Algorithmic foundations and ethics in AI: from theory to practice course

Toolkit for synchronous sessions

Project-based exercises

Project number | 2022-1-ES01-KA220-HED-000085257

Project-based exercises

Project-based exercises

	Description
Task description	In teams of 4-5, you will work on evaluating one of six AI solutions over the course of five course units. This comprehensive task involves a detailed analysis of the AI solution from multiple perspectives following the course structure and topics handled in each course unit.
Description of how to do the task	**Instructions** • **Your team will work on one AI solution** from the list in the page 4. • **Analyze the chosen solution** based on the five perspectives provided in page 6 • **Document your findings** in a structured report, addressing each evaluation criterion in detail. Each point should be presented in max. 2 slides • **Present your analysis** to the class in the last synchronous session, highlighting key insights and recommendations for improving the solution. • Your presentation to team should last between **5-10** minutes
Estimated time to do the task	Approx. 1-2 hrs / CU (5-10 hrs total)
Suggestion of sources for doing the task	Interactive booklets and other course material + any other sources
Detailed description of how to deliver the task	See next pages.
Information on the deadline for the task delivery	Presented in the last synchronous session.
Contact information or how to clarify doubts	The teacher must provide a form of contact (it could be an email address, a telephone number...).

Project-based exercises
Timeline and instructions

PROJECT EXERCISE TIMELINE

Introducing the project and forming teams.

- Preferably, begin working on the project after the 1st synchronous session, agree with the team on meetings to work on the project.
- Alternatively, start after the 2nd synchronous session covering CUs 1 and 2.

Review progress

- On the 3rd and 4th synchronous sessions possibility to review progress and discuss any challenges with the teacher.

Final Presentation

- On the 5th synchronous session present the project exercise as a team.
  - Presentation 5-10 minutes.
  - You can choose the PowerPoint or Canva layout, remember to make an appealing presentation.

NOTE | In case the student is unavailable for group work (or not participating actively to the project) the project exercise should be submitted individually to the teacher (10 slides).

Further instructions

- Your topic is presented on the next slide. As a team, you'll get to decide the name, purpose and more detailed content of your service. However, the application should be in the field given to you in the next page.
- Continue working on the project at the agreed pace with the project team, however the content is always related to Course units, thus working after group members have finalized e-learnings per topic, is preferred.
- You can decide on how you build your presentation; thus, the questions are there as a guidance, but you should create a content, which is engaging for the audience.

Project-based exercises
AI Solutions divided per team

The AI solution your team should be working on:

AI-powered healthcare diagnosis tool
- A tool that assists doctors by providing preliminary diagnoses based on patient symptoms and medical history.
  - XX groups working on the topic (add here names of the students, divided into groups of 4-5 by teacher)

Streaming service recommendation system
- An application that suggests movies and series to users based on their viewing history and preferences.
  - XX groups working on the topic (add here names of the students, divided into groups of 4-5 by teacher)

Smart home automation system
- A solution that automates home devices such as lights, thermostats, and security systems based on user routines and preferences.
  - XX groups working on the topic (add here names of the students, divided into groups of 4-5 by teacher)

AI-driven financial investment advisor
- An advisor that provides investment recommendations based on market data and user financial goals.
  - XX groups working on the topic (add here names of the students, divided into groups of 4-5 by teacher)

Online shopping tool application
- An app that recommends products to users based on their browsing and purchasing history.
  - XX groups working on the topic (add here names of the students, divided into groups of 4-5 by teacher)

Personalized e-Learning platform

A platform that customizes learning paths and content for students based on their learning styles, progress, and preferences.
  - XX groups working on the topic (add here names of the students, divided into groups of 4-5 by teacher)

Project-based exercises
CU 1 related content

Start by defining the following (This task should not take more than 2 presentation slides)

Problem statement: Which problem is your application solving?
Goal: What is the Goal of your AI application? What is the application called? What are the success criteria of your solution?

Create a stakeholder map, which should define at least the usage stakeholders, the external stakeholders and the development stakeholders (use the CU1 material as guidance).

