The future of deep learning, according to its pioneers

By

[Ben Dickson](#)

July 1, 2021

6 min read

Deep learning pioneers Yoshua Bengio (left), Geoffrey Hinton (center), and Yann LeCun (right)

This article is part of our [reviews of AI research papers](#), a series of posts that explore the latest findings in artificial intelligence.

Deep neural networks will move past their shortcomings without help from [symbolic artificial intelligence](#), three pioneers of deep learning argue in a paper published in the July issue of the *Communications of the ACM* journal.

In their paper, Yoshua Bengio, Geoffrey Hinton, and Yann LeCun, recipients of the 2018 Turing Award, explain the current [challenges of deep learning](#) and how it differs from learning in humans and animals. They also explore recent advances in the field that might provide blueprints for the future directions for research in deep learning.

Titled "Deep Learning for AI," the paper envisions a future in which deep learning models can learn with little or no help from humans, are flexible to changes in their environment, and can solve a wide range of reflexive and cognitive problems.

The challenges of deep learning

Image credit: Depositphotos

Deep learning is often compared to the brains of humans and animals. However, the past years have proven that artificial neural networks, the main component used in deep learning models, lack the efficiency, flexibility, and versatility of their biological counterparts.

In their paper, Bengio, Hinton, and LeCun acknowledge these shortcomings. "Supervised learning, while successful in a wide variety of tasks, typically requires a large amount of human-labeled data. Similarly, when reinforcement learning is based only on rewards, it requires a very large number of interactions," they write.

Supervised learning is a popular subset of machine learning algorithms, in which a model is presented with labeled examples, such as a list of images and their corresponding content. The model is trained to find recurring patterns in examples that have similar labels. It then uses the learned patterns to associate new examples with the right labels. Supervised learning is especially useful for problems where labeled examples are abundantly available.

Reinforcement learning is another branch of machine learning, in which an "agent" learns to maximize "rewards" in an environment. An environment can be as simple as a tic-tac-toe board in which an AI player is rewarded for lining up three Xs or Os, or as complex as an urban setting

in which a self-driving car is rewarded for avoiding collisions, obeying traffic rules, and reaching its destination. The agent starts by taking random actions. As it receives feedback from its environment, it finds sequences of actions that provide better rewards.

In both cases, as the scientists acknowledge, machine learning models require huge labor. Labeled datasets are hard to come by, especially in specialized fields that don't have public, open-source datasets, which means they need the hard and expensive labor of human annotators. And complicated reinforcement learning models require massive computational resources to run a vast number of training episodes, which makes them available to a few, very wealthy AI labs and tech companies.

Bengio, Hinton, and LeCun also acknowledge that current deep learning systems are still limited in the scope of problems they can solve. They perform well on specialized tasks but "are often brittle outside of the narrow domain they have been trained on." Often, slight changes such as a few modified pixels in an image or a very slight alteration of rules in the environment can cause deep learning systems to go astray.

The brittleness of deep learning systems is largely due to machine learning models being based on the "independent and identically distributed" (i.i.d.) assumption, which supposes that real-world data has the same distribution as the training data. i.i.d also assumes that observations do not affect each other (e.g., coin or die tosses are independent of each other).

"From the early days, theoreticians of machine learning have focused on the iid assumption… Unfortunately, this is not a realistic assumption in the real world," the scientists write.

Real-world settings are constantly changing due to different factors, many of which are virtually impossible to represent without causal models. Intelligent agents must constantly observe and learn from their environment and other agents, and they must adapt their behavior to changes.

"[T]he performance of today's best AI systems tends to take a hit when they go from the lab to the field," the scientists write.

The i.i.d. assumption becomes even more fragile when applied to fields such as computer vision and natural language processing, where the agent must deal with high-entropy environments. Currently, many researchers and companies try to overcome the limits of deep learning by training neural networks on more data, hoping that larger datasets will cover a wider distribution and reduce the chances of failure in the real world.

