error and have limited utility.

In contrast to narrow AI applications, *strong* AI or AI applications that exhibit *general intelligence* are considered the holy grail of the AI field. These are theorized applications that approach the cognitive reasoning capabilities of humans, demonstrate a common sense understanding of how the world works, are able to solve novel problems or perform novel tasks, and can *learn* with little or no prior data or information about the current problem or task context. At this time, AI systems that exhibit general intelligence are still an aspiration of the AI community.

The remainder of this paper focuses on successful and promising applications of narrow AI in education to augment teacher capacity, highlighting both their benefits to teaching and their current limitations.

Artificial Intelligence in Education

There are two broad categories of narrow AI that have been applied thus far in education. The first category encompasses rule-based applications used to power adaptive instructional software systems. The second encompasses applications that use machine-based learning, such as those used in the automated scoring of student essays.

Rule-Based Expert Systems

Research on AI got its start in the 1950s with funding primarily from the U.S. Department of Defense. One of the early products of this work was the development of *rule-based expert systems* (that is, systems that mimic the decisionmaking ability of human experts) to support military decisionmaking and planning.[7] Commercial applications of expert systems began in earnest in the 1970s, including the custom design of computers based on client specifications, maintenance diagnostics for heavy machinery and oil-drilling operations, and field support for service technicians and analysts assessing personal credit risk.[8] Two key components of expert systems are the *knowledge base* (collection of encoded expert knowledge and experience necessary for problem-solving in a particular domain, often in the form of if-then statements or rules) and the *inference engine* (programmable decision rules that are applied to incoming streams of data about a specific case and that produce a recommendation or solution that is fully explainable by the rules in the expert knowledge base). In most applications, rule-based expert systems support decisionmaking processes that are well understood and from which decision rules can almost always be derived. Building and evaluating these expert systems requires extensive programming, access to expert knowledge, and reliable and accurate measures of the input and output variables on which the decisionmaking process depends (e.g., temperature, interest rates, medical diagnoses).

Intelligent Tutoring Systems

Intelligent tutoring systems (ITSs) are an early application of rule-based expert systems in education; the first ITS appeared in schools in the early 1980s.[9] The original goal of early ITS research and development was to simulate the instructional experience and interactions between a student and a human tutor or coach.[10] An ITS adjusts the content presented to each student based on the student's current state of knowledge in a particular domain, such as in mathematics, and provides the level of support and feedback needed so that the student can learn and progress through the content. Because of the personalized nature of the learning environment within an ITS, many of these systems are used in schools to help teachers accommodate a wide range of student abilities in heterogenous classrooms—a well-known challenge for many teachers.

The original goal of early ITS research and development was to simulate the instructional experience and interactions between a student and a human tutor.

Some ITS applications, such as those used to power an online course, are designed to be used as the primary mode of instruction, while others are designed to be integrated into instructor-led classrooms or used as a separate in-class or homework activity.[11] Common features of many of the ITS applications available in the education market today include mastery learning of individual concepts and skills (i.e., students are required to demonstrate a predetermined level of understanding of a concept before moving on to the next concept in the sequence), self-pacing, instructional content that adapts to the individual student's knowledge state, frequent assessment of student knowledge and understanding, continuous monitoring of student progress as students demonstrate mastery of individual concepts and skills in the content domain, and automated feedback related to student performance within a task. Many of today's online adaptive learning platforms (e.g., ALEKS, MATHia, Dreambox Learning, STMath, Achieve3000) use an ITS rule-based AI architecture, although the sophistication and comprehensiveness of the various models vary by system. Like all expert systems, an ITS requires access to experts and extensive programming to represent the expert domain knowledge and reasoning in the system.

A typical ITS architecture comprises three interacting models: the domain model, the student model, and the tutoring or pedagogical model. The *domain model* (also known as the cognitive or expert knowledge model) captures all of the concepts and skills to be learned in a domain (e.g., algebra) and their interrelationships. It also includes the correct solution steps to the problem-based tasks that are given to students to demonstrate their knowledge and skills in the domain. Content domains are currently limited to those where demonstration of knowledge in a domain can be reduced to the successful solution of multistep problems or tasks that require students to learn and apply a set of rules or strategies determined by experts. Such content areas include (1) reading and math in primary and secondary education and (2) statistics, physics, and computer science in postsecondary education, making these the most suitable domains for rule-based AI instructional approaches.[12]

The ITS *student model* uses student responses to problem-based tasks and statistical models of student learning to estimate and monitor the student's current state of knowledge of the concepts and skills in the domain. Typically, student-learning data are captured at the subconcept or micro-skill level associated with the individual steps in a multistep problem solution, and the system flags whether a student has entered a correct response to each step. The student model may also capture data on student task performance, including the number of tasks completed, the time needed to complete the tasks, and the number of errors made.[13]

The *tutor model* then takes inputs from the domain and student models to determine how to interact with the student to help improve his or her performance based on which knowledge elements the student has learned or not learned and what feedback or additional instruction a student needs after answering a particular problem or problem step incorrectly or using an incorrect strategy. The level, type, specificity, and timing of the feedback provided by the tutor model is determined by the developer and varies by system. Some systems provide feedback immediately on the correctness of the student response at each step, while others provide feedback only after students complete all steps in a task. Some systems also provide error-specific feedback to help a student learn from a mistake, automated hints after one or more incorrect responses to a particular problem step, or hints only at the request of the student.[14]

When students attempt but do not demonstrate mastery of a concept, the system typically provides additional opportunities to do so, in the form of a new problem or task. Depending on the ITS, before providing additional opportunities, the system may provide a remedial instruction activity, ask the student to review instruction on prerequisite concepts and skills, or provide a model example of how an expert would have solved the problem. Once a concept is mastered, the tutor model then allows the student to advance to the next knowledge element in the domain model's concept map, typically of increasing difficulty. Although these are general capabilities of an ITS, the extent to which these features are present in any one ITS varies by system.

