CS-439

# Optimization for machine learning

Flammarion Nicolas, Jaggi Martin

| Cursus | Sem. | Type |
|---|---|---|
| Computer science | MA2, MA4 | Opt. |
| Cybersecurity | MA2, MA4 | Opt. |
| Data Science | MA2, MA4 | Obl. |
| Data science minor | E | Opt. |
| Electrical Engineering | | Opt. |
| Minor in statistics | E | Opt. |
| Neuro-X minor | E | Opt. |
| Neuro-X | MA2, MA4 | Opt. |
| Quantum Science and Engineering | MA2, MA4 | Opt. |
| SC master EPFL | MA2, MA4 | Opt. |
| Statistics | MA2, MA4 | Opt. |

| | |
|---|---|
| Language of teaching | English |
| Credits | 8 |
| Session | Summer |
| Semester | Spring |
| Exam | Written |
| Workload | 240h |
| Weeks | 14 |
| **Hours** | **5 weekly** |
| Courses | 2 weekly |
| Exercises | 2 weekly |
| Lab | 1 weekly |
| **Number of positions** | |


## Summary

This course teaches an overview of modern optimization methods, for applications in machine learning and data science. In particular, scalability of algorithms to large datasets will be discussed in theory and in implementation.


## Content

This course teaches an overview of modern optimization methods, for applications in machine learning and data science. In particular, scalability of algorithms to large datasets will be discussed in theory and in implementation.
Fundamental Contents:

- Convexity, Gradient Methods, Proximal algorithms, Stochastic and Online Variants of mentioned methods, Coordinate Descent Methods, Subgradient Methods, Non-Convex Optimization, Frank-Wolfe, Accelerated Methods, Primal-Dual context and certificates, Lagrange and Fenchel Duality, Second-Order Methods, Quasi-Newton Methods, Gradient-Free and Zero-Order Optimization.

Advanced Contents:

- Non-Convex Optimization: Convergence to Critical Points, Saddle-Point methods, Alternating minimization for matrix and tensor factorizations
- Parallel and Distributed Optimization Algorithms, Synchronous and Asynchronous Communication
- Lower Bounds

On the practical side, a graded **group project** allows to explore and investigate the real-world performance aspects of the algorithms and variants discussed in the course.

## Keywords

*Optimization, Machine learning*

## Learning Prerequisites

### Recommended courses

- CS-433 Machine Learning

### Important concepts to start the course

- Previous coursework in calculus, linear algebra, and probability is required.
- Familiarity with optimization and/or machine learning is useful.

## Learning Outcomes

By the end of the course, the student must be able to:

- Assess / Evaluate the most important algorithms, function classes, and algorithm convergence guarantees
- Compose existing theoretical analysis with new aspects and algorithm variants.
- Formulate scalable and accurate implementations of the most important optimization algorithms for machine learning applications
- Characterize trade-offs between time, data and accuracy, for machine learning methods

## Transversal skills

- Use both general and domain specific IT resources and tools
- Summarize an article or a technical report.

## Teaching methods

- Lectures
- Exercises with Theory and Implementation Assignments

## Expected student activities

Students are expected to:

- Attend the lectures and exercises
- Give a short scientific presentation about a research paper
- Read / watch the pertinent material
- Engage during the class, and discuss with other colleagues

## Assessment methods

- Continuous control (course project)
- Final Exam

## Resources

### Websites

- https://github.com/epfml/OptML_course

### Moodle Link

- https://go.epfl.ch/CS-439

### Videos

- https://www.youtube.com/playlist?list=PL4O4bXkI-fAeYrsBqTUYn2xMjJAqlFQzX