

Project Guidelines

Machine Learning for Behavioral Data
February 17, 2026

Overview

- Teams of 3 people
- We will provide the data sets (from EdTech Start-Ups)
- We will provide example research questions
- You will suggest an additional analysis/extension to the selected research question
- We will give feedback during the semester (see milestones)
- You will do a **poster presentation** in the last week of the semester on May 26 (13:15 – 16:00), **attendance is mandatory**
- Final project (Code + Report) delivered by **June 4, 2026 23:59 CET**

Scope

Work on 2-3 research questions:

- At least 1 question needs to be on predictive modeling or unsupervised learning
- At least 1 question needs to be on ethical machine learning (fairness or explainability)

Research Questions

Research question 1a): Predictive Modeling

Focus on developing and evaluating predictive models (Bayesian Knowledge Tracing (BKT), Deep Knowledge Tracing (DKT), Neural Networks,...)

OR

Research Question 1b): Unsupervised Learning

Focus on discovering patterns or structures in data (Clustering and Pattern Discovery)

AND

Research Question 2: Ethical Machine Learning

Focus on fairness or explainability (Bias detection and mitigation techniques, Fairness metrics, Explainable AI methods)

| Week | Lecture & Lab (13:15-16:00) | Project Hours (16:15-17:00) | Milestones |
|------|-----------------------------------|--|---|
| 1 | Introduction + Data Exploration | Project Presentation + Environment Setup | - |
| 2 | Regression | Introduction to tasks for M2 | <i>M1: Preferences on team members and data sets (Feb 25, 23:59)</i> |
| 3 | Classification + Model Evaluation | Office hours – Individual Drop-In | |
| 4 | Time Series Prediction | Office hours – Individual Drop-In | |
| 5 | Time Series Prediction | Introduction to tasks for M3/M4 | <i>M2: Individual exploration of selected data set (March 19, 23:59)</i> |
| 6 | Time Series Prediction | Office hours – Individual Drop-In | <i>M3: selection of research question and approach (March 26, 23:59)</i> |
| 7 | Unsupervised Learning | Individual discussion with teams | |
| 8 | Spring Break | | |
| 9 | Unsupervised Learning | Office hours – Individual Drop-In | |
| 10 | Ethical Machine Learning | Team Coaching | |
| 11 | Ethical Machine Learning | Office hours – Individual Drop-In | <i>M4: submission of results for first research question (April 30, 23:59) M5: ideas for extension and approach (April 30, 23:59)</i> |
| 12 | Reserve | Individual discussion with teams | |
| 13 | No lecture | Office hours – Individual Drop-In | |
| 14 | No lecture | Office hours – Individual Drop-In | |
| 15 | Poster Session | | <i>M6: Poster Session, May 26 (13.15-17.00, EPFL Campus)</i> |
| 16 | Semester End | Semester End | <i>M7: Hand in report and code base (June 4, 23:59)</i> |

Mandatory Attendance

- Individual team discussion (feedback on M2/M3)
 - March 31, 2026: 10-15 minutes slot between 15:15-17:00
- Individual team discussion (M4 progress)
 - April 21, 2026: 10-15 minutes slot between 15:15-17:00
- Individual team discussion (feedback on M4/M5)
 - May 5, 2026: 10-15 minutes slot between 15:15-17:00
- Poster presentations
 - May 26, 2026: 10-minutes slot between 13:15-15:00

Grading

- M2 (individual, notebook): 15%
- M4 (code, short report): 25%
- M6 (poster, presentation): 20%, attendance mandatory
- M7 (code, report): 40%