In-depth tutorials with Practical Sessions will take place on DAY 4 & 5

40 participants per session: registrations after application acceptance.
The sessions will be one day long.
A laptop is required.

Bayesian modeling and inference

Schedule: 1-day long session repeated twice (June 27-28)
Organizers: Sinead WILLIAMSON, Evan OTT
Abstract: Bayesian methods offer natural ways to express uncertainty about model parameters, to share information between model components in a principled manner, and to incorporate prior knowledge into our learning problem. In this tutorial we will focus on Bayesian models for supervised learning, from Bayesian linear and logistic regression, to Gaussian processes. While the focus of the course is on modeling, we will also discuss common inference methods such as MCMC and Variational Inference. We will apply the methods we learn about to some real-world datasets, and compare with common non-Bayesian analogues.
Requirements: Requirements: Python, Tensorflow, Tensorflow Probability, Jupyter Notebooks.

Causality

Schedule: 1-day long session repeated twice (June 27-28)
Organizers: Jonas PETERS, Rune CHRISTIANSEN
Abstract: In the field of causality we want to understand how a system reacts under interventions (e.g. in gene knock-out experiments). These questions go beyond statistical dependences and can therefore not be answered by standard regression or classification techniques. In this part of the program you will learn about the interesting problem of causal inference and recent developments in the field. No prior knowledge about causality is required.

Part 1: We introduce structural causal models and formalize interventional distributions. We define causal effects and show how to compute them if the causal structure is known.
Part 2: We present three ideas that can be used to infer causal structure from data: (1) finding (conditional) independences in the data, (2) restricting structural equation models and (3) exploiting the fact that causal models remain invariant in different environments.
Part 3: We show how causal concepts could be used in more classical statistical and machine learning problems.
Requirements: We will use jupyter notebooks (joint work with Niklas Pfister) during the course. Please download them here http://web.math.ku.dk/~peters/jonas_files/2019-01-23-causal-notebooks.zip and try to run setupNotebook.ipynb Further details are shown below. I am looking forward to meet all of you in Palaiseau!
Infos on jupyter notebooks: Niklas Pfister and me have prepared some jupyter notebooks, which you will be able to work on during the session. We would therefore encourage you to install

jupyter with an R kernel on your laptop (see below). Please try to get things working before the winter school but if there are persistent problems, it suffices to (a) find a colleague who has a running version of jupyter or to (b) use R together with the pdf-versions of the notebooks.

1. For installing anaconda, I am using http://docs.anaconda.com/anaconda/install/linux/ and https://irkernel.github.io/installation/#linux-panel The sites also contain relevant links if you use Windows or Mac. Installing anaconda requires a lot of disk space and there are more minimalistic options, too.

2. Please download the notebooks here: http://web.math.ku.dk/~peters/jonas_files/2019-01-23-causal-notebooks.zip (Please let me know if you believe that I forgot to add a file.)

3. Once you have a running version of jupyter, start it, e.g., by using jupyter notebook in your terminal. You can then check if everything is correctly set up by running setupNotebook.ipynb notebook (use the R kernel). This also tells you which additional R packages you need to install. If steps 1.-3. fail, run setupNotebook.r in R.

4. Remind yourself on some R syntax: https://www.rstudio.com/resources/cheatsheets/

Classical algorithms and matrix factorization
Schedule: 1-day long session repeated twice (June 27-28)
Organizer: Olivier KOCH
Abstract: TBA
Requirements: TBA

Deep generative models
Schedule: 1-day long session repeated twice (June 27-28)
Organizers: Mario LUCIC, Marcin MICHALSKI
Abstract: TBA
Requirements: TBA

Fundamentals of text analysis for user generated content

Schedule: 1-day long session repeated twice (June 27-28)

Organizer: Steven R. WILSON

Abstract: How do bloggers in different countries express their personal beliefs? What are Twitter users saying about Brexit? Which community on Reddit uses the most positive language? In this tutorial, we will explore the basic tools needed to apply natural language processing techniques to answer these types of questions. Dealing with user-generated text brings unique challenges, such as the use of non-standard language (e.g., slang, hashtags, and emoji), and also unique opportunities, such as the ability to automatically discover trends in the views and sentiments huge numbers of users. During this tutorial, participants will have the chance to formulate their own research questions and employ useful natural language processing methods to start to answer them. Topics to be covered include:

- Preprocessing noisy text data
- Content analysis of user-generated text
- Supervised learning using user-generated text
- Getting insights from statistical NLP models

Requirements: Basic programming knowledge in Python. (Details about specific Python package requirements to follow)

Hyper-parameter selection with Bayesian optimization

Schedule: 1-day long session repeated twice (June 27-28)

Organizers: Matthew B. BLASCHKO, Amal RANNEN TRIKI

Abstract: In this module, we will cover the theory and practice of hyperparameter selection using Bayesian optimization. Bayesian optimization is closely related to optimal experimental design, and iteratively refines a proxy model by selecting a new point to evaluate. In the application of hyperparameter selection in machine learning, the evaluation can be performed by training and testing a model with hyperparameters determined by the Bayesian optimization procedure. The resulting procedure is more efficient than grid search, and more principled than stochastic search algorithms such as evolutionary computing. The theory section will cover aspects of Gaussian process modeling (the most common model underlying Bayesian optimization), acquisition functions, and model selection in machine learning. In the practical section, you will get hands on experience setting up and applying state-of-the-art Bayesian optimization software packages to hyperparameter search. The practical section will be given in Python.

