Max Planck Lecture 2020

Untethering Unsupervised Learning

Geoffrey Hinton

Google Research & Vector Institute & University of Toronto

There are currently two main approaches to unsupervised learning. In the first approach, exemplified by BERT and Variational Autoencoders, a deep neural network is used to reconstruct its input. This is problematic for images because the deepest layers of the network need to encode the detailed phase and noise in the image. An alternative approach, introduced by Becker and Hinton in 1992, is to train two copies of a deep neural network to produce output vectors that have high mutual information when given two different crops of the same image as their inputs. This approach was designed to allow the representations to be "untethered" from irrelevant details of the input. The method of optimizing mutual information used by Becker and Hinton was flawed (for a subtle reason that I will explain) so Pacannaro and Hinton (2001) replaced it by a discriminative objective in which one vector representation must select a corresponding vector representation from among many alternatives. This contrastive objective function gave rise to t-SNE (2008). With faster hardware, contrastive learning of representations has recently become very popular and is proving to be very effective, but it suffers from a major flaw: To learn pairs of representation vectors that have N bits of mutual information we need to contrast the correct corresponding vector with about $2^N$ incorrect alternatives. I will describe a novel and effective way of dealing with this limitation. I will also show that this leads to a simple way of implementing perceptual learning in cortex and suggests an explanation for one of the most peculiar properties of the brain.

Geoffrey Hinton received his PhD in Artificial Intelligence from Edinburgh in 1978. After five years as a faculty member at Carnegie-Mellon he became a fellow of the Canadian Institute for Advanced Research and moved to the Department of Computer Science at the University of Toronto where he is now an Emeritus Distinguished Professor. He is also a Vice President & Engineering Fellow at Google and Chief Scientific Adviser of the Vector Institute.

Geoffrey Hinton was one of the researchers who introduced the backpropagation algorithm and the first to use backpropagation for learning word embeddings. His other contributions to neural network research include Boltzmann machines, distributed representations, time-delay neural nets, mixtures of experts, variational learning and deep learning. His research group in Toronto made major breakthroughs in deep learning that revolutionized speech recognition and object classification.

Geoffrey Hinton is a fellow of the UK Royal Society, a foreign member of the US National Academy of Engineering and a foreign member of the American Academy of Arts and Sciences. His awards include the ACM A. M. Turing Award (jointly with Yoshua Bengio and Yann LeCun), the David E. Rumelhart prize, the IJCAI award for research excellence, the Killam prize for Engineering, the IEEE Frank Rosenblatt medal, the IEEE James Clerk Maxwell Gold medal, the NEC C&C award, the BBVA Foundation Frontiers of Knowledge award, and the NSERC Herzberg Gold Medal.