מכון ויצמן למדע

THE WEIZMANN INSTITUTE OF SCIENCE

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Machine Learning and Statistics Seminar

Room 1 ,Ziskind Building
on Wednesday, Nov 14, 2018
at 11:15

Kfir Levy
ETH

Beyond SGD: Data Adaptive Methods for Machine Learning

Abstract:

The tremendous success of the Machine Learning paradigm heavily relies on the development of powerful optimization methods. The canonical algorithm for training learning models is SGD (Stochastic Gradient Descent), yet this method has its limitations. It is often unable to exploit useful statistical/geometric structure, it might degrade upon encountering prevalent non-convex phenomena, and it is hard to parallelize. In this talk I will discuss an ongoing line of research where we develop alternative methods that resolve some of SGD"s limitations. The methods that I describe are as efficient as SGD, and implicitly adapt to the underlying structure of the problem in a data dependent manner.

In the first part of the talk, I will discuss a method that is able to take advantage of hard/easy training samples. In the second part, I will discuss a method that enables an efficient parallelization of SGD. Finally, I will briefly describe a method that implicitly adapts to the smoothness and noise properties of the learning objective.