Research on the Effectiveness of Intelligent Tutoring Systems

A 2014 review of the effectiveness research on a variety of ITSs found that these systems can be relatively effective sources of classroom instruction and support for student learning for topics that are amenable to a rule-based AI architecture.[15] The meta-analysis reviewed the research going back to 1997 and covered a range of content areas in K–12 and higher education—primarily math, physics, computer science, language, and literacy. The authors found that, when they compared scores on standardized or researcher-developed tests, ITS-based instruction (1) resulted in higher test scores than did traditional formats of teacher-led instruction and non-ITS online instruction and (2) produced learning results similar to one-on-one tutoring and small-group instruction.[16] In general, these results held across grade levels (elementary through higher education), content domains, and study quality (e.g., randomized controlled trials and quasi-experiments).

Limitations of Intelligent Tutoring Systems in Education

There are several important limitations to the use of ITSs in education. As mentioned previously, the target subject area must be amenable to a rule-based AI architecture. As a result, the subject areas that have been targeted by ITS developers have been primarily constrained to math, literacy, the physical sciences, and computer science. In addition, within a subject area, ITSs are best suited to support the learning of aspects of the content that are appropriate for rule-based approaches, including facts, methods, operations, algorithms, and procedural skills. However, the systems are less able to support the learning of complex, difficult-to-assess, higher-order skills—such as critical thinking, effective communication, explanation, argumentation, collaboration, self-management, social awareness, and professional ethics—that are increasingly emphasized in state education standards and valued by employers.

The self-paced and mastery-learning features of most ITSs that allow such a system to accommodate a range of different learners and abilities can also pose challenges for teachers who want to integrate ITS instruction as an in-class activity that is part of a broader coherent curriculum. In a classroom of students with diverse competencies in a particular subject area, the students will progress through the ITS content at different rates. This can make it difficult for teachers to align the ITS's content and instruction with teacher-led group instruction. Although some ITS applications are designed to be modular and allow teachers to assign students discrete units of ITS content that align with their daily lessons, other applications are closed systems and do not have this capability. As a result, teachers often relegate ITS-based instruction to independent learning time, in which students use the ITS to help remediate their own skills, gain exposure to advanced topics, or complete homework. Finally, ensuring that all students are making adequate progress in an ITS learning environment requires careful monitoring by the teacher. Although all systems provide some level of automated feedback and support to students and attempt to adapt the content to meet the needs of individual students, the level of automated support provided by the systems may be insufficient to support the learning of all students. Thus, district and school administrators must allocate time for teachers to regularly review student progress within an ITS, using the system's reports on student performance to identify students who may be struggling to make progress and intervene before these students experience frustration, lose initiative, and disengage.

Machine Learning

In contrast to deterministic rule-based expert systems, *machine learning* is a technical approach to AI that uses statistical algorithms to build (or *learn*) a prediction model by processing large amounts of multivariate data related to the phenomena of interest.[17] Instead of humans programming a set of expert rules into the system, the system discovers patterns among the predictor variables, or *features*, on the one hand and the output variable of interest on the other. For example, a machine-learning approach might be used to identify the relationships between student characteristics from their early school years (e.g., school attendance, credits earned, and early test scores, all of which would be considered predictor variables, or input features of the system) and subsequent on-time graduation from high school, an output variable of interest. By doing so, the system can learn whether there are early signs that a student will likely drop out at a future date, and educators can then try to intervene. This kind of early warning system is described later in this paper.

With machine learning, the goal is to create a model that can accurately predict the outcome for a set of input data that the application has not seen before. Applications of machine-learning approaches are most suitable for tasks that have complex solutions; depend on many different factors; and, in contrast to rule-based AI techniques, cannot be addressed with a simple computation or the coding of predetermined rules.

Although commercial applications of machine-learning techniques first appeared in the 1990s, recent advances in hardware processing speeds and access to large stores of digital data generated by various sources—including internet search engines, social media websites, online shopping platforms, and digital medical records and instrumentation—have accelerated applications across many fields. To build a statistical model that can reliably predict outcomes (e.g., whether an X-ray image contains a tumor) requires access to existing data sets to train the system and to independently validate (or *ground truth*) the accuracy of the predicted outcomes. This process of training and validation is called *supervised learning*. Most current commercial AI applications that use machine learning use a supervised learning model.

The training data set must be sufficiently large to allow the system to use a portion of the data to build a prediction model and then use an independent portion of the data to evaluate the accuracy of the model's predictions compared with actual outcomes. Developers then use that information to optimize or tune the model to improve its accuracy. Depending on the complexity of the problem to be solved and the requirements of the statistical algorithms used, machine-learning applications that require highly

With machine learning, the goal is to create a model that can accurately predict the outcome for a set of input data that the application has not seen before.

accurate prediction capability may need access to thousands or millions of data records associated with a particular task to properly train a system to reach an acceptable level of performance. A requirement of the training set is that the data associated with the model's input features (e.g., type of treatment regimen and diet) and the desired target output variable (e.g., whether a patient recovered from an illness) must be accurately measured and correctly labeled.[18] In some instances, the data may need to be labeled by hand, a very time-intensive task, to create the training data set, or it may be possible, depending on the task, to write a software program to identify and automatically label features in the data set.[19] The training process also assumes stability over time in the relationship between input features and output variables. If the relationship changes and causes the predictions or recommendations to no longer be accurate or relevant (e.g., when an assessment that is being used to collect data on student skills changes or the system is being applied to a different type of student population), the model needs to be retrained and tested on an updated data set. Unlike humans, a machine-learning application has no common sense and thus no built-in ability to detect when its prediction models may no longer be relevant for the task it was designed to perform.[20]

Machine-learning techniques are able to efficiently process and recognize complex patterns in large data sets with thousands

Deep-learning computational models use artificial neural networks, which are interconnected layers of algorithms, to loosely simulate the processing capabilities of the human brain and recognize complex patterns in large multivariate data sets.

