

Scientists teach machines to learn like humans

December 10 2015





This paper compares human and machine learning for a wide range of simple



visual concepts, or handwritten characters selected from alphabets around the world. This is an artist's interpretation of that theme. Art by Danqing Wang. Credit: Danqing Wang

A team of scientists has developed an algorithm that captures our learning abilities, enabling computers to recognize and draw simple visual concepts that are mostly indistinguishable from those created by humans. The work, which appears in the latest issue of the journal *Science*, marks a significant advance in the field—one that dramatically shortens the time it takes computers to 'learn' new concepts and broadens their application to more creative tasks.

"Our results show that by reverse engineering how people think about a problem, we can develop better algorithms," explains Brenden Lake, a Moore-Sloan Data Science Fellow at New York University and the paper's lead author. "Moreover, this work points to promising methods to narrow the gap for other machine learning tasks."

The paper's other authors were Ruslan Salakhutdinov, an assistant professor of Computer Science at the University of Toronto, and Joshua Tenenbaum, a professor at MIT in the Department of Brain and Cognitive Sciences and the Center for Brains, Minds and Machines.

When humans are exposed to a new concept—such as new piece of kitchen equipment, a new dance move, or a new letter in an unfamiliar alphabet—they often need only a few examples to understand its makeup and recognize new instances. While machines can now replicate some pattern-recognition tasks previously done only by humans—ATMs reading the numbers written on a check, for instance—machines typically need to be given hundreds or thousands of examples to perform with similar accuracy.



"It has been very difficult to build machines that require as little data as humans when learning a new concept," observes Salakhutdinov. "Replicating these abilities is an exciting area of research connecting machine learning, statistics, computer vision, and cognitive science."

Salakhutdinov helped to launch recent interest in learning with 'deep neural networks,' in a paper published in *Science* almost 10 years ago with his doctoral advisor Geoffrey Hinton. Their algorithm learned the structure of 10 handwritten character concepts—the digits 0-9—from 6,000 examples each, or a total of 60,000 training examples.

In the work appearing in *Science* this week, the researchers sought to shorten the learning process and make it more akin to the way humans acquire and apply new knowledge—i.e., learning from a small number of examples and performing a range of tasks, such as generating new examples of a concept or generating whole new concepts.

To do so, they developed a 'Bayesian Program Learning' (BPL) framework, where concepts are represented as simple computer programs. For instance, the letter 'A' is represented by computer code—resembling the work of a computer programmer—that generates examples of that letter when the code is run. Yet no programmer is required during the <u>learning process</u>: the algorithm programs itself by constructing code to produce the letter it sees. Also, unlike standard computer programs that produce the same output every time they run, these probabilistic programs produce different outputs at each execution. This allows them to capture the way instances of a concept vary, such as the differences between how two people draw the letter 'A.'





Can you tell the difference between humans and machines? Humans and machines were given an image of a novel character (top) and asked to produce new exemplars. The nine-character grids in each pair that were generated by a machine are (by row) B, A; A, B; A, B. Credit: Brenden Lake



While standard pattern recognition algorithms represent concepts as configurations of pixels or collections of features, the BPL approach learns "generative models" of processes in the world, making learning a matter of 'model building' or 'explaining' the data provided to the algorithm. In the case of writing and recognizing letters, BPL is designed to capture both the causal and compositional properties of real-world processes, allowing the algorithm to use data more efficiently. The model also "learns to learn" by using knowledge from previous concepts to speed learning on new concepts—e.g., using knowledge of the Latin alphabet to learn letters in the Greek alphabet. The authors applied their model to over 1,600 types of handwritten characters in 50 of the world's writing systems, including Sanskrit, Tibetan, Gujarati, Glagolitic—and even invented characters such as those from the television series Futurama.

In addition to testing the algorithm's ability to recognize new instances of a concept, the authors asked both humans and computers to reproduce a series of handwritten characters after being shown a single example of each character, or in some cases, to create new characters in the style of those it had been shown. The scientists then compared the outputs from both humans and machines through 'visual Turing tests.' Here, human judges were given paired examples of both the human and machine output, along with the original prompt, and asked to identify which of the symbols were produced by the computer.





Can you tell the difference between humans and machines? Humans and machines were given an image of a novel character (top) and asked to produce new exemplars. The nine-character grids in each pair that were generated by a machine are (by row) 1, 2; 2, 1; 1, 1. Credit: Brenden Lake



While judges' correct responses varied across characters, for each visual Turing test, fewer than 25 percent of judges performed significantly better than chance in assessing whether a machine or a human produced a given set of symbols.

"Before they get to kindergarten, children learn to recognize new concepts from just a single example, and can even imagine new examples they haven't seen," notes Tenenbaum. "I've wanted to build models of these remarkable abilities since my own doctoral work in the late nineties. We are still far from building machines as smart as a human child, but this is the first time we have had a machine able to learn and use a large class of real-world concepts—even simple visual concepts such as handwritten characters—in ways that are hard to tell apart from humans."

More information: "Human-level concept learning through probabilistic program induction," by B.M. Lake et al. *Science*, <u>www.sciencemag.org/lookup/doi/ ... 1126/science.aab3050</u>

Provided by New York University

Citation: Scientists teach machines to learn like humans (2015, December 10) retrieved 27 April 2024 from <u>https://phys.org/news/2015-12-scientists-machines-humans.html</u>

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.