University of Chicago and Toyota Technological Institute at Chicago
Machine Learning Seminar Series

PRESENTS:

Sebastian Stitch
EPFL, Switzerland

Title: Error Feedback for Communication Efficient SGD

Abstract: Huge scale machine learning problems are nowadays tackled by distributed optimization algorithms, i.e. algorithms that leverage the compute power of many devices for training. The communication overhead is a key bottleneck that hinders perfect scalability. Various recent works proposed to use quantization or sparsification techniques to reduce the amount of data that needs to be communicated. We analyze Stochastic Gradient Descent (SGD) with k-sparsification (for instance top-k or random-k) and compression (for instance quantization) and show that these schemes converge at the same rate as vanilla SGD when equipped with error compensation (i.e. keeping track of accumulated errors in memory). That is, communication can be reduced by a factor of the dimension of the problem (sometimes even more) whilst still converging at the same rate.

Bio: Dr. Sebastian Stich is a postdoctoral researcher at EPFL in Switzerland, working at the Machine Learning and Optimization Laboratory of Prof. Martin Jaggi. He received a MSc in Mathematics with distinction from ETH Zurich in 2010 and a PhD in Theoretical Computer Science from ETH Zurich in 2014. Before joining EPFL, he held for two years a postdoctoral position at UCLouvain to work with Prof. Yurii Nesterov and Prof. François Glineur on coordinate descent methods for large scale optimization problems. Dr. Stich is broadly interested in the complexity analysis of the optimization algorithms that are used in nowadays machine learning applications, with recent focus on distributed algorithms that allow to tackle high dimensional problems. Website: www.sstich.ch

Host: Nati Srebro

Wednesday, March 13 ~ 1 – 2 pm
Harper Center (Booth) Room 219
Pizza provided by UChicago CS Department