of input features. One of the most efficient and accurate types of machine-learning techniques is *deep learning*.[21] An idea originally conceived in the 1940s, deep-learning computational models use artificial neural networks, which are interconnected layers of algorithms, to loosely simulate the processing capabilities of the human brain and recognize complex patterns in large multivariate data sets.[22] Each layer consists of processing nodes that are connected with nodes in adjacent layers above and below it. Each node within a layer is responsible for performing the same relatively simple computation, applying weights to a string of incoming data variables (e.g., age, zip code, income) and producing a single output that then becomes the input signal for nodes in the layer above. Each layer of algorithms is responsible for reducing the complexity of the data for the layer above until, eventually, the final layer produces the system's predicted outcome (e.g., whether an applicant is a credit risk). By processing massive amounts of input and output data during the training stage, the deep neural network learns from its errors, comparing the predicted outcome with the known, actual outcome and making slight adjustments to the weights of each layer to minimize the prediction error. The typical deep-learning application is computationally intensive, requiring thousands or millions of examples for training and substantial computing power. Deep-learning applications that use supervised learning have made possible the significant gains in performance reported in several fields, including autonomous vehicles, image and voice recognition technologies, and automated language-to-language translation.[23] For many standard image-, voice-, and text-recognition tasks, deep-learning systems are beginning to exceed human performance.

Machine Learning in Education

Two of the most promising applications of machine-learning techniques in education are automated systems that score student essays and early warning detection systems that identify students who are struggling academically and at risk of dropping out and not graduating.[24]

Automated Essay Scoring

Automated essay scoring (AES) is one of the most mature applications of AI in education; the first commercial AES systems—including *Intellimetric*, developed by Vantage Learning, and the *e-rater* engine, developed by the Educational Testing Service—reached the market in the 1990s. Human reading and scoring of writing are extremely time-intensive, so many teachers are often reluctant to frequently assign extended writing projects of more than a few paragraphs. In addition, few teachers outside of the English department are trained to evaluate writing and provide feedback to students to help them improve their writing. At the same time, state standards for student knowledge have placed more emphasis on writing and communication, particularly in K–12 education; furthermore, across the grade levels, standardized assessments are becoming more writing-intensive. One of the primary motivations for developing AES applications was the need to score student writing, including assignments and exams for large, lecture-based college courses and entrance exams that are used in the admissions process for many higher education institutions. More recently, providers of massive open online courses (or MOOCs), including EdX, Coursera, and Udacity, have integrated automated scoring engines into their platforms to score the writing of the thousands of students who may be enrolled in a single course.[25] While many of the original AES systems returned an overall holistic writing quality score only, some current systems also provide students with basic feedback, guidance, and model writing samples to help students improve and revise their writing. These systems include, for example, the Educational Testing Service's Criterion Online Writing Evaluation Service, Turnitin's Revision Assistant, Pearson's Write to Learn, Grammarly, and Chegg's WriteLab. The type and specificity of the feedback vary by system.

AES and student writing analysis systems are made possible by a field of AI known as *natural language processing*. This field applies AI techniques, including machine learning, to the analysis of written language. Natural language processing algorithms, first developed in the 1950s, also make possible popular text-to-speech applications; language-to-language translation applications; and, when combined with speech-recognition algorithms, virtual personal assistants, such as Amazon's Alexa and Apple's Siri. Rather than attempting to program the specific rules of language, the AES systems' algorithms extract features of the text and, using supervised learning data from human-scored essays, learn the pattern of relationships between the features and different levels of writing quality or score points. Depending on the AES system and its goals for scoring and providing feedback, input features may include linguistic and nonlinguistic surface features (e.g., types and total number of grammatical errors, number of words, average word length), sentence-level qualities (e.g., use of passive voice, unnecessary words, use of concrete verbs), and essay-level qualities (e.g., coherence, style, organization). Some AES systems are designed to provide holistic scores, while others provide both scores and feedback on individual aspects of the writing, such as grammar, style, and mechanics. Typically, training a system takes several hundred to several thousand essays and hundreds of hours of labor for expert readers to annotate and score essays in the training set.

AES systems have their critics, and research has shown that it is possible to fool some AES systems to generate high scores with nonsensical writing (also known as *adversarial input*), thus raising concerns about their use in high-stakes testing situations.[26] However, many of these systems have been proven to perform similarly to human scorers on standard writing tasks.[27] AES systems will never replace the quality of the feedback that can be provided by a good writing teacher who has the time to carefully and thoughtfully critique a student's writing. But AES may make it possible for many teachers to assign more extended writing assignments for students with the assurance that the writing will be scored and that students will receive some form of timely feedback, primarily around writing style and conventions—something that may not be possible without AES.

Early Warning Systems

School districts' and administrators' use of student attendance, behavior, and course performance data to identify students who are at risk of dropping out and not graduating has become widespread in the past decade. A 2016 report based on a U.S. Department of

Education survey estimated that slightly more than half of public high schools in the United States had implemented such *early warning systems*.[28] Early warning systems are even more widely adopted in higher education: In 2014, one study found that an estimated 90 percent of four-year institutions have some type of system in place.[29] Historically, most early warning systems, an application of predictive analytics, used fairly simple, rule-based prediction models, monitoring one or more key measures that had been identified in the research literature as important indicators of students straying off track and dropping out (e.g., number of absences, course pass rates, number of disciplinary actions, cumulative grade point average, credits earned). The University of Chicago's Consortium on School Research was an early pioneer in the identification and use of *on-track* indicators for high school graduation in Chicago public schools.[30] Typically, when the indicators reach or drop below a certain threshold, the system flags the student, and someone from the institution may follow up with individualized support or other intervention.

While these simple warning systems may be highly effective in some contexts, the potential for the misclassification of students may have important negative consequences for students, teachers, and administrators.[31] As a result, some educational systems have begun to explore the use of machine learning to leverage the vast quantities of longitudinal data in student information systems to develop probabilistic models to identify at-risk students earlier in their school careers and to improve the accuracy of the warning systems.[32] The systems are trained on digital data archived by districts or higher education institutions from prior cohorts of students, which allows the machine-learning algorithms to determine the most-relevant indicators and their weights in the model. The machine-learning algorithms are then fed data about the current cohort of students to produce a probability score for each student—typically, the probability of dropping out of school before graduation.

