

# Human Centered Machine Learning

Course code	INFOMHCML
Academic year	2020-2021
Period	4
Level	MSc
ECTS Credits	7.5
Department	Department of Information and Computing Sciences, Utrecht University
Notification	All information in this Course Syllabus as well as on Blackboard overrules and supersedes less updated information from other sources, especially course information on Osiris and the CS education page.

<b>Introduction</b>	<b>3</b>
<b>Practical information</b>	<b>3</b>
Literature	3
Microsoft Teams and Blackboard	3
Study load	3
Feedback	4
<b>Midterm</b>	<b>4</b>
<b>Paper presentations</b>	<b>5</b>
Instructions	5
Grading	5
<b>Programming assignments</b>	<b>5</b>
<b>Project</b>	<b>6</b>
Objectives	6
Instructions	6
Result, assessment and feedback	6
<b>Grading</b>	<b>8</b>
<b>Contact</b>	<b>8</b>
<b>Schedule</b>	<b>9</b>

# Introduction

The impact of machine learning (ML) systems on our society has been increasing rapidly, ranging from systems that influence the content that we see online (e.g., ranking algorithms, advertising algorithms) to systems that enhance or even replace human decision making (e.g. in hiring processes). However, machine learning systems often perpetuate or even amplify societal biases—biases we are often not even aware of. What's more, most machine learning systems are not transparent, which hampers their practical uptake and makes it challenging to know when to trust (or not trust) the output of these systems.

This course will familiarize students with a growing set of concepts and techniques to develop and assess machine learning systems, so that these systems are fair and interpretable, and can be used in responsible ways. More specifically, we will discuss methods to measure the fairness of ML systems and to make ML systems fairer. The course also covers different approaches to creating and evaluating interpretable or explainable ML models.

The course will cover examples from various areas of AI. Given the expertise of the lecturers we will also zoom in on specific examples from natural language processing and multimodal affective computing research. Our discussion will also be informed by relevant literature from the social sciences. An interest in these areas is therefore desirable.

## Practical information

### Literature

See Blackboard for the required literature. During the lectures, we may also reference additional background material.

### Microsoft Teams and Blackboard

**Blackboard:** For announcements and all course material (slides, exercises, literature). We will also use Blackboard for assignment submissions and grading.

**Microsoft Teams:** The lectures and office hours will be held using Microsoft Teams. You can also use Microsoft Teams for more questions, e.g. about the organization. Feel free to also use Microsoft Teams to share useful background material, or interesting news/academic articles that are relevant to the course.

### Study load

This course is 7.5 ECTs which means 210 hours per person.

## Feedback

This is the first year that this course is being offered (but hopefully there will be many more editions!). Please do not hesitate to share your feedback with us, as this helps us to improve the course. Besides the Caracal evaluation (at the end), we will also send around a short feedback form about halfway of the course. Of course, feel free to also send us feedback during other moments of the course.

## Midterm

The midterm is a *closed book* exam. The midterm will be on location and you are allowed to bring one A4 (double sided) with notes. You should provide your answers in English. The midterm covers all material up until the midterm.

# Paper presentations

## Instructions

During the course you will present two papers, one on fairness and one on explainability. These paper presentations will be carried out in groups of 4. For both paper presentations, we will assign the groups *randomly* (so you will be in two different random groups).

### **Presentation format**

A paper presentation should be at most 12 minutes ( +3 minutes for questions). Make sure to go beyond “just” a summary of the paper. For example, relate the paper contents to what has been discussed in the course, highlight strong and weak points of the paper, follow up questions, ideas for future work, etc. Keep in mind that the other students have most likely not read the paper! You should prepare the presentation together. All members of the group need to be present during the presentation, but not everyone needs to actually present (up to you!).

### **Paper selection + deadlines**

We have prepared a list with papers to select from that will be provided on Blackboard (“Course Content”). As a group, provide your top 3 paper preferences to dr. Nguyen and dr. Kaya by email. Your email subject should be in the following format: “[*INFOMHCML*] Group X paper preference”, where X is your group number. Include all group members in CC. We’ll try to assign groups their preferred papers, but we’ll also try to balance so that the same paper won’t be presented by many groups. See the schedule for the deadline.