Consider how well the following ethical principles are fulfilled with your application, main focus should be on societal and environmental aspects:

Beneficence
Ensure the AI solution enhances individual well-being and societal good. Key areas to consider:

- Protection of fundamental rights: safeguard individual rights, reinforcing ethical integrity.
- Environmental sustainability: minimize environmental impact through sustainable practices.
- Justification: clearly define and link the system's purpose to tangible societal benefits.

Non-maleficence:

- Ensure that the problem identification and proposed solution do not harm individuals or communities.
- Carefully consider and monitor the system's impact on physical and mental well-being.
- Assess potential safety risks and strategies for mitigation.

Justice:
Evaluate the system's impact on institutions, democracy, and society.

Project-based exercises
CU 1 (Help slide)

Use this table to see more in details the possible considerations under each ethical principle, as combined by Morley et. al (2020)

Table 5 showing the connection between high-level ethical principles and tangible system requirements as adapted from the methodology outlined in Chapter II of the European Commission's "Ethics Guidelines for Trustworthy AI"

Principle	Beneficence	Non-Maleficence	Autonomy	Justice	Explicability
Requirements	**Stakeholder participation:** to develop systems that are trustworthy and support human flourishing, those who will be affected by the system should be consulted **Protection of fundamental rights** **Sustainable and environmentally friendly AI:** the system's supply chain should be assessed for resource usage and energy consumption **Justification:** the purpose for building the system must be clear and linked to a clear benefit—system's should not be built for the sake of it	**Resilience to attack and security:** AI systems should be protected against vulnerabilities that can allow them to be exploited by adversaries. **Fallback plan and general safety:** AI systems should have safeguards that enable a fallback plan in case of problems. **Accuracy:** for example, documentation that demonstrates evaluation of whether the system is properly classifying results. **Privacy and Data Protection:** AI systems should guarantee privacy and data protection throughout a system's entire lifecycle. **Reliability and Reproducibility:** does the system work the same way in a variety of different scenarios. **Quality and integrity of the data:** when data is gathered it may contain socially constructed biases, inaccuracies, errors and mistakes—this needs to be addressed **Social Impact:** the effects of system's on people's physical and mental wellbeing should be carefully considered and monitored	**Human agency:** users should be able to make informed autonomous decisions regarding AI systems **Human oversight:** may be achieved through governance mechanisms such as human-on-the-loop, human-in-the-loop, human-in-command	**Avoidance of unfair bias** **Accessibility and universal design** **Society and democracy:** the impact of the system on institutions, democracy and society at large should be considered **Auditability:** the enablement of the assessment of algorithms, data and design processes. **Minimisation and reporting of negative impacts:** measures should be taken to identify, assess, document, minimise and respond to potential negative impacts of AI systems **Trade-offs:** when trade-offs between requirements are necessary, a process should be put in place to explicitly acknowledge the trade-off, and evaluate it transparently **Redress:** mechanism should be in place to respond when things go wrong	**Traceability:** the data sets and the processes that yield the AI system's decision should be documented **Explainability:** the ability to explain both the technical processes of an AI system and the related human decisions **Interpretability**

SOURCE | Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. Science and Engineering Ethics, 26(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Project-based exercises
CUs related content

CU 2
Privacy Concerns

- Identify the privacy concerns associated with the solution.
- Consider how data is collected, stored, and used.
- Suggest measures to protect user privacy and ensure data security.

CU 3
Algorithm Used

- Describe the algorithm that powers the solution.
- Explain how the algorithm is built and trained based on the data collected (cu2).
- Discuss any specific technical details that are important for understanding the solution.

CU 4
Possible Biases

- Identify potential biases in the system.
- Discuss what could cause these biases (e.g., training data, algorithm design).
- Suggest ways to prevent or mitigate these biases.

CU 5

- TU delft has created a checklist for ethical AI development. Check the tool via this link and utilize the AI APP project checklist in the pages 12-13 and fill in the checklist in the page 13. Present your checklist at the end of your project presentation.

- The final slide should present what you as a team learned and what new insights did you gain by doing this presentation.

THANK YOU

Project number: 2022-1-ES01-KA220-HED-000085257

Co-funded by the European Union