

CS-421

**Machine learning for behavioral data**

Käser Tanja

Cursus	Sem.	Type
Computer science	MA2, MA4	Opt.
Cybersecurity	MA2, MA4	Opt.
Data Science	MA2, MA4	Opt.
Digital Humanities	MA2, MA4	Opt.
Learning Sciences		Obl.
SC master EPFL	MA2, MA4	Opt.

Language of teaching	English
Credits	4
Session	Summer
Semester	Spring
Exam	Written
Workload	120h
Weeks	14
<b>Hours</b>	<b>4 weekly</b>
Courses	2 weekly
Project	2 weekly
<b>Number of positions</b>	

**Summary**

Computer environments such as educational games, interactive simulations, and web services provide large amounts of data, which can be analyzed and serve as a basis for adaptation. This course will cover the core methods of user modeling and personalization, with a focus on educational data.

**Content**

The users of computer environments such as intelligent tutoring systems, interactive games, and web services are often very heterogeneous and therefore it is important to adapt to their specific needs and preferences.

This course will cover the core methods of adaptation and personalization, with a focus on educational data. Specifically we will discuss approaches to the task of accurately modeling and predicting human behavior within a computer environment. Furthermore, we will also discuss data mining techniques with the goal to gain insights into human behavior. We will cover the theories and methodologies underlying the current approaches and then also look into the most recent developments in the field.

1. 'Cycle' of adaptation : representation, prediction, intervention (e.g. recommendation)
2. Data Processing and Interpretation (missing data, feature transformations, distribution fitting)
3. Performance evaluation (cross-validation, error measures, statistical significance, overfitting)
4. Representation & Prediction (probabilistic graphical models, recurrent neural networks, logistic models, clustering-classification approaches)
5. Recommendation (collaborative filtering, content-based recommendations, multi-armed bandits)
6. Stealth Assessment (seamless detection of user traits)
7. Multimodal analytics (represent & analyze data from non-traditional sources. i.e. sensors, classroom analytics, human-robot interaction)

**Learning Prerequisites****Required courses**

The student must have passed a course in probability and statistics and a course including a programming project

**Recommended courses**

- CS-433 Machine learning or
- CS-233a / CS-233b Introduction to machine learning

**Important concepts to start the course**

Probability and statistics, basic machine learning knowledge, algorithms and programming, Python

**Learning Outcomes**

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

- Explain the main machine learning approaches to personalization, describe their advantages and disadvantages and explain the differences between them
- Implement algorithms for these machine learning models
- Apply them to real-world data
- Assess / Evaluate their performance
- Explain and understand the fundamental theory underlying the presented machine learning models

### Teaching methods

- Lectures
- Weekly lab sessions
- Course project

### Expected student activities

- Attend the lectures
- Attend the lab sessions and work on the homework assignments
- Project work

### Assessment methods

- Project work (50%)
- Final exam (50%)

### Supervision

Office hours	Yes
Assistants	Yes
Forum	Yes