## Grading

Each presentation will be graded based on the following criteria:

- content: 60%
  - Understanding of scientific content - 30%
  - Critical reflection (e.g., strengths, weaknesses, follow up questions) - 15%
  - Relation to course material - 15%
- slides: 20%
- delivery (e.g., pace, voice): 20%

# Programming assignments

There are two programming assignments (one on fairness, one on explainability). The programming assignments will be carried out in pre-assigned pairs. The programming language is Python. You will need to deliver both the code as well as a short report for each assignment. We will have two sessions where the TA is available for questions on these programming assignments (see schedule). However, please start early with the assignments so that you can get the most out of these sessions. The exact assignments will be posted on Blackboard. For questions about the assignments, please use the 'Lab' channel on MS teams.

## Project

### Objectives

This is the final project of the course carried out in small student groups. In the first part of this course you have learned about core concepts and techniques from human-centered ML. The overall objective is to synthesize the techniques and concepts learned in a larger project. More specifically you will:

- Gain hands-on experience with carrying out research and applying the learned techniques.
- Gain experience in communicating your findings in a written report and presentation.

### Instructions

Groups should be composed of 4 members and can be self-assigned. You are encouraged to compose groups with members that complement each other's skills and background. You will work closely together throughout the whole research process (designing and motivating a research question, diving into related work, carrying out an experiment, writing a paper).

Your project should build on the concepts and techniques learned in this course. Given the time frame, *do not propose a project for which you still need to collect data to train a machine learning model.*

### Result, assessment and feedback

See Blackboard for submission instructions.  
For the deadlines, see the [schedule](#).

### Project proposal

One paragraph (max. 200 words) with the project idea + group members.

*Assessment:* Instructors will review the proposals and send short feedback.

*Grading:* Go/no go. *No go* meaning: come up with a new project idea or heavily revise the current one. In case of a *no go*, you have 5 days to resubmit your proposal.

## Project updates

In week 8 (see schedule) each group gives a short pitch presentation of the project (3 min). This allows you to get feedback from the lecturers and peers.

## Project office hours

Teachers are available for advice on the projects.

## Project (poster) presentations

The projects will be presented in the poster area of [Gather Town](#). Each group should prepare a poster and submit the poster to dr. Kaya by email the previous day at 17:00 (see the schedule for the deadline), so that the posters will be uploaded to the area before the session. Your email subject should be in the following format: “[*INFOMHCML*] Group X poster presentation”, where X is your group number. All group members should be present in Gather Town between 9:00 and 11:00 to present the project and respond to questions asked by visitors.

The goal is to share the main findings but also gain feedback for the final bits of the project.

- Present your summary and conclusions.
- Present things that are most debatable/interesting to the group.
- Any puzzling findings?
- What questions do you have to your fellow students? What would you like to have input on to your fellow students?
- Self-evaluation: What went well? What could be improved?

## Project-report

The report should have the structure of an academic research paper. Use the paper format from ACM FAccT <https://facctconference.org/2021/cfp.html> (**max 8 pages, including references**). Also include one or two paragraphs or a table in which you shortly describe how each project member contributed to each component of the project.

The report will be **assessed** based on the following components:

- Introduction/problem statement
- Related work
- Methodology
- Execution of the experiments
- Discussion and conclusion (including elaboration on findings and limitations)
- Writing
- Ethical considerations (include a short section in your report discussing ethical considerations)

Each component will be assessed on a 5-point scale:

1. Insufficient.
2. Barely sufficient.
3. Sufficient. The typical student will achieve this level.
4. Above standard.
5. Outstanding. Meant for work that exceeds expectations by far. Only rarely awarded.

## Overall grade

The grade is based on the final grade of the project report. All project deliverables are required to pass the course.

# Grading

Your grade will be determined as follows:

Paper presentation: fairness	10%
Paper presentation: explainability	10%
Midterm	40%
Project	40%
Programming assignments	<i>Pass or fail</i>

We offer a few repair options:

- **Midterm:** We will offer a retake, but only if the  $4.0 \leq \text{midterm grade} < 5.5$ . The retake for the midterm is on the 13th of July.
- **Paper presentations:** It is not possible to repair the paper presentations.
- **Project:** It is not possible to repair the project. However, if you pass all other components, you only need to redo the project when you retake the course.
- **Programming assignments:** It is not possible to repair the programming assignments. They need to be submitted by the provided deadlines. Each week delay leads to a reduction of 0.2 points (out of 10) from the final grade.

