

Smarter training of neural networks

7 May 2019, by Adam Conner-Simons



(L-R) MIT Assistant Professor Michael Carbin and PhD student Jonathan Frankle. Credit: Jason Dorfman/MIT CSAIL

These days, nearly all the artificial intelligence-based products in our lives rely on "deep neural networks" that automatically learn to process labeled data.

For most organizations and individuals, though, deep learning is tough to break into. To learn well, neural networks normally have to be quite large and need massive datasets. This training process usually requires multiple days of training and expensive graphics processing units (GPUs)—and sometimes even custom-designed hardware.

But what if they don't actually have to be all that big, after all?

In a new paper, researchers from MIT's Computer Science and Artificial Intelligence Lab (CSAIL) have shown that neural networks contain subnetworks that are up to one-tenth the size yet capable of being trained to make equally accurate predictions—and sometimes can learn to do so even faster than the originals.

The team's approach isn't particularly efficient

now—they must train and "prune" the full [network](#) several times before finding the successful subnetwork. However, MIT Assistant Professor Michael Carbin says that his team's findings suggest that, if we can determine precisely which part of the original network is relevant to the final prediction, scientists might one day be able to skip this expensive process altogether. Such a revelation has the potential to save hours of work and make it easier for meaningful models to be created by individual programmers, and not just huge tech companies.

"If the initial network didn't have to be that big in the first place, why can't you just create one that's the right size at the beginning?" says Ph.D. student Jonathan Frankle, who presented his new paper co-authored with Carbin at the International Conference on Learning Representations (ICLR) in New Orleans. The project was named one of ICLR's two best papers, out of roughly 1,600 submissions.

The team likens traditional [deep learning](#) methods to a lottery. Training large neural networks is kind of like trying to guarantee you will win the lottery by blindly buying every possible ticket. But what if we could select the winning numbers at the very start?

"With a traditional neural network you randomly initialize this large structure, and after training it on a huge amount of data it magically works," Carbin says. "This large structure is like buying a big bag of tickets, even though there's only a small number of tickets that will actually make you rich. The remaining science is to figure how to identify the winning tickets without seeing the winning numbers first."

The team's work may also have implications for so-called "transfer learning," where networks trained for a task like image recognition are built upon to then help with a completely different task.

Traditional transfer learning involves training a network and then adding one more layer on top

that's trained for another task. In many cases, a network trained for one purpose is able to then extract some sort of general knowledge that can later be used for another purpose.

For as much hype as [neural networks](#) have received, not much is often made of how hard it is to train them. Because they can be prohibitively expensive to train, data scientists have to make many concessions, weighing a series of trade-offs with respect to the size of the model, the amount of time it takes to train, and its final performance.

To test their so-called "lottery ticket hypothesis" and demonstrate the existence of these smaller subnetworks, the team needed a way to find them. They began by using a common approach for eliminating unnecessary connections from trained networks to make them fit on low-power devices like smartphones: They "pruned" connections with the lowest "weights" (how much the network prioritizes that connection).

Their key innovation was the idea that connections that were pruned after the network was trained might never have been necessary at all. To test this hypothesis, they tried training the exact same network again, but without the pruned connections. Importantly, they "reset" each connection to the weight it was assigned at the beginning of training. These initial weights are vital for helping a lottery ticket win: Without them, the pruned networks wouldn't learn. By pruning more and more connections, they determined how much could be removed without harming the network's ability to learn.

To validate this hypothesis, they repeated this process tens of thousands of times on many different networks in a wide range of conditions.

"It was surprising to see that resetting a well-performing network would often result in something better," says Carbin. "This suggests that whatever we were doing the first time around wasn't exactly optimal, and that there's room for improving how these models learn to improve themselves."

As a next step, the team plans to explore why certain subnetworks are particularly adept at

learning, and ways to efficiently find these subnetworks.

"Understanding the 'lottery ticket hypothesis' is likely to keep researchers busy for years to come," says Daniel Roy, an assistant professor of statistics at the University of Toronto, who was not involved in the paper. "The work may also have applications to network compression and optimization. Can we identify this subnetwork early in training, thus speeding up [training](#)? Whether these techniques can be used to build effective compression schemes deserves study."

More information: The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks. openreview.net/forum?id=rJl-b3RcF7

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

APA citation: Smarter training of neural networks (2019, May 7) retrieved 28 June 2022 from <https://techxplore.com/news/2019-05-smarter-neural-networks.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.