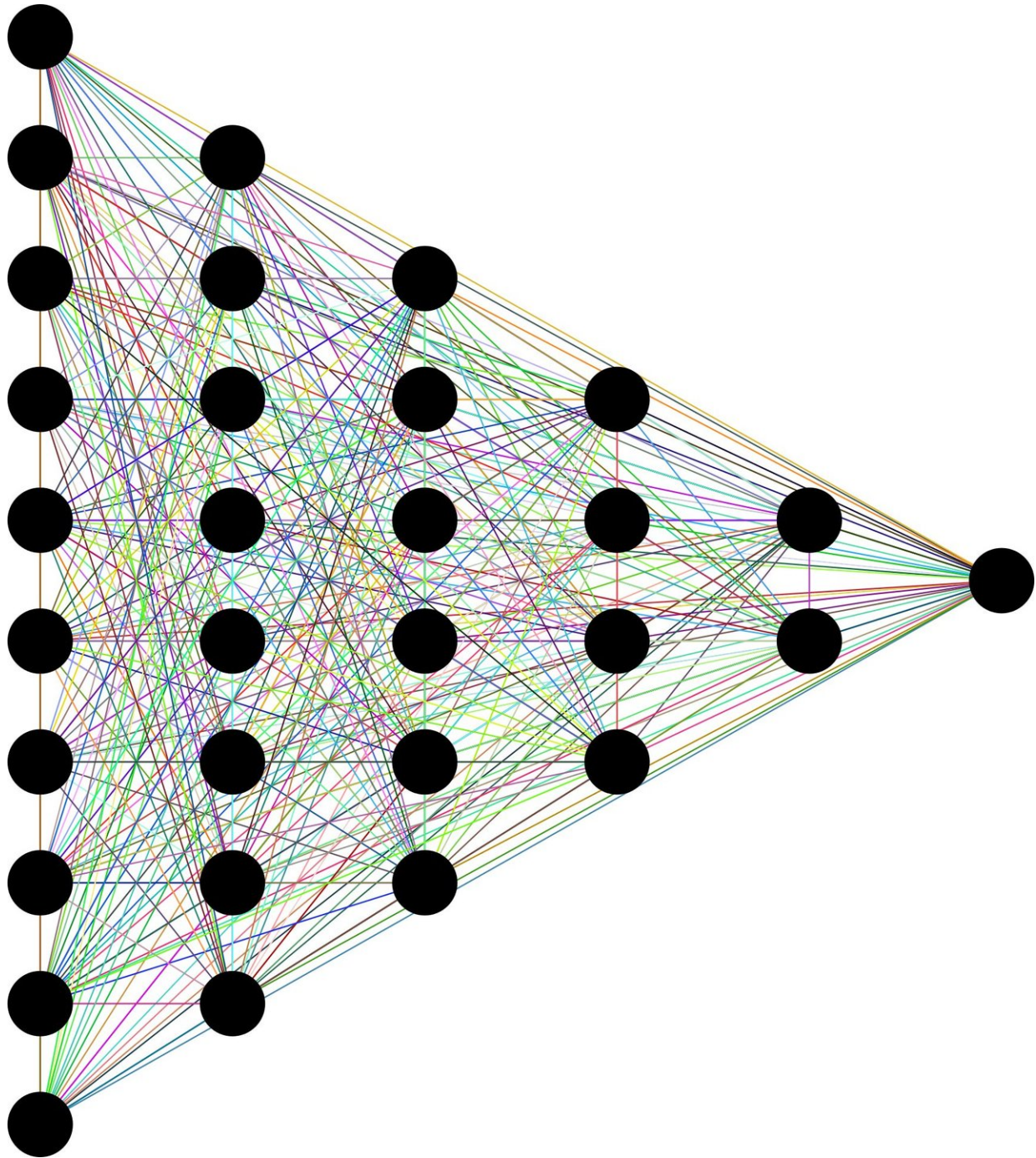


Predicting the accuracy of a neural network prior to training

December 14 2018, by Roxana Istrate



Credit: CC0 Public Domain

Constructing a neural network model for each new dataset is the ultimate

nightmare for every data scientist. What if you could forecast the accuracy of the neural network earlier thanks to accumulated experience and approximation? This was the goal of a recent project at IBM Research and the result is [TAPAS](#) or Train-less Accuracy Predictor for Architecture Search ([click for demo](#)). Its trick is that it can estimate, in fractions of a second, classification performance for unseen input datasets, without training for both image and text classification.

In contrast to previously proposed approaches, TAPAS is not only calibrated on the topological network information, but also on the characterization of the [dataset](#) difficulty, which allows us to re-tune the prediction without any training.

This task was particularly challenging due to the heterogeneity of the datasets used for training neural networks. They can have completely different classes, structures, and sizes, adding to the complexity of coming up with an approximation. When my colleagues and I thought about how to address this, we tried not to think of this as a problem for a computer, but instead to think about how a human would predict the accuracy.

We understood that if you asked a human with some knowledge of deep learning whether a network would be good or bad, that person would naturally have an intuition about it. For example, we would recognize that two types of layers don't mix, or that after one type of layer, there is always another one which follows and improves the accuracy. So we considered whether adding features resembling this human intuitions into a computer could help it do an even better job. And we were correct.

We tested TAPAS on two datasets performed in 400 seconds on a single GPU, and our best discovered networks reached 93.67% accuracy for CIFAR-10 and 81.01% for CIFAR-100, verified by training. These

networks perform competitively with other automatically discovered state-of-the-art networks, but needed only a small fraction of the time to solution and computational resources. Our predictor achieves a performance which exceeds 100 networks per second on a single GPU, thus creating the opportunity to perform large-scale architecture search within a few minutes. We believe this is the first tool which can do predictions based on unseen data.

TAPAS is one of the AI engines in IBM's new breakthrough capability called [NeuNetS](#) as part of [AI OpenScale](#), which can synthesize custom [neural networks](#) in both text and image domains.

In NeuNetS, users will upload their data to the IBM Cloud and then TAPAS can analyze the data and rate it on a scale of 0-1 in terms of complexity of task, 0 meaning hard and 1 being simple. Next TAPAS starts to gather knowledge from its reference library looking for similar datasets based on what the user uploaded. Then based on this, TAPAS can accurately predict how a new [network](#) will perform on the new dataset, very similar to how a human would determine it.

Today's demand for data science skills already exceeds the current supply, becoming a real barrier towards adoption of AI in industry and society. TAPAS is a fundamental milestone towards the demolition of this wall. IBM and the Zurich Research Laboratory are working to make AI technologies as easy to use, as a few clicks on a mouse. This will allow non-expert users to build and deploy AI models in a fraction of the time it takes today—and without sacrificing [accuracy](#). Moreover, these tools will gradually learn over utilization in specialized domains and automatically improve over time, getting better and better.

More information: TAPAS: Train-less Accuracy Predictor for Architecture Search. arxiv.org/abs/1806.00250

This story is republished courtesy of IBM Research. Read the original story [here](#).

Provided by IBM

Citation: Predicting the accuracy of a neural network prior to training (2018, December 14)
retrieved 20 May 2024 from <https://phys.org/news/2018-12-accuracy-neural-network-prior.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.