Computational Neuroscience Seminar

"Dark Knowledge"

Geoffrey Hinton

Computer Science Department, University of Toronto, Canada
Distinguished Researcher, Google Inc.

Wednesday, October 8, 2014

4:00-5:00pm
Auditorium, San Diego Supercomputer Center
University of California, San Diego

Abstract: A simple way to improve classification performance is to average the predictions of a large ensemble of different classifiers. This is great for winning competitions but requires too much computation at test time for practical applications such as speech recognition. In a widely ignored paper in 2006, Caruana and his collaborators showed that the knowledge in the ensemble could be transferred to a single, efficient model by training the single model to mimic the log probabilities of the ensemble average. This technique works because most of the knowledge in the learned ensemble is in the relative probabilities of extremely improbable wrong answers. For example, the ensemble may give an image of a BMW a probability of one in a billion of being a garbage truck but this is still far greater (in the log domain) than its probability of being a carrot. This "dark knowledge", which is practically invisible in the class probabilities, defines a similarity metric over the classes that makes it much easier to learn a good classifier.

I will describe a new variation of this technique called "distillation" and will show some surprising examples in which good classifiers over all of the classes can be learned from data in which some of the classes are entirely absent, provided the target probabilities come from an ensemble that has been trained on all of the classes. I will also show how this technique can be used to improve a state-of-the-art acoustic model and will discuss its application to learning large sets of specialist models without overfitting. This is joint work with Oriol Vinyals and Jeff Dean.

Organized by:
Institute for Neural Computation: http://inc.ucsd.edu
University of California San Diego

Sponsored by:
Qualcomm: http://www.qualcomm.com
Brain Corporation: http://www.braincorporation.com