Mathematics of Machine Learning

Aurélien Garivier and Sébastien Gerchinovitz

This course is an introduction to the mathematics of statistical machine learning. The main mathematical concepts and tools involved in the study of classical and more recent learning algorithms are presented. Two types of scenarios are addressed: supervised classification (whose applications range from pattern or image recognition to automatic medical diagnosis) and online learning, where data is observed and processed sequentially.

In addition to various statistical models, these lectures will present probabilistic and information-theoretic tools of primary importance in statistical research: non-asymptotic deviation inequalities (in connection with the course on concentration inequalities), which may be used to obtain mathematical guarantees on the performance of learning methods, and lower-bound techniques (Fano's and Assouad's lemmas, Pinsker's inequality) that are necessary to investigate the optimality of learning methods.

Outline of the lectures:

1. Supervised classification: mathematical model
2. Deviation inequalities and Vapnik's theory
3. Optimality in learning theory
4. Online learning: mathematical framework and algorithms
5. Online convex optimization, aggregation of expert advice and stochastic gradient descent
6. Boosting, and connections with game theory

Some references:

1. N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games,* Cambridge University Press (2006).

2. Luc Devroye, László Györfi, Gabor Lugosi *A probabilistic theory of pattern recognition,* Springer (1997).

3. S. Boucheron, O. Bousquet, G. Lugosi, *Introduction to Statistical Learning Theory,* In *Advanced Lectures on Machine Learning*, Springer (2004).

4. A. S. Tsybakov, *Introduction to Nonparametric Estimation,* Springer (2009).