Annual Summer School on Mathematical Aspects of Data Science
Mini Course Abstracts

22–26 Jun 2026

June 8, 2026

Foundations of Online Learning with Application to Digital Markets

Nicolò Cesa-Bianchi

University of Milan, Italy

Online learning explores algorithms that acquire knowledge sequentially, through repeated interactions with an unknown environment. The general goal is to understand how fast an agent can learn based on the information received from the environment. Digital markets, with their complex ecosystems of algorithmic agents, offer a rich landscape of sequential decision-making problems, characterized by diverse decision spaces, utility functions, and feedback mechanisms. This mini-course introduces fundamental algorithms for online learning under partial feedback, demonstrating their practical application in algorithmic trading and auction bidding.

Statistical Physics of Random Computational Problems

Shuangping Li
Yale University, USA

Random computational problems often exhibit rich phenomena and sharp phase transitions in the high-dimensional limit. As the density of constraints varies, problems such as random k-SAT, graph coloring, independent set, and perceptron models undergo transitions not only in satisfiability, but also in the geometry of their solution spaces and in the performance of algorithms. Statistical physics provides a powerful language for describing these phenomena through notions such as free energy, overlaps, clustering, condensation, freezing, and replica symmetry breaking.

In these lectures, I will introduce this viewpoint, with the perceptron as the central model problem. The perceptron is a particularly useful example because it lies at the intersection of random constraint satisfaction, high-dimensional probability, spin glass theory, and statistical learning. Despite its simple definition, it displays many of the central phenomena found in random computational problems, while also differing in important ways from sparse models such as random k-SAT.

The first lecture will introduce the statistical-physics perspective on random computational problems and the typical phase-transition picture. It will also develop some of the basic tools and concepts needed for the subsequent lectures. The second lecture will focus on spherical, binary, and related perceptron models, emphasizing capacity, margin, overlap structure, and the geometry of the solution space. The final lecture will discuss algorithmic questions, including the gap between existence and efficient search, geometric obstructions such as clustering, freezing, and the overlap gap property, and open problems at the interface of probability, statistical physics, and theoretical computer science.

Moment and Deviation Inequalities for Random Matrices and their Applications

Stanislav Minsker

University of Southern California, USA

This short course will introduce the audience to concentration-of-measure and moment inequalities for sums of independent random matrices, highlighting several classical applications. We will present simple proofs of matrix Khintchine's inequality, matrix Bernstein's inequality and matrix Rosenthal's inequality. We will then explore various extensions such as the dimensional dependence in the bounds and introduce tools developed by M. Talagrand that enable versions of the aforementioned inequalities under weaker moment assumptions.

Dynamical Systems, Graphical Models, and Language Models

Ankur Moitra

Massachusetts Institute of Technology, USA

In this tutorial I will revisit some classic learning problems in a new light:

1. Learning linear dynamical systems. Is it possible to learn from a single trajectory even when there are long range correlations?

2. Learning graphical models. Is learning from the Glauber dynamics actually computationally easier than learning from iid samples?

3. Learning sequence models. Can we hope for algorithms that work in greater generality when we are given access to a conditional sampling oracle?

All of these examples feature interesting ways that classical statistical tools can be used in dynamical settings.

Mathematical Frameworks for Generalization in Learning Theory

Shay Moran

Technion, Israel

We will discuss several related mathematical frameworks for generalization in supervised learning, with a focus on classification. We begin with the classical PAC framework of Vapnik–Chervonenkis and Valiant, which provides a clean and powerful worst-case theory of learning. While highly successful, this framework does not fully capture some aspects of modern practice, where algorithms often generalize far better than worst-case guarantees suggest.

I will then present two seemingly modest refinements of the PAC model that go a long way. These allow us to capture, in a clean mathematical way, situations in which the data distribution has "nice" structure that facilitates learning, and help explain when and why faster learning rates are possible.

In contrast to classical PAC learning - whose algorithmic landscape is by now well understood - these refined frameworks remain largely unexplored. I will conclude, time permitting, with open questions highlighting what we still do not understand about algorithms and optimal learning rates in these settings.

Optimal Mean Estimation and Robust Statistics

Roberto Imbuzeiro Oliveira
IMPA, Brazil

Take a statistical problem: say, estimating the mean of a random vector, or performing least-squares regression. What is the best non-asymptotic high-probability error guarantee that one can obtain for this problem? This kind of question is usually studied under the assumption of sub-Gaussian or otherwise light-tailed data. In this minicourse, by contrast, we allow for data that has heavy tails and contain a small number of outliers. The general theme of this course is that very strong results are still possible, including "sub-Gaussian" style bounds under finite moment assumptions. Topics to be covered include mean estimation in 1d; mean estimation for families of functions and high-dimensional vectors; application to regression with mean-squared loss; extensions to other problems; and some interesting open questions.

Gradient-free Stochastic Optimization

Alexandre Tsybakov
CREST-ENSAE Paris, France

This mini-course will deal with optimization problems in a statistical learning setup where the learner has no access to unbiased estimators of the gradient of the objective function. It includes stochastic optimization with zero-order oracle, continuum bandit and contextual continuum bandit problems. I'll give an overview of recent results on minimax optimal algorithms for these problems.