Deep learning vs hybrid AI
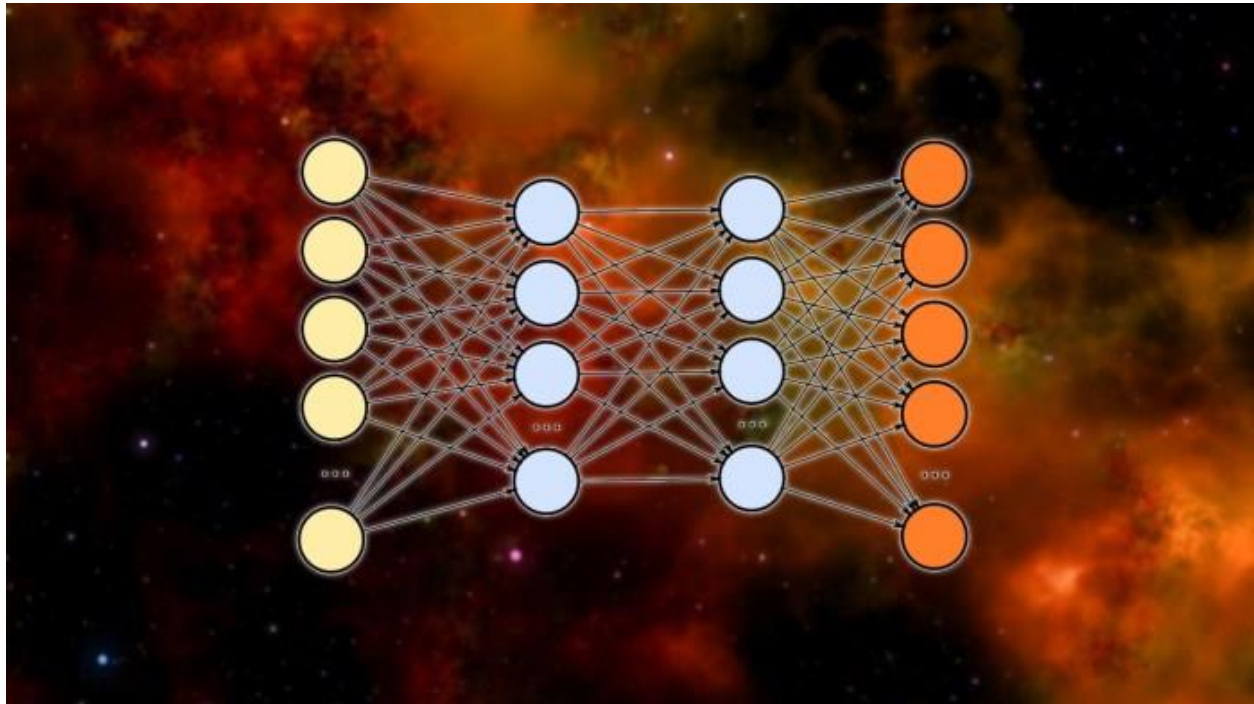
The ultimate goal of AI scientists is to replicate the kind of general intelligence humans have. And we know that humans don't suffer from the problems of current deep learning systems.

"Humans and animals seem to be able to learn massive amounts of background knowledge about the world, largely by observation, in a task-independent manner," Bengio, Hinton, and LeCun write in their paper. "This knowledge underpins common sense and allows humans to learn complex tasks, such as driving, with just a few hours of practice."

Elsewhere in the paper, the scientists note, "[H]umans can generalize in a way that is different and more powerful than ordinary iid generalization: we can correctly interpret novel combinations of existing concepts, even if those combinations are extremely unlikely under our training distribution, so long as they respect high-level syntactic and semantic patterns we have already learned."

Scientists provide various solutions to close the gap between AI and human intelligence. One approach that has been widely discussed in the past few years is hybrid artificial intelligence that combines neural networks with classical symbolic systems. Symbol manipulation is a very important part of humans' ability to reason about the world. It is also one of the great challenges of deep learning systems.

Bengio, Hinton, and LeCun do not believe in mixing neural networks and symbolic AI. In a video that accompanies the ACM paper, Bengio says, "There are some who believe that there are problems that neural networks just cannot resolve and that we have to resort to the classical AI, symbolic approach. But our work suggests otherwise."

The deep learning pioneers believe that better neural network architectures will eventually lead to all aspects of human and animal intelligence, including symbol manipulation, reasoning, causal inference, and common sense.

Promising advances in deep learning

Video link: https://vimeo.com/554817366

In their paper, Bengio, Hinton, and LeCun highlight recent advances in deep learning that have helped make progress in some of the fields where deep learning struggles. One example is the Transformer, a neural network architecture that has been at the heart of language models such as OpenAI's GPT-3 and Google's Meena. One of the benefits of Transformers is their capability to learn without the need for labeled data. Transformers can develop representations through unsupervised learning, and then they can apply those representations to fill in the blanks on incomplete sentences or generate coherent text after receiving a prompt.

More recently, researchers have shown that Transformers can be applied to computer vision tasks as well. When combined with convolutional neural networks, transformers can predict the content of masked regions.

A more promising technique is contrastive learning, which tries to find vector representations of missing regions instead of predicting exact pixel values. This is an intriguing approach and seems to be much closer to what the human mind does. When we see an image such as the one below, we might not be able to visualize a photo-realistic depiction of the missing parts, but our mind can come up with a high-level representation of what might go in those masked regions (e.g., doors, windows, etc.). (My own observation: This can tie in well with other research in the field aiming to align vector representations in neural networks with real-world concepts.)

The push for making neural networks less reliant on human-labeled data fits in the discussion of self-supervised learning, a concept that LeCun is working on.

Can you guess what is behind the grey boxes in the above image?

The paper also touches upon "system 2 deep learning," a term borrowed from Nobel laureate psychologist Daniel Kahneman. System 2 accounts for the functions of the brain that require conscious thinking, which include symbol manipulation, reasoning, multi-step planning, and solving complex mathematical problems. System 2 deep learning is still in its early stages, but if it becomes a reality, it can solve some of the key problems of neural networks, including out-of-distribution generalization, causal inference, robust transfer learning, and symbol manipulation.

The scientists also support work on "Neural networks that assign intrinsic frames of reference to objects and their parts and recognize objects by using the geometric relationships." This is a reference to "capsule networks," an area of research Hinton has focused on in the past few years. Capsule networks aim to upgrade neural networks from detecting features in images to detecting objects, their physical properties, and their hierarchical relations with each other. Capsule networks can provide deep learning with "intuitive physics," a capability that allows humans and animals to understand three-dimensional environments.

"There's still a long way to go in terms of our understanding of how to make neural networks really effective. And we expect there to be radically new ideas," Hinton told ACM.

Reference: https://bdtechtalks.com/2021/07/01/deep-learning-future-bengio-hinton-lecun/