Preliminary research has shown that, in at least one case, machine learning–based early warning systems for academics can improve the prediction accuracy offered by existing rule-based systems.[33] Researchers used machine-learning methods to develop an early warning indicator system to predict students who were at risk for not graduating high school on time because they would either drop out or need more than four years to receive a diploma. The research team trained the model on longitudinal data (grades 6 through 12) from 11,000 students attending a large U.S. school district. To test the relative precision of the system's predictions, the authors compared the accuracy of predicted outcomes for the machine-learning system and the district's existing rule-based system for students with the highest risk of not graduating on time (the top 10 percent with highest risk scores). For grades 10 through 12, the precision of the machine-learning systems was almost twice that of the rule-based system. For example, for 10th-grade students, 75 percent of those with the highest risk scores estimated by the machine-learning model did not graduate on time, compared with 38 percent of the students identified by the rule-based system.

Challenges and Risks with Machine-Learning Approaches

There are several potential challenges and risks associated with the use of machine-learning approaches to develop solutions for the classroom. In this section, I describe three: access to the appropriate data to train the models, biases in the models introduced through training data sets, and a lack of transparency into how the models work. Public concerns regarding model bias and transparency are particularly relevant for machine-learning applications designed to support decisionmaking in contexts that can have real consequences for peoples' lives and livelihoods, including those of students.

Access to Training Data on Teaching and Learning

Most developers interested in applying machine-learning techniques to develop intelligent, adaptive instruction products for the classroom lack access to the large digital data sets needed to train the models. These data sets must include high-quality measures of the relevant teaching and learning variables for the application's target population of schools and students. As in all fields in which machine-learning approaches are being proposed, the types of machine-learning applications that developers will pursue will be constrained by the types of data sets that they have access to or can accurately simulate. Although the student information systems of most large school districts and higher education institutions include a significant amount of digital data on students' family characteristics, courses taken, teachers, end-of-course grades, disciplinary and special education referrals, and standardized achievement scores (typically math and reading only), these systems do not include the fine-grained information on instruction and learning that is required to train a machine learning–based adaptive instruction system. Such data are typically available only from existing online instructional platforms, and access to these metadata is restricted to the district and the platform provider.

Data access issues are likely to be further complicated by the growing international movement to protect consumer data privacy and restrict how user-derived data and information are used and protected by technology companies, even when the data are anonymized. Following a string of widely publicized data breaches at several large corporations that involved the exposure of the personal information of millions of consumers, the European Union (via the 2016 General Data Protection Regulation) and the state of California (via the California Consumer Privacy Act, set to go into effect on January 1, 2020) have adopted regulations that formalize the digital data rights of consumers.[34] The current implications of these regulations for the education technology industry and the development of AI applications for education are yet to be determined. But these regulations, coupled with growing public awareness and concern about the protection and use of student data collected by education institutions and education technology companies, will likely make it more difficult for developers to access the data needed to train machine-learning models for advanced instructional systems.[35]

Learned Bias

Several high-profile incidents of racial and gender bias associated with decisions produced by machine-learning algorithms—including a facial recognition program that disproportionately

Data access issues are likely to be further complicated by the growing international movement to protect consumer data privacy and restrict how user-derived data and information are used and protected by technology companies, even when the data are anonymized.

misidentified some African-American and Latino U.S. senators as convicted felons—have revealed one of the most important potential limitations of machine-learning AI: The statistical models will encode the biases that are embedded in the training data.[36] In general, the quality of prediction models built with machine-learning approaches is tied to the quality of the data that those models are trained on and the biases and competency of the humans who are evaluating and correcting the models during the training phase. Biased decisionmaking algorithms can have serious implications for the people whose fate may depend on the output of these systems. Uses where this is particularly important include, for example, deciding which job applicant to interview or hire, determining the type of treatment that might be more effective for a particular cancer patient, deciding whether to grant parole to an incarcerated person in the criminal justice system, and assessing whether a loan applicant is worthy of credit.[37]

Of the three education applications discussed in this paper (ITSs, AES, and early warning systems), the type that is most vulnerable to potential bias is the early warning system. If certain groups of students are over- or under-represented in the training data set or are associated with a higher likelihood of dropping out because of structural biases or prejudices in greater society, these systems may overidentify one group as needing academic and social services based on individual surface features, such as race or gender. In these cases, the potential is for systems to misclassify students—as needing services when they do not (false positives) or as not needing services when they do (false negatives)—resulting in wasted resources and missed opportunities to intervene with individual students and help get them on track. Of course, one must also consider the biases introduced through the alternatives: early warning systems that do not use machine learning, instead relying on pure human judgment, and simple rule-based systems. These systems have their own biases and misclassification issues. One of the promises of moving to a machine-learning approach is that the prediction models will be more accurate (fewer false positives and negatives), have less bias, and potentially be more cost-effective than the alternatives, although there currently is not enough robust evidence from studies comparing alternative approaches to assess the validity of this claim.

The quality of prediction models built with machine-learning approaches is tied to the quality of the data that those models are trained on and the biases and competency of the humans who are evaluating and correcting the models during the training phase.

AI researchers at Microsoft, IBM, and elsewhere are exploring various strategies for identifying and eliminating bias in machine-learning applications.[38] At minimum, developers need to make sure that their training sets represent the diversity of the application's target population. Using the example of early warning systems, this means that the systems need to be trained separately for each district on data sets generated from each district's student information system. In addition, during the system training phase, the initial models must be carefully evaluated for potential bias and corrected as needed.

Transparency and the Trust Problem

The issues of model bias and misclassification associated with machine-learning applications, and particularly with deep-learning neural network approaches, are compounded by the inability to explain why a machine-learning application made a particular prediction.[39] This has important implications for the use of these applications in many fields, including education. Without an understanding of how a model arrived at a particular decision, it is difficult to identify the source of any bias and inaccuracies and then correct them. It also makes it difficult for users to trust the system's output, particularly when a prediction or recommendation is counterintuitive. For example, if an early warning system gives a student a probability score for dropping out that is inconsistent with the beliefs of an administrator or teacher based on their personal knowledge of the student and there is no information to explain how the system arrived at the prediction, it can easily cause educators to discount or mistrust the prediction. The importance of this issue has been acknowledged by many in the field. Researchers around the world, including at Google and the Defense Advanced Research Projects Agency, are aggressively pursuing strategies, known as *explainable AI*, to help increase the transparency of decision models while striking the right balance between transparency and system performance.[40]

Insights and Recommendations

The life of a teacher is demanding, especially in under-resourced school systems with large class sizes and heterogenous student abilities. In this paper, I summarize the most-promising AI applications to date that assist, not supplant, teachers, helping them be more effective at supporting student learning. The paper covers three kinds of AI-based applications—intelligent tutoring systems, automated essay scoring, and early warning systems—that can be used to support teachers and teaching, as well as the associated AI methods and their limitations. In this final section, I offer three recommendations for product developers, district and school administrators, and researchers to consider as they contemplate the role of AI in the classrooms of the future.

