

# Artificial intelligence blends algorithms and applications

January 3 2019, by Kevin McElwee

---



Through advances in artificial intelligence and machine learning, Princeton researchers are building technologies in health care, transportation, energy and many other areas. Credit: Daniel Hertzberg

Artificial intelligence is already a part of everyday life. It helps us answer questions like "Is this email spam?" It identifies friends in online photographs, selects news stories based on our politics and helps us deposit checks via our phones—if all somewhat imperfectly.

But these applications are just the beginning. Through advances in computer science, researchers are creating new capabilities that have the potential to improve our lives in ways we have yet to imagine. Princeton researchers are at the forefront of this research, from the theoretical underpinnings to the new apps and devices to the ethical considerations.

Attempts at building [intelligent systems](#) are as old as computers themselves. Early efforts often involved directly programming rules of behavior into a system. For example, researchers might input the laws of motion to control a robotic arm. But the resulting behaviors usually fell short.

With artificial intelligence, computers learn from experience. Through "machine learning," a subfield of [artificial intelligence](#), computers are programmed to make choices, learn from the outcomes, and adjust to feedback from the environment.

Machine learning is transforming scholarship across campus, said Jennifer Rexford, Princeton's Gordon Y.S. Wu Professor in Engineering and chair of the computer science department.

"Princeton has a very long tradition of strong work in computer science and mathematics, and we have many departments that are just top notch, combined with an emphasis on serving humanity," Rexford said. "You just don't get that everywhere."

## Positive outcomes

One societal challenge that artificially intelligent machines are addressing is how to make better health care decisions. Barbara Engelhardt, an associate professor of [computer science](#), is creating algorithms to help doctors adopt practices most likely to have positive patient outcomes.

For example, when should a patient be weaned from a ventilator? Used by one in three patients in intensive care units, a ventilator is a life-saving device, but is invasive, costly and can spread infection. Doctors often wait longer than necessary to remove a patient from a ventilator, because if they are wrong, they could complicate the patient's health further.

In partnership with researchers at the University of Pennsylvania's hospital system, Engelhardt and her team aim to move patient care away from a one-size-fits-all approach to one that is tailored to individual patients. Their algorithm considers many patient factors and then calculates when and how to remove the patient from the ventilator. It makes numerous decisions, including how much sedative to give prior to the procedure and how to test whether the patient can breathe unassisted.

Machine learning could also help in situations where high-quality human healthcare is not immediately available, such as with patients in palliative care, who could be monitored around the clock as if by a specialist.

## Reinforcement learning

Engelhardt uses a machine-learning approach called reinforcement learning, a departure from the older but still widely used practice of "supervised learning," where programmers provide computers with training sets of data and ask the machines to generalize to new situations. For example, to teach a computer to identify dogs in photos, programmers provide tens of thousands of images, from which the computer develops its own rules to figure out whether new photos contain a dog.

Reinforcement learning, by contrast, is more like the trial-and-error learning that young children use. A toddler who tries to pet the family cat and receives a sharp swipe will learn to stay away from cats. Similarly, the computers try things and interpret the results.

Mengdi Wang, an assistant professor of operations research and financial engineering, studies this approach. She has used reinforcement learning to limit risk in financial portfolios, help a local hospital predict complications in knee replacement surgery, and partner with Microsoft Research to produce story-quality dialogue.

One challenge when implementing reinforcement learning is data overload. Computers don't have the advantage of human forgetfulness, so they must process all incoming data. In practice, experts often have to step in to put some bounds on the number of items that need to be considered.

"Having too many variables is the bottleneck of reinforcement learning," Wang said. "Even if you have all the information in the world, you have a limited amount of processing power."

Wang developed a method for helping computers figure out what is and



what is not important. It's an algorithm that reduces complexity by mathematically compressing a large collection of possible states into a small number of possible clusters. The approach, which she developed with Anru Zhang of the University of Wisconsin-Madison, uses statistics and optimization to group the likely scenarios for each stage of a decision-making process.

## **AI to the rescue**

Although [reinforcement learning](#) is powerful, it offers no guarantees when an algorithm confronts a new environment. For example, an autonomous aerial vehicle (drone) trained to perform search-and-rescue missions in a certain set of environments may fail dramatically when deployed in a new one.

Developing approaches to guarantee drone safety and performance is the goal of Anirudha Majumdar, an assistant professor of mechanical and aerospace engineering. Due to safety and technological limitations, most drones today require a human to control the craft using its cameras and sensors. But steering drones through destroyed buildings, like those in the radioactivity-damaged Fukushima Daiichi power station in Japan, presents challenges.

