Description of the Lecture

Unique choice box:

- introductory course;
- advanced course;
- exercise sessions;
- programming sessions;
- other interactive sessions

Title:

- Optimization for Machine Learning

Short description/abstract:

- The advent of deep neural networks has triggered a number of important issues in machine learning, which have a direct impact on the associated optimization problems and methods. This course aims at presenting recent developments in this area, with a focus on techniques with theoretical underpinnings that apply in a machine learning setting. In this lecture, we will focus on first-order methods, that form the backbone of modern optimization techniques in data science. The first part of the course will review gradient descent techniques, by showcasing recent advances in applying those methods to convex and nonconvex problems. The basic gradient descent algorithm as well as several extensions (proximal, accelerated, manifold,…) will be presented, along with typical problems of interest. In the second part of the course, we will dive into stochastic gradient methods and their relevance to data science problems. Building on the first part, we will investigate the main features of stochastic gradient methods, such as inclusion of momentum and mini-batching. We will also study popular variants that have been widely used in deep learning applications. Finally, if time permits, we will explore the development of distributed algorithms in the context of massive amounts of data, and the algorithmic challenges posed by this setup.

Sessions: 6 (1 hour per session > Total : 6 Hours).

o Domain from Arxiv (e.g., Algebraic Geometry): Optimization and Control, Machine Learning

o MSC (ex. 35K57 & 35R30): 65K05, 90C06, 90C60

o Keywords (separated by #): Stochastic gradient methods,