Lesson plan/Syllabus:

Both weeks of the school will follow the following format.

- *9:30-11:00. Lecture 1.*
- *11:00-11:30. Coffee break*
- *11:30-13:00. Problem Session lead by lecturers and teaching assistants*
- *13:00-14:30. Lunch*
- *14:30-16:00. Lecture 2.*
- *16:00-16.30. Coffee break*
- *16.30-18.00 Problem Session lead by TAs and lecturers*

Program:

The course by Andrea Montanari the following topics will be covered:

1.Empirical risk minimization and empirical process theory.
Objective: Understand how learning is formalized in terms of generalization guarantees, and how the latter emerge from uniform control over suitably restricted function classes.
a.Uniform convergence guarantees for learning
b.Complexity bounds for neural networks
c.Tools: Strong laws of large numbers, concentration of measure, Radamacher complexity.

2.Generalization in the linear regime.
Objective: Understand how and when good generalization can be achieved while violating the (naive) assumptions of uniform convergence. In particular, characterize the generalization error of kernel and random features methods.
a.Interpolation and benign overfitting
b.Kernel methods
c.Random features and neural tangent models
d.Tools: Random matrix theory

3.Feature learning in large networks.
Objective: Analysis of simple neural network models (two-layer network) outside the linear/neural tangent regime.
a.Mean field theory
b.Multi-index models
c.Tools: Propagation of chaos, Wasserstein gradient flows.

4.Sampling and generative methods.
Objective: Describe and justify mathematically architecture for generative models. Presents correctness guarantees, as well as limits in some simple settings.
a.Autoregressive models, diffusion models, probability flows.
b.Sampling guarantees under assumptions on the score estimation.
c.Impossibility results.

d.Tools: Stochastic localization, functional inequalities, computation-information gaps.

The lectures by Francesca Mignacco will present an overview on the statistical physics approach to high-dimensional learning problems. The main concepts and methods from mean-field theory of disordered systems will be introduced through prototypical examples that are amenable to analytic characterization. The course syllabus will be as follows:

1. Motivation and background.
Objective: Introduce the statistical physics perspective on high-dimensional learning problems and the main concepts and settings covered throughout the course.
(a) The jargon of statistical physics: thermodynamic limit, order parameters, typical-case scenario.
(b) The perceptron.
(c) The teacher-student paradigm.

2. Statics of learning.
Objective: Understand the Bayes-learning framework and how to apply it to study the impact of data structure and architectural bias on the performance.
(a) Generalization error in the perceptron: Bayes-optimal performance vs empirical risk minimization. Introduction to the replica method.
(b) Memory capacity: from random points to neural manifolds. Applications to deep learning and neuroscience.
(c) Models of data structure.
(d) Deep linear networks and back-propagating kernel renormalization

.

3. Dynamics of learning.
Objective: Understand how to derive effective low-dimensional descriptions of the learning dynamics and apply them to study commonly-used training algorithms such as stochastic gradient descent.
(a) Online learning in two-layer neural networks.
(b) Batch learning: dynamical mean-field theory, cavity method and path integral formulation.