I’ve been using machine-learning methods for more than two decades for a variety of problems, the most challenging of which has been to design a set of algorithms that learn from data how to trade securities profitably on their own and with minimal human input. I started this research to answer a basic question, namely, whether machine-learning algorithms could be shaped to make investment decisions comparable to the best humans with access to the same data. I hitched myself to the machine-learning post out of the conviction that financial markets are too complex, subtle, and “noisy” to be described by humans in simple terms, and that they evolve over time along multiple dimensions. By noisy, I mean that seemingly identical market conditions at different times can result in very different outcomes. My thinking was that machine-learning algorithms, by virtue of their ability, in principle, to learn or generalize from large amounts of noisy data, should have advantages over humans when it comes to making decisions more frequently.

The broad approach I took is commonly referred to as “machine learning” (ML). The creative work of crafting useful descriptions about the problem that are fed to the machine is done by humans, based on intuition and prior theory if it exists. The computer then learns to generalize from this type of curated input. For example, I might take a few price series and do some calculations on them that constitute the curated input fed to the machine. If it turns out that a calculation such as “increasing volatility” of the series (which may be indicative of a jumpy market) is generally bearish for stocks, the ML algorithms should be able to infer this by doing some arithmetic on the subsequent outcomes associated with the market under jumpy conditions.

But what if instead we gave the computer just the raw data—prices, news, earnings, and macro numbers—and asked it to learn from them directly, without any human curation of the raw data?

Sound like science fiction?

Perhaps, but before we dismiss it as outlandish, it is worth considering the surge in interest among the Internet giants such as Google and Facebook, which is matched in fervor, if not financial resources, by smaller innovators in the area of “deep learning.” The UK-based company DeepMind was acquired by Google for roughly $400 million last year. What was so significant about this acquisition?

Consider the following experiment on learning to play video games. With access only to the video output on the screen, game controllers, and the game score—the only feedback available for learning—a deep learning-based program learned how to play three games better than humans.

It is hard to underestimate the significance of this feat. The program required no prior knowledge about the problem it would be trying to solve and no preprocessing of inputs by humans. The machine “saw” the same input that a human would see, in the same form—sequences of pixels—and learned to play the games very well.

What is the essential difference between traditional ML and deep learning that enables this “different in kind” capability? And perhaps more tantalizingly, what are the promises and challenges associated with this difference in kind?

A fundamental difference is in how the problem is formulated and how the “decision surface” is generated from the data. Unlike standard ML approaches (such as the ones I’ve used to develop financial prediction systems) in which the raw data are curated into features that direct and constrain the algorithms’ explorations in a specified manner, there is zero to minimal pre-processing of the raw data in the deep-learning context. Raw data on, say, pixels of the screen in a video game or a digitized facial image or raw text are used to learn useful features implicitly. Raw data and nothing else. Features derived from pixels might be lines, shadows, distance between similar
patterns, lighting discontinuities, and so on that are useful in distinguishing objects, but it is the machine’s challenge to figure this out. In the standard ML approach, these types of features are carefully crafted measurements that are fed into the learner in the place of the raw data. But not in the deep-learning paradigm.

Arguably, deep learning is motivated by a closer intended correspondence with how the human brain processes visual and auditory inputs to create useful higher-level features by itself for the problem it is learning to solve. The machine learns how these features relate to each other or to objects or decisions of interest. As humans, we learn in a more fluid and robust manner than computers have been able to thus far for many basic recognition tasks. Many problems that challenge computers even today are trivial for humans. We easily recognize faces in varying lighting conditions, orientations, and more, even when seeing with them for the first time. And we generally use the available data intelligently—ignoring superfluous data when it is plentiful and making the best of it when it is in short supply, being impervious to rotation and scale, focusing on changes rather than the full montage from period to period, and so on—and using elementary mechanisms such as reinforcement and dampening, which organisms as primitive as ants have employed with great success in learning about danger or finding useful supplies.

An even deeper learning challenge than recognition is one that involves the linking of a series of events or actions taken over time and synthesizing them into useful lessons that relate to achieving outcomes. For example, consider a video game where one must learn a concept such as “always take cover from a threat” as a key ingredient in survival and hence achieving a high score. Or “fire immediately at yellow objects” and so on. Some actions might even be costly in the short term but essential to long-term success. For example, coming up for air in an underwater mission where one is racing the clock might be costly at the time it is taken, but essential to eventual success when trying to save trapped survivors. While these tradeoffs may seem like common sense to us as adults, they are difficult learning problems for a machine in that the connection between the “attribution” of actions (such as moving to the right or left, hiding, firing a weapon, etc.) and outcomes involve complex temporal relationships and similarly complex objective functions. This type of ex post high-dimensional “path attribution” is very difficult to learn, yet, humans do it remarkably easily, even if not always optimally.

The most common technology underlying deep-learning algorithms is some variant of a multilayer neural network. Neural nets are more than half a century old, and have been a staple in the ML community for decades. DeepMind’s demonstration that a multilayer neural network could learn to play video games in the same way as humans, using the computer equivalent of hand-to-eye coordination. This was a major breakthrough in deep learning, and stood in stark contrast to the bulk of ML approaches that typically use some representation of the game parameters as input rather than unfiltered images from the screen. While the standard neural network algorithms such as “back propagation” are by now well known, the skills associated with the newer tricks that allow deep-learning tools to work this way are in short supply. Indeed, skeptics of deep learning might reasonably contend that to date we do not completely understand how much engineering effort is needed to get a deep-learning machine to work for a particular problem, or how broadly it will excel. It seems clear that at the moment, with the state of the art, for tasks such as image classification (among others), engineering the system to build its own features performs better than engineering the features manually.

Is deep learning well suited to specific types of problems and more challenging for others? This is certainly true of most of the stable of ML technologies today. And it is true for most humans as well, who are excellent at facial recognition but on average not very good at learning to win at chess, which is a challenging intellectual task, or in making good investment decisions, which is inherently difficult for a number of reasons that I have discussed elsewhere. The answer to this question for deep learning is likely to be driven by the nature of the relationship between the form of the raw input and the output that the deep-learning algorithm is seeking to extract. The more explicit and direct the relationship between the raw data with outcomes of interest, the more likely that a deep learner will be able to learn to associate them. For example, a deep learner for image recognition that is trained on a limited number of images can subsequently recognize objects from multiple orientations and noisy input if it is able to learn the deeper relationships among elements in the raw pixel level data. These relationships can be spatial, temporal, or logical. At the other extreme, prediction problems such as those in financial markets pose serious challenges to learning systems for a number of reasons, including how to even define...
and represent the raw inputs. Should a price series be transformed at all, such as by taking differences, or represented in its original values? When there are multiple series involved, can they be represented in a way that exposes the relationship between the raw inputs and the variable being predicted? These are challenging questions that are the subject of current research.

Current industrial applications are pursuing the basic machinery of deep learning tools to solve specific problems of interest. These areas have copious amounts of data from humans and machines feeding into learning methods that are capable of associating patterns in the raw data with specific outcomes. It may be some time, however, before we see machines rise to a level at which they learn to handle deeply intellectual or noisy problems without the kind of curated assistance we provide to existing machine learning based systems. One thing seems certain. The level of intelligence we are going to see in machines performing motor, translation, and recognition types of tasks will accelerate in the emerging era of big data and powerful automated learning methods.

References
2. Regalado A. Is Google cornering the market on deep learning: a cutting-edge corner of science is being wooed by Silicon Valley, to the dismay of some academics. Available at www.technologyreview.com/news/524026/is-google-cornering-the-market-on-deep-learning/


Abbreviation Used
ML = machine learning