



De: Centre de recherches mathématiques crm@crm.umontreal.ca
Objet: Conférence de Yoshua Bengio - lundi 16 avril - Programme du 50e anniversaire du CRM - Les mathématique de l'apprentissage machine
Date: 10 avril 2018 15:18
À: Activités CRM activites@crm.umontreal.ca

Programme du 50e anniversaire du CRM
Les mathématique de l'apprentissage machine

14 avril - 11 mai 2018 / April 14 - May 11, 2018

50th anniversary program
Mathematics of machine learning

Conférence inaugurale / Opening keynote lecture

lundi 16 avril / Monday, April 16
11:30 - 12:30
Université de Montréal, Pavillon André-Aisenstadt, salle / room 1360

Yoshua Bengio (Université de Montréal)

"Deep Learning for AI"

There has been rather impressive progress recently with brain-inspired statistical learning algorithms based on the idea of learning multiple levels of representation, also known as neural networks or deep learning. They shine in artificial intelligence tasks involving perception and generation of sensory data like images or sounds and to some extent in understanding and generating natural language. We have proposed new generative models which lead to training frameworks very different from the traditional maximum likelihood framework, and borrowing from game theory. Theoretical understanding of the success of deep learning is work in progress but relies on representation aspects as well as optimization aspects, which interact. At the heart is the ability of these learning mechanisms to capitalize on the compositional nature of the underlying data distributions, meaning that some functions can be represented exponentially more efficiently with deep distributed networks compared to approaches like standard non-parametric methods which lack both depth and distributed representations. On the optimization side, we now have evidence that local minima (due to the highly non-convex nature of the training objective) may not be as much of a problem as thought a few years ago, and that training with variants of stochastic gradient descent actually helps to quickly find better-generalizing solutions. Finally, new interesting questions and answers are arising regarding learning theory for deep networks, why even very large networks do not necessarily overfit and how the representation-forming structure of these networks may give rise to better error bounds which do not absolutely depend on the iid data hypothesis.

<http://www.crm.math.ca/crm50/>