Requirements: Basic programming knowledge in Python.

Introduction to deep learning with Keras
Schedule: 1-day long session repeated twice (June 27-28)
Organizer: Olivier GRISEL
Abstract: TBA
Requirements: TBA

Learning with positive definite kernels: Theory, algorithms and applications
Schedule: 1-day long session repeated twice (June 27-28)
Organizers: Bharath K. SRIPERUMBUDUR, Dougal J. SUTHERLAND
Abstract: The course provides a broad introduction to the topic of learning with positive definite kernels from the view points of theory, algorithms and applications. The course is conceptually divided into 3 parts. In the first part, we will motivate the overall course through a simple nonlinear classification problem, leading to the notion of a positive definite kernel (kernel, in short). We will explore this notion of kernel from feature space and function space points of view with the former being particularly useful to develop algorithms and the latter being useful to understand the related mathematical aspects. Using both these view points, we will investigate the role of kernels in popular machine learning and statistical methodologies such as M-estimation and Principal component analysis. The second part deals with modern aspects and novel applications of kernels to non-parametric hypothesis testing (including goodness-of-fit, homogeneity, independence and conditional independence), which hinges on the notion of kernel embedding of probability measures. We will explore the mathematical aspects of kernel embedding and discuss the aforementioned applications. The last part exposes the recent developments on computational vs. statistical trade-off in learning with kernels. This is an important line of ongoing research which addresses the inherent computational difficulties with kernel algorithms.

The topics covered in the lectures will be further developed and explored in lab sessions handled by Dr. Dougal Sutherland.
Requirements: TBA

Mathematics of data: From theory to computation

Schedule: 1-day long session (June 28)

Organizers: Volkan CEVHER, Armin EFTEKHARI, Thomas SANCHEZ, Paul ROLLAND

Abstract: Convex optimization offers a unified framework in obtaining numerical solutions to data analytics problems with provable statistical guarantees of correctness at well-understood computational costs. To this end, this course reviews recent advances in convex optimization and statistical analysis in the wake of Big Data. We provide an overview of the emerging convex data models and their statistical guarantees, describe scalable numerical solution techniques such as stochastic, first-order and primal-dual methods. Throughout the course, we put the mathematical concepts into action with large scale applications from machine learning, signal processing, and statistics.

Learning outcomes: By the end of the course, the students are expected to understand the so-called time-data tradeoffs in data analytics. In particular, the students must be able to:

1. Choose an appropriate convex formulation for a data analytics problem at hand.

2. Estimate the underlying data size requirements for the correctness of its solution.

3. Implement an appropriate convex optimization algorithm based on the available computational platform.

4. Decide on a meaningful level of optimization accuracy for stopping the algorithm.

5. Characterize the time required for their algorithm to obtain a numerical solution with the chosen accuracy.

Requirements: Previous coursework in calculus, linear algebra, and probability is required. Familiarity with optimization is useful.

Teaching material and website: https://lions.epfl.ch/mathematics_of_data/
                                   http://lions.epfl.ch/teaching

Neural networks and causal recommendation

Schedule: 1-day long session repeated twice (June 27-28)

Organizer: Flavien VASILE

Abstract: TBA

Requirements: TBA

Optimization for machine learning and deep learning
Schedule: 1-day long session (June 27)
Organizers: Martin JAGGI, Thijs VOGELS
Abstract: This course will give an overview of modern mathematical optimization method for applications in machine learning and deep learning. In particular, scalability of algorithms to large datasets will be discussed in theory and in implementation (Python).

Contents: Gradient Methods (including Proximal, Subgradient, Stochastic) for ML and deep learning, Convex and Non-convex Convergence analysis, Derivative-Free Optimization.
Parallel and Distributed Optimization Algorithms for ML and DL, Communication efficient methods, Decentralized (server-less) methods.
Optional: Coordinate Descent, Frank-Wolfe, Accelerated Methods, Second-Order Methods including Quasi-Newton Methods.
Practical Python exercises, lecture notes & slides available
here:https://github.com/epfml/OptML_course
Requirements:*- Mathematical Background (linear algebra and basic probability)
- Basic Python/numpy/matplotlib with Jupyter notebooks

Shallow dive in deep reinforcement learning
Schedule: 1-day long session repeated twice (June 27-28)
Organizers: Bilal PIOT, Diana BORSA, Pierre H. RICHEMOND
Abstract: TBA
Requirements: TBA

Theory and applications of submodular optimization: From discrete to continuous and back

Schedule: 1-day long session repeated twice (June 27-28).

Organizer: Hamed HASSANI

Abstract: Many scientific and engineering models feature inherently discrete decision variables — from phrases in a corpus to objects in an image. The study of how to make (near-)optimal decisions from a massive pool of possibilities is at the heart of combinatorial optimization problems. In this regard, submodularity has proven to be a key combinatorial structure that can be exploited to provide efficient algorithms with strong theoretical guarantees. This tutorial aims to provide a deep understanding of the various frameworks that have been recently developed for submodular optimization in the presence of the modern challenges in machine learning and data science. In particular, we will discuss challenges such as large-scale, online, distributed, streaming, robust, and stochastic submodular maximization/minimization and illustrate the discrete and continuous based frameworks to address these challenges. A particular emphasis is on the current research directions as well as concrete exemplar applications in data science.

Requirements: Basic background in Python programming + laptop with Python environment (including numpy, scipy, matplotlib, jupyter notebook).