Publishers, product developers, and education administrators, recognizing that AI applications are not well suited to all content areas or the full array of educational activities in which teachers engage, should focus on applications that leverage the capabilities of AI to help solve important problems in teaching. The work of teachers and act of teaching, unlike repetitive tasks on the manufacturing floor, cannot be completely automated. Good teaching is complex and requires creativity, flexibility, improvisation, and spontaneity. At the same time, teachers need to be able to think logically and apply common sense, compassion, and empathy to deal with the everyday nonacademic issues and problems that arise in the classroom—abilities famously lacking in even the most-advanced AI systems. In addition to providing students with opportunities to develop narrow procedural knowledge and skills across a range of content areas (something that AI is particularly good at), schools and teachers must support the development of the whole child and provide students with rich opportunities to develop higher-order critical thinking and communication skills, as well as important social and emotional skills and mindsets (such as interpersonal skills, self-efficacy, and resiliency).

AI applications are best suited to tasks that are repetitive and predictable—with narrow, well-defined rules—or for look-

Research should focus on understanding the unintended consequences that these systems might have on instructional decisions and opportunities as a result of possible learned bias in the algorithmic models or of inaccuracies in model predictions, recommendations, and feedback.

ing for patterns in large multivariate data sets to help support decisionmaking for a specific purpose. Product developers should continue to look for opportunities within K–12 education that exploit these capabilities and that can have a significant impact on teaching and learning.

I have identified three areas in which AI-based solutions have shown promise for supporting teachers in challenging areas of instruction: adaptive instructional systems that allow teachers to differentiate instruction at the student level for certain topic areas and skills; automated scoring of student writing assignments, which supports teachers' ability to assign more writing in the classroom; and early warning systems, which alert administrators and teachers when students may need additional support to stay on track and progress toward graduation. Developers should focus on ways to apply advanced machine-learning techniques to improve existing capabilities in all three areas. The primary limiting factor will be developers' access to high-quality data sets for training that represent the populations of interest and that are unbiased.

Publishers and product developers should provide administrators, teachers, parents, and students with information that makes the workings and performance of machine-learning applications transparent. High-profile incidences of racial and gender bias associated with some machine-learning applications have brought the issue to the public's attention. Products being developed for the education market will likely come under greater scrutiny by administrators and parents, as well as by state, federal, and international regulators.[41] To gain the trust of system users, developers need to be transparent about the limitations and accuracies of their models; the consequences of inaccurate decisions for students and teachers; and how the models were trained, including details of the data sets used and how the learned models were evaluated for potential bias.

Independent and objective research is needed to understand the effects of advanced AI-based products on teaching and learning. Although machine-learning applications have made dramatic impacts in many fields, there currently is little evidence that these techniques are yet adding value in the classroom. This is mostly because such applications are currently in the early stages of product development and adoption. As the availability and adoption of various products begin to scale, new federal-, state-, and industry-funded research should focus on understanding the effects of the products on teaching and learning and the products' cost-effectiveness relative to existing approaches. In addition, research should focus on understanding the unintended consequences that these systems might have on instructional decisions and opportunities as a result of possible learned bias in the algorithmic models or of inaccuracies in model predictions, recommendations, and feedback.

The best use of AI in education is to augment teacher capacity by helping teachers deliver more-effective classroom instruction.

Applications of AI in the classroom, while promising, will be limited to a narrow set of instructional practices, supports, and topic areas like those highlighted in this paper. As a result, AI's overall influence on instruction and learning will likely be modest relative to its influence in other fields, such as autonomous transportation, medical diagnostics, robotics, and genomic research.

The most-effective AI applications will continue to play an assistive role, supporting rather than replacing teachers in their work with students in a limited set of content and topic areas that are most amenable to AI approaches. The most prevalent use cases will continue to be blended forms of instruction, in which the use of AI applications is integrated into teacher-led instruction and classroom activities. Advances in machine learning will likely lead to improvements in existing rule-based adaptive instruction systems, automated writing analyses, and early warning systems, although there is currently little robust evidence to support this claim. To leverage machine-learning capabilities for education as the AI field moves forward, product developers and publishers will need to address important challenges and concerns, including securing access to the relevant training data sets, navigating and complying with data privacy regulations, guarding against algorithmic bias, and improving model transparency to increase the confidence and trust of users.

Notes

1 Osoba and Welser, 2017b; Ng, 2016; Summerson, 2018; Bump, 2018; National Science and Technology Council, Committee on Technology, 2016.

2 Arnett, 2016; Dickson, 2017; Stone et al., 2016.

3 These types of assistive AI applications, where the application is used to help human actors (such as teachers) make better decisions and diagnoses and, in general, be more effective on the job, are also referred to as *augmented intelligence*, *human-machine teaming* (an initiative of the U.S. Department of Defense), and *clinical decision support systems* (in the medical field). Although, in other sectors, there are concerns over large-scale worker displacement resulting from advanced AI, including AI-based coding applications possibly displacing most programmers in the future, these same concerns are not being expressed with respect to teachers, particularly at the pre-kindergarten and K–12 school levels.

4 While the focus of this paper is on AI applications that directly support K–12 teachers, other promising applications in K–12 education target improving performance in the administrative and logistic functions of educational agencies and institutions, such as teacher recruitment and hiring, transportation, class scheduling, food services, and school safety and security.