If due to circumstances beyond your control you are not able to take the midterm, you will be offered the opportunity to participate in the retake, given that you have notified us about the situation as soon as possible via an email to both dr. Kaya and dr. Nguyen. The email should state your: 1) name, 2) student number, 3) reason for your absence.

To pass the course, the programming assignments need to be graded with a 'pass' and all other components must be graded with a 5.5 or higher.

## Contact

We prefer posting questions that would benefit other students (e.g., about organization or content) on Microsoft Teams. If you have questions that concern you specifically, you can contact us by email. If you send emails about group work, make sure to also include your group members in the cc, so that every team member has the same information.

- Dong Nguyen (course coordinator): [d.p.nguyen@uu.nl](mailto:d.p.nguyen@uu.nl)
- Heysem Kaya: [h.kaya@uu.nl](mailto:h.kaya@uu.nl)
- Yupei Du: [y.du@uu.nl](mailto:y.du@uu.nl)



# Schedule

Week	Date	Activity
1 (17)	Tuesday 27/4	No lecture! Koningsdag
	Thursday 29/4	<b>Lecture 1:</b> Introduction to the course + introduction to fairness
2 (18)	Tuesday 4/5	<b>Lecture 2:</b> Fairness <i>Release of group assignments for fairness paper presentations</i> <i>Release of fairness programming assignment</i>
	Thursday 6/5	<b>Lecture 3:</b> Fairness
	Friday 6/5	<b>Deadline (17.00):</b> Submit your top 3 paper choices for the fairness paper presentations. We will send out the paper assignments on Monday.
3 (19)	Tuesday 11/5	<b>Lab:</b> TA will be available for support with programming assignments from 9-12.45.
	Thursday 13/5	No lecture! Hemelvaartsdag
4 (20)	Monday 17/5	<b>Deadline (17.00):</b> Programming assignment fairness
	Tuesday 18/5	<b>Paper presentations:</b> Fairness <i>Release of group assignments for explainability paper presentations</i>
	Thursday 20/5	<b>Lecture 4:</b> Explainability <i>Release of explainability programming assignment</i>
	Friday 21/5	<b>Deadline (17.00):</b> Submit your top 3 paper choices for the explainability paper presentations. We will send out the paper assignments on Monday.
5 (21)	Tuesday 25/5	<b>Lecture 5:</b> Explainability
	Thursday 27/5	<b>Lecture 6:</b> Explainability
	Friday 28/5	<b>Deadline (17.00):</b> Project proposal
6 (22)	Tuesday 1/6	<b>Lab:</b> TA will be available for support with programming assignments from 9-12.45.
	Thursday 3/6	<b>Paper presentations:</b> Explainability
7 (23)	Monday 7/6	<b>Deadline (17.00):</b> Programming assignment explainability
	Tuesday 8/6	<b>Lecture 7: Guest lecture:</b> David Graus (Randstad) on fairness and recruiting systems + <b>Recap</b> /questions/project

	Thursday 10/6	<b>Midterm</b>
8 (24)	Tuesday 15/6	Optional: On campus meeting. Meet your teachers and fellow students.
	Thursday 17/6	<b>Project updates:</b> Students present short updates of projects
9 (25)	Tuesday 22/6	<b>Guest lecture:</b> Maranke Wieringa (UU) on algorithmic accountability <b>Project office hour</b>
	Thursday 24/6	<b>Guest lectures:</b> Marcel Robeer (UU/National Police Lab AI) on explainability and AI at the Dutch Police Christine Bauer (UU) on fairness in music recommendation systems
10 (26)	Monday 28/6	<b>Deadline (17.00):</b> Project poster submission
	Tuesday 29/6	<b>Project presentations:</b> Gathertown
11 (27)	Monday 5/7	<b>Deadline (17.00):</b> Project report submission