- Faculté des sciences
- www.unine.ch/sciences

Fairness and Privacy in Machine Learning (3IN2067)

Filières concernées	Nombre d'heures	Validation	Crédits ECTS
Master en informatique	**Cours: 2 ph** **Exercice: 2 ph**	**Voir ci-dessous**	5

ph=période hebdomadaire, pg=période globale, j=jour, dj=demi-jour, h=heure, min=minute

Période d'enseignement:

- Semestre Automne

Contenu

This course gives a thorough introduction to algorithms and theory for achieving privacy and fairness in modern computing systems. In particular, we will focus on one formal model of privacy, differential privacy, with solid theoretical properties. We will then study various notions of fairness, how they can be satisfied, and at what cost. Finally, we will consider the connection of privacy to reproducibility in research.

- Privacy introduction: k-anonymity
-Differential privacy: the randomised response mechanism
-Central privacy: the Laplace mechanism
-Privacy in auctions: the Exponential mechanism
-Case study: Federated learning
-Fairness introduction: parity
-Discrimination: Bias, calibration and conditional independence
-Fairness as smoothness, meritocracy
-Explainability and actionable recourse
-Case study: college admissions
-Reproducibility introduction: p-value hacking
-Privacy, robustness, and reproducibility

Forme de l'évaluation

Assessment: Project work (80%: including group project and assignments) and exam (20%)

Documentation

https://github.com/olethrosdc/ml-society-science

Pré-requis

Requirements:
- Introductory probability (probability, expectation, conditional probability and expectation)
- calculus (derivatives, integrals)
- programming (basic python skills)

It also helps to have some background in statistics and machine learning. For students wanting to do a thesis on the subject, it is advised to also take the Advanced Seminar on Privacy, Fairness and Reproducibility.

Forme de l'enseignement

Lectures, live coding, exercises, assignments

URLs	1) https://github.com/olethrosdc/ml-society-science