5 Levy, 2018; Brynjolfsson and Mitchell, 2017; Knight, 2017a.

6 Lohr, 2016.

7 Researchers at the RAND Corporation were early pioneers in AI and expert systems. For a summary of this work, see Klahr and Waterman, 1986, Chapter 1.

8 Leonard-Barton and Sviokla, 1988.

9 McFarland and Parker, 1990.

10 Nkambou, Mizoguchi, and Bourdeau, 2010; Seidel and Park, 1994.

11 One of the most well-known and widely researched ITSs used in math education is the Cognitive Tutor Algebra (now known as MATHia), the product of research conducted in the 1990s by researchers at Carnegie Mellon University and now published by Carnegie Learning. A unique aspect of the MATHia curriculum is that it is a blended model, combining teacher-led whole-class instruction, small-group work, and individual time on the ITS platform. Approximately 40 percent of students' instructional time is spent on the online platform. During this time, teachers may be working with individual students one on one and providing instruction to small groups.

12 Ma et al., 2014.

13 VanLehn, 2006.

[14] VanLehn, 2006.

[15] Ma et al., 2014.

[16] Ma et al. (2014) reported impacts as an effect size, calculated as the difference between mean scores on the student outcome measure for the two groups divided by the pooled standard deviation for the overall sample. One or more effect sizes were estimated for each study reviewed, and then the effect sizes were averaged across studies by the type of comparison group—non-ITS computer-based instruction, teacher-led large-group instruction, teacher-led small-group instruction, individualized human tutoring, and textbooks or workbooks only. The reported average effect size was $g = +0.042$ standard deviation units, favoring the ITS condition, when achievement outcomes for students in the ITS group were compared with scores of students who received teacher-led large-group instruction; $g = +0.57$ when the comparison was non-ITS computer-based instruction; and $g = +0.35$ when the comparison was textbooks or workbooks only. These effects were all statistically different from 0. When the comparison condition was individualized human tutoring, the effect size was $g = -0.11$; when the comparison condition was small-group instruction, $g = +0.05$. These effects were not statistically significant.

[17] Brynjolfsson and Mitchell, 2017; National Science and Technology Council, Committee on Technology, 2016; Chen, 2016; Kwok, undated.

[18] Brynjolfsson and Mitchell, 2017.

[19] To make the process of hand-labeling large data sets more efficient, some developers have enlisted the help of crowd-sourcing platforms. See, for example, Chang, Amershi, and Kamar, 2017.

[20] Thompson, 2018.

[21] LeCun, Bengio, and Hinton, 2015.

[22] Osoba and Davis, 2018.

[23] National Science and Technology Council, Committee on Technology, 2016.

[24] IBM's Watson, which uses advanced machine-learning techniques, has been deployed in several education applications, including an open-education resource search tool for teachers (Teacher Advisor with Watson), an early-learning vocabulary application developed in partnership with Sesame Street, and instructional supports for students and instructors in higher education in a partnership with Pearson. See IBM Watson Education, undated. These applications are in the early stages of development and piloting and are not mature enough at this time to be adequately assessed for their utility or effectiveness.

[25] Stone et al., 2016.

[26] Powers et al., 2001; Kolowich, 2014.

[27] Shermis, 2014.

[28] Policy and Program Studies Service, 2016.

[29] Hanover Research, 2014.

[30] Allensworth and Easton, 2005.

[31] Gleason and Dynarski, 2002.

[32] Aguiar et al., 2015; Jayaprakash et al., 2014.

[33] Aguiar et al., 2015.

[34] European Union General Data Protection Regulation, undated; California Consumer Privacy Act, undated.

[35] Best and Pane, 2018. In 2014, after pressure by parent groups and privacy advocates, eight states rescinded their participation in a $100 million project funded by the Bill & Melinda Gates Foundation and the Carnegie Corporation to help districts and teachers make better use of student data. The groups raised concerns about sharing sensitive student information collected by public education authorities with a third-party corporation, InBloom, that was contracted to clean, store, and manage participants' detailed student data records in a cloud-based data warehouse and provide administrators and teachers with dashboards that would simplify the manipulation, analysis, and viewing of the data (Singer, 2014). In 2018, the Federal Bureau of Investigation issued a public service announcement warning the public about student data privacy and security risks associated with the use of education technologies in K–12 schools (Federal Bureau of Investigation, 2018).

[36] Osoba and Welser, 2017a; Bass and Huet, 2017; Singer, 2018.

[37] O'Neil, 2016; Angwin et al., 2016.

[38] Dickson, 2018.

[39] Osoba and Welser, 2017a; Knight, 2017b.

[40] Voosen, 2017.

[41] In 2018, the European Commission released a set of policy recommendations for the safe and secure use of AI and automated decisionmaking in consumer applications. The recommendations focus on security and liability, data protection and privacy, and transparency to ensure nondiscrimination (European Consumer Consultative Group, 2018).

References

Aguiar, Everaldo, Himabindu Lakkaraju, Nasir Bhanpuri, David Miller, Ben Yuhas, and Kecia L. Addison, "Who, When, and Why: A Machine Learning Approach to Prioritizing Students at Risk of Not Graduating High School on Time," in Association for Computing Machinery, *LAK '15, Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, Poughkeepsie, N.Y., 2015, pp. 93–102.

Allensworth, Elaine, and John Q. Easton, *The On-Track Indicator as a Predictor of High School Graduation*, Chicago: University of Chicago Consortium on School Research, 2005.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks," *ProPublica*, May 23, 2016. As of September 18, 2018: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Arnett, Thomas, *Teaching in the Machine Age: How Innovation Can Make Bad Teachers Good and Good Teachers Better*, Lexington, Mass.: Christensen Institute, December 2016. As of September 18, 2018: https://www.christenseninstitute.org/wp-content/uploads/2017/03/Teaching-in-the-machine-age.pdf

Bass, Dina, and Ellen Huet, "Researchers Combat Gender and Racial Bias in Artificial Intelligence," *Bloomberg*, December 4, 2017. As of September 18, 2018: https://www.bloomberg.com/news/articles/2017-12-04/researchers-combat-gender-and-racial-bias-in-artificial-intelligence

Best, Katharina Ley, and John F. Pane, *Privacy and Interoperability Challenges Could Limit the Benefits of Education Technology*, Santa Monica, Calif.: RAND Corporation, PE-313-RC, 2018. As of October 23, 2018: https://www.rand.org/pubs/perspectives/PE313.html

Brynjolfsson, Erik, and Tom Mitchell, "What Can Machine Learning Do? Workforce Implications," *Science*, Vol. 358, No. 6370, 2017, pp. 1530–1534.

