



Human-Centered Artificial Intelligence

Inteligência Artificial centrada no Humano

Mestrado em Engenharia e Ciência de Dados

University of Coimbra

2024/25

Teachers



Luis Macedo
macedo@dei.uc.pt

2



Hugo Oliveira
hroliv@dei.uc.pt

Learning outcomes

- The course aims to provide students with knowledge about an AI where human-AI cooperation should predominate, based on considering adequate levels of automation and machine autonomy, without loss of human control, and guided by human objectives and values, towards the promotion of the human condition in its various dimensions, including those outlined in declarations and reference guides such as the United Nations Sustainable Development Goals, the EU AI Act, or the UNESCO Recommendation on the Ethics of AI.
- It is expected that students will acquire knowledge and skills to study, design and develop this category of human-centered AI systems, in the context of a multi-agent system formed by humans and AI systems, towards a path of hybrid and collective intelligence, human-AI, that amplifies human intelligence without replacing it.

Recommended prerequisites

- Programming, mathematical, artificial intelligence and machine learning skills

Syllabus

- 1. Human-centered Artificial Intelligence
- 2. Human-AI Cooperation
- 3. Human in the loop machine learning
- 4. Human-AI Communication
- 5. Human-AI Trustworthiness
- 5. Applications – Recommender Systems

Teaser



<https://www.youtube.com/watch?v=tiwVMrTLUWg>

Teaser 1-b

- <https://www.youtube.com/watch?v=9nF0K2nJ7N8>
- <https://www.nytimes.com/2023/08/21/technology/waymo-driverless-cars-san-francisco.html>

Assessment rules

- 3 components:
 - Written exam [8/20]
 - Project [10/20]
 - Continuous assessment [2/20]
- Minimum for approval:
 - 35% in the Written Exam
 - 35% in the Project
 - 45% in the sum of all components (9.5/20)

Assessment rules

- Written exam [8/20]
 - With consultation (only paper, digital means not allowed)
- Project [10/20]
 - Development of a project in two topics covered by the course
 - Groups of three students
 - Two phases: Project Proposal + Project
 - Written document describing what the group intends to accomplish
 - The Project
- Continuous assessment [2/20]
 - Worksheets, Quizzes, participation in the discussions, etc.

Assessment rules

- Project schedule:
 - Theme choice: 30 september (informal)
 - Submissions of project proposal: 07 october
 - Submission of practical works: 28 november
 - Presentation of results (defenses; in class): 2 and 9 december

Communication channels

- Course's webpage in Inforestudante
- Notifications from Inforestudante/UCStudent (read your email regularly!)
- Discord:
 - <https://discord.gg/pYVpOke7> (expires in 6 days)
- Teachers' emails

How classes will work

- T classes:
 - For presentation and discussion of Syllabus' topics
 - Discussion guided by Worksheets (assignments provided previously, to be answered as homework)
 - Face-to-face
- PL classes:
 - Mostly for project work
 - Face-to-face
 - 2 teachers in the room for the two PLs
- Homework:
 - 6 ECTS mean an effort of 8h/week: **4h in classes + 4h outside of classes**

Bibliography

- S. Russell, P. Norvig. Artificial Intelligence: A Modern Approach. Pearson, Upper Saddle River, NJ, 2009/2020.
- Shneiderman, Ben. Human-centered AI. Oxford University Press, 2022.
- A. Holzinger, P. Kieseberg, E. Weippl & A Min Tjoa. Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction: From Machine Learning to Explainable AI. Springer Lecture Notes in Computer Science LNCS 11015. Cham: Springer, pp. 1-8, 2018.
- C. Molnar. Interpretable machine learning. A Guide for Making Black Box Models Explainable, 2019. <https://christophm.github.io/interpretable-ml-book/>.
- T. Miller. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38, 2019.
- F. Ricci, L. Rokach, B. Shapira. Recommender Systems Handbook. Springer, 2015.