Autonomous aerial vehicles could aid search-and-rescue efforts in tight spaces where the risk of human error is great. Majumdar is exploring how to apply a set of tools from machine learning known as "generalization theory" to guarantee drone safety in new environments. Roughly speaking, generalization theory provides ways to narrow the difference between performance on the training data and performance on new data.

## **Language learning**

Teaching computers to recognize shapes is one thing, but teaching them to understand everyday language is quite another. To get at the question of how the brain processes language, Princeton researchers scanned the brains of volunteers who watched episodes of the BBC television series "Sherlock" to see what the brain is doing while its owner is taking in new information.

The challenge was how to aggregate results from several brains to identify trends. Each brain is shaped slightly differently, leading to slight differences in their functional magnetic resonance imaging (fMRI) scans. "It is as if you send a thousand tourists to take a photo of the Eiffel Tower. Each photo will be slightly different depending on the camera, the spot where the tourist stood to take the picture, and so forth," said Peter Ramadge, the Gordon Y.S. Wu Professor of Engineering and director of the Center for Statistics and Machine Learning. "You need machine learning to understand what is common to the response of all the subjects," he said.

Ramadge and other computer scientists, including then-undergraduate Kiran Vodrahalli of the Class of 2016, worked with researchers at the Princeton Neuroscience Institute to aggregate brain scans using a method for finding commonalities called a "shared response model." They then mapped brain activity to the dialogue in the episodes using a [natural language](#) processing technique—which extracts meaning from speech—developed by Sanjeev Arora, Princeton's Charles C. Fitzmorris Professor in Computer Science, and his team.

While a typical [speech recognition](#) method needs huge numbers of examples, the new method is capable of drawing meaning from a relatively small collection of words, such as the few hundred found in the script of the TV show. In a paper published in the journal *NeuroImage* in June 2017, the researchers demonstrated that they could determine from looking at the fMRI scans which scene was being

watched with about 72 percent accuracy.

## **Into the black box**

Machine learning has the potential to unlock questions that humans find difficult or impossible to answer, especially ones involving large data sets. For really complex questions, researchers have developed a method called deep learning, inspired by the human brain. This method relies on artificial neural networks, collections of artificial neurons that, like real brain cells, can receive a signal, process it, and produce an output to hand off to the next neuron.

While deep learning has been successful, researchers are still discovering what tasks it is best suited for, said Arora, who recently founded a program in theoretical machine learning at the nearby Institute for Advanced Study. "The field has derived a lot of use from treating deep learning as a black box," he said. "The question is what will we see when we open the black box."

## **Unintended consequences**

In addition to broad ethical questions about the use of AI and the implications of intelligent machines in society, near-term worries about AI systems taking jobs from people are becoming more common. Enter Ed Felten, who is researching policies to curb the unintended consequences of AI.

Felten, the Robert E. Kahn Professor of Computer Science and Public Affairs and director of Princeton's Center for Information Technology Policy, served as deputy U.S. chief technology officer in the Obama White House, where he led federal policy initiatives on AI and [machine learning](#).

With researchers at New York University, Felten has explored whether concerns about AI's impact on jobs and the economy can be supported by data. The researchers used standard benchmarks published by AI researchers. For visual recognition, for example, the team evaluated how many images an AI algorithm correctly categorized. Felten and his colleagues paired this estimation with data sets provided by the Bureau of Labor Statistics.

The question is whether AI will replace workers, or complement their efforts and lead to even greater opportunities? History shows that new technologies often prove beneficial for workers in the long term, but not without short-term pains for workers replaced by technology.

While some researchers think that low-skill jobs will experience the greatest threat from artificially intelligent machines, Felten's numbers suggest otherwise. Airline pilots and lawyers may be at least as threatened by automation as the person behind the counter at the local 7-Eleven, he said.

"Things like house cleaning are very difficult to automate," Felten said. "The person doing that job needs to make a lot of contextual decisions. Which objects on the floor are trash and which objects on the floor are valued objects that have fallen on the floor?"

Felten and his team plan to pair their findings with geographic information, giving a kind of heat map on what regions of the country will be most affected, to allow companies and governments to prepare for the coming changes.

"I'm an optimist in that I think there's huge opportunity," Felten said. "AI is going to lead to tremendous progress in a lot of different areas. But it does come with risks, and we could easily do it badly."



**More information:** Kiran Vodrahalli et al. Mapping between fMRI responses to movies and their natural language annotations, *NeuroImage* (2017). [DOI: 10.1016/j.neuroimage.2017.06.042](https://doi.org/10.1016/j.neuroimage.2017.06.042)

Provided by Princeton University

Citation: Artificial intelligence blends algorithms and applications (2019, January 3) retrieved 4 May 2024 from <https://techxplore.com/news/2019-01-artificial-intelligence-blends-algorithms-applications.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.