Bump, Pamela, "Deep Learning in the Enterprise—Current Traction and Challenges," TechEmergence, 2018. As of September 18, 2018: https://www.techemergence.com/deep-learning-in-the-enterprise-current-traction-and-challenges

California Consumer Privacy Act, homepage, undated. As of October 30, 2018: https://www.caprivacy.org

Chang, J. C., S. Amershi, and E. Kamar, "Revolt: Collaborative Crowdsourcing for Labeling Machine Learning Datasets," in Association for Computing Machinery, *CHI '17: Proceedings of the 2017 ACM SIGCHI Conference on Human Factors in Computing Systems*, Denver, Colo., 2017, pp. 2334–2346.

Chen, Frank, "AI, Deep Learning, and Machine Learning: A Primer," presentation, Andreessen Horowitz, June 10, 2016. As of September 18, 2018: http://a16z.com/2016/06/10/ai-deep-learning-machines

Dickson, Ben, "How Artificial Intelligence Is Shaping the Future of Education," *PCMag*, November 20, 2017. As of September 18, 2018: https://www.pcmag.com/article/357483/how-artificial-intelligence-is-shaping-the-future-of-educati

———, "Artificial Intelligence Has a Bias Problem, and It's Our Fault," *PCMag*, June 14, 2018. As of September 18, 2018: https://www.pcmag.com/article/361661/artificial-intelligence-has-a-bias-problem-and-its-our-fau

European Union General Data Protection Regulation, homepage, undated. As of October 30, 2018: https://www.eugdpr.org/

European Consumer Consultative Group, *Policy Recommendations for a Safe and Secure Use of Artificial Intelligence, Automated Decision-Making, Robotics and Connected Devices in a Modern Consumer World*, European Commission, May 16, 2018. As of September 18, 2018: https://ec.europa.eu/info/sites/info/files/eccg-recommendation-on-ai_may2018_en.pdf

Federal Bureau of Investigation, "Education Technologies: Data Collection and Unsecured Systems Could Pose Risks to Students," public service announcement, Alert No. I-091318-PSA, September 13, 2018. As of October 30, 2018: https://www.ic3.gov/media/2018/180913.aspx

Gleason, P., and M. Dynarski, "Do We Know Whom to Serve? Issues in Using Risk Factors to Identify Dropouts," *Journal of Education for Students Placed at Risk*, Vol. 7, No. 1, 2002, pp. 25–41.

Hanover Research, *Early Alert Systems in Higher Education*, Washington, D.C., November 2014. As of September 18, 2018: https://www.hanoverresearch.com/wp-content/uploads/2017/08/Early-Alert-Systems-in-Higher-Education.pdf

IBM Watson Education, homepage, undated. As of October 30, 2018: https://www.ibm.com/watson/education

Jayaprakash, S. M., E. W. Moody, E. J. Lauría, J. R. Regan, and J. D. Baron, "Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative," *Journal of Learning Analytics*, Vol. 1, No. 1, 2014, pp. 6–47.

Klahr, Philip, and Donald Waterman, eds., *Expert Systems: Techniques, Tools and Applications*, Boston, Mass.: Addison-Wesley Longman Publishing, 1986.

Knight, Will, "*Machine Learning and Data Are Fueling a New Kind of Car,*" *MIT Technology Review*, March 14, 2017a. As of December 18, 2018: https://www.technologyreview.com/s/603853/machine-learning-and-data-are-fueling-a-new-kind-of-car

———, "The Dark Secret at the Heart of AI," *MIT Technology Review*, April 11, 2017b. As of September 18, 2018: https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai

Kolowich, Steven, "Writing Instructor, Skeptical of Automated Grading, Pits Machine vs. Machine," *Chronicle of Higher Education*, April 24, 2014. As of October 30, 2018: https://www.chronicle.com/article/Writing-Instructor-Skeptical/146211

Kwok, Sam, "Artificial Intelligence: A Primer," Garage Technology Ventures, blog, undated. As of September 18, 2018: https://www.garage.com/artificial-intelligence

LeCun, Y., Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, Vol. 521, May 28, 2015, pp. 436–444.

Leonard-Barton, Dorothy, and John Sviokla, "Putting Expert Systems to Work," *Harvard Business Review*, March–April 1988. As of September 18, 2018: https://hbr.org/1988/03/putting-expert-systems-to-work

Levy, Frank, "Computers and Populism: Artificial Intelligence, Jobs, and Politics in the Near Term," *Oxford Review of Economic Policy*, Vol. 34, No. 3, 2018, pp. 393–417.

Lohr, Steve, "IBM Is Counting on Its Bet on Watson, and Paying Big Money for It," *New York Times*, October 17, 2016. As of January 3, 2019: https://www.nytimes.com/2016/10/17/technology/ibm-is-counting-on-its-bet-on-watson-and-paying-big-money-for-it.html

Ma, W., O. O. Adesope, J. C. Nesbit, and Q. Liu, "Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis," *Journal of Educational Psychology*, Vol. 106, No. 4, 2014, pp. 901–918.

McFarland, Thomas, and Reese Parker, *Expert Systems in Education and Training*, Englewood, N.J.: Educational Technologies Publications, 1990.

National Science and Technology Council, Committee on Technology, *Preparing for the Future of Artificial Intelligence*, Washington, D.C.: Executive Office of the President, October 2016. As of September 18, 2018: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

Ng, Andrew, "What Artificial Intelligence Can and Can't Do Right Now," *Harvard Business Review*, November 9, 2016. As of September 18, 2018: https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now

Nkambou, Roger, Riichiro Mizoguchi, and Jacqueline Bourdeau, eds., *Advances in Intelligent Tutoring Systems*, New York: Springer, 2010.

O'Neil, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown Publishing Group, 2016.

Osoba, Osonde A., and Paul K. Davis, *An Artificial Intelligence/Machine Learning Perspective on Social Simulation: New Data and Challenges*, Santa Monica, Calif.: RAND Corporation, WR-1213-DARPA, 2018. As of September 18, 2018: https://www.rand.org/pubs/working_papers/WR1213.html

Osoba, Osonde A., and William Welser IV, *An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence*, Santa Monica, Calif.: RAND Corporation, RR-1744-RC, 2017a. As of August 29, 2018: https://www.rand.org/pubs/research_reports/RR1744.html

———, *The Risks of Artificial Intelligence to Security and the Future of Work*, Santa Monica, Calif.: RAND Corporation, PE-237-RC, 2017b. As of August 27, 2018: https://www.rand.org/pubs/perspectives/PE237.html

Policy and Program Studies Service, *Issue Brief: Early Warning Systems*, Washington, D.C.: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, September 2016. As of September 18, 2018: https://www2.ed.gov/rschstat/eval/high-school/early-warning-systems-brief.pdf

Powers, Donald, Jill Burstein, Martin Chodorow, Mary Fowles, and Karen Kukich, *Stumping E-Rater: Challenging the Validity of Automated Essay Scoring*, Princeton, N.J.: Educational Testing Service, Research Report 01-03, March 2001.

Seidel, Robert, and Ok-Choon Park, "An Historical Perspective and a Model for Evaluation of Intelligent Tutoring Systems," *Journal of Educational Computing Research*, Vol. 10, No. 2, 1994, pp. 103–128.

Shermis, Mark D., "State-of-the-Art Automated Essay Scoring: Competition, Results, and Future Directions from a United Stated Demonstration," *Assessing Writing*, Vol. 20, 2014, pp. 53–76.

Singer, Natasha, "A Student-Data Collector Drops Out," *New York Times*, April 26, 2014. As of October 30, 2018:
https://www.nytimes.com/2014/04/27/technology/a-student-data-collector-drops-out.html

———, "Amazon's Facial Recognition Wrongly Identifies 28 Lawmakers, ACLU Says," *New York Times*, July 26, 2018. As of September 18, 2018:
https://www.nytimes.com/2018/07/26/technology/amazon-aclu-facial-recognition-congress.html

Stone, Peter, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, Julia Hirschberg, Shivaram Kalyanakrishnan, Ece Kamar, Sarit Kraus, Kevin Leyton-Brown, David Parkes, William Press, Anna Lee Saxenian, Julie Shah, Milind Tambe, and Astro Teller, *Artificial Intelligence and Life in 2030—One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel*, Stanford, Calif.: Stanford University, September 2016. As of September 18, 2018:
https://ai100.stanford.edu/sites/default/files/ai_100_report_0916fnl_single.pdf

Summerson, Karen, "AI in Pathology—Use Cases in Slide Imaging, Tissue Phenomics, and More," TechEmergence, June 18, 2018. As of September 18, 2018:
https://www.techemergence.com/ai-in-pathology-use-cases-in-slide-imaging-tissue-phenomics-and-more

Thompson, Clive, "How to Teach Artificial Intelligence Some Common Sense," *Wired*, November 13, 2018. As of November 27, 2018:
https://www.wired.com/story/how-to-teach-artificial-intelligence-common-sense

VanLehn, Kurt, "The Behavior of Tutoring Systems," *International Journal of Artificial Intelligence in Education*, Vol. 16, No. 3, 2006, pp. 227–265.

Voosen, Paul, "How AI Detectives Are Cracking Open the Black Box of Deep Learning," *Science*, July 6, 2017. As of September 18, 2018:
http://www.sciencemag.org/news/2017/07/how-ai-detectives-are-cracking-open-black-box-deep-learning

Acknowledgments

The author would like to thank the individuals who helped improve the quality of this paper. The reviewers, Jeremy Roschelle and John S. Davis II, shared insightful comments and constructive feedback. In addition, the author thanks his colleagues, Osonde Osaba, Laura Hamilton, John Pane, and Heather Schwartz for helpful feedback on early drafts of the paper and Allison Kerns for her excellent editing. The author is solely responsible for the content of the paper, and the opinions and recommendations expressed here are those of the author and do not necessarily reflect the views of the reviewers.

About the Author

Robert F. Murphy is a senior education researcher at the RAND Corporation. His education research focuses on the evaluation of innovative educational programs and technologies and the policies and external factors that affect their adoption and efficacy.

About This Perspective

Recent applications of artificial intelligence (AI) have been successful in performing complex tasks in health care, financial markets, manufacturing, and transportation logistics, but the influence of AI applications in the education sphere has been limited. However, that may be changing. In this paper, the author discusses several ways that AI applications can be used to support the work of K–12 teachers by augmenting teacher capacity rather than replacing teachers. Three promising applications are intelligent tutoring systems, automated essay scoring, and early warning systems. The author also discusses some of the key technical challenges that need to be addressed in order to realize the full potential of AI applications for educational purposes. The paper should be of interest to education journalists, publishers, product developers, researchers, and district and school administrators.

This research was undertaken in RAND Education and Labor, a division of the RAND Corporation that conducts research on early childhood through postsecondary education programs, workforce development, and programs and policies affecting workers, entrepreneurship, financial literacy, and decisionmaking.

More information about RAND can be found at www.rand.org. Questions about this Perspective should be directed to robertm@rand.org, and questions about RAND Education and Labor should be directed to educationandlabor@rand.org.

RAND Ventures

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND Ventures is a vehicle for investing in policy solutions. Philanthropic contributions support our ability to take the long view, tackle tough and often-controversial topics, and share our findings in innovative and compelling ways. RAND's research findings and recommendations are based on data and evidence and therefore do not necessarily reflect the policy preferences or interests of its clients, donors, or supporters.

Funding for this venture was provided by gifts from RAND supporters and income from operations.

For more information on this publication, visit www.rand.org/t/PE315.

RAND CORPORATION

www.rand.org

PE-315-RC (2019)