



# CS-E4740 - Federated Learning D, Lectures, 28.2.2024-29.5.2024

Started on	Thursday, 4 April 2024, 8:01 PM
State	Finished
Completed on	Thursday, 4 April 2024, 9:41 PM
Time taken	1 hour 40 mins
Grade	6.00 out of 11.00 (54.55%)

## Question 1

The European Commission set up the High-Level Expert Group on Artificial Intelligence (AI HLEG) in 2018. What requirement for trustworthy AI was **not** put forward?

Note: the requirements are rephrased. Pay attention to the meaning, not wording.

- ☒ a. A robot may not injure a human being or, through inaction, allow a human being to come to harm. 

The First Law of robotics by Isaac Asimov (1942).
- ☐ b. AI systems must reliably behave as intended while minimizing unintentional and unexpected harm, and preventing unacceptable harm.
- ☐ c. Accountability must be ensured for AI systems and their outcomes, both before and after their development, deployment and use.
- ☐ d. Prevention of harm to privacy and adequate data governance.

Your answer is correct.

The seven key requirements for trustworthy AI are listed in the Section 8.2 "Seven Key Requirements by the EU" in the Lecture Notes (link).

The correct answer is:

A robot may not injure a human being or, through inaction, allow a human being to come to harm.

## Question 2

What does the explainability of a trained personalized model imply?

- ☐ a. A human can replicate the model and perform the prediction delivery process.
- ☒ b. A human can guess the prediction delivered by the model with a certain level of accuracy.
- ☐ c. The model structure must be available to public.
- ☐ d. The model must replicate the acknowledged formulas.

Your answer is correct.

Explainability is subjective. It tells how well can a user anticipate (or guess) the prediction delivered by a hypothesis for a given data point.

See the corresponding Section 8.4 in the Lecture Notes (link).

The correct answer is:

A human can guess the prediction delivered by the model with a certain level of accuracy.

## Question 3

This question refers to **the student task #1** in the "Trustworthy AI" assignment.

In the student task #1 the perturbations follow the normal distribution. What type of relationship is between the **variances** of perturbations and the sum of the Euclidian distances between the weight vectors obtained for the original and perturbed data?

- ☐ a. There is no relationship between the specified variables.
- ☐ b. Positive linear relationship.
- ☒ c. Non-linear relationship.
- ☐ d. Negative linear relationship.

Your answer is correct.

Please, join Slack channel (link) if you have any questions regarding the coding assignment.

The correct answer is:

Non-linear relationship.

## Question 4

This question refers to **the student task #1** in the "Trustworthy AI" assignment.

In the student task #1 the perturbations follow the normal distribution. What type of relationship is between the **means** of perturbations and the sum of the Euclidian distances between the weight vectors obtained for the original and perturbed data?

- ☐ a. Positive linear relationship.
- ☐ b. Negative linear relationship.
- ☐ c. There is no relationship between the specified variables.
- ☒ d. Non-linear relationship.

Your answer is incorrect.

Please, join Slack channel (link) if you have any questions regarding the coding assignment.

The correct answer is:

Positive linear relationship.

## Question 5

This question refers to **the student task #1** in the "Trustworthy AI" assignment.

In the student task #1 the perturbations follow the normal distribution. Consider two cases:

1. the perturbations  $\sim \mathcal{N}(1, 10)$ , and
2. the perturbations  $\sim \mathcal{N}(10, 1)$ .

For what case does the the sum of the Euclidian distances between the weight vectors obtained for the original and perturbed data **is larger**?

- ☐ a. For both cases the sum Euclidian distances is the same.
- ☒ b. 1
- ☐ c. 2

Your answer is incorrect.

Please, join Slack channel (link) if you have any questions regarding the coding assignment.

The correct answer is:

2

Finish review

### Previous activity

◀ "Graph Learning"

### Next activity

"Privacy in FL" ▶



### Tuki / Support

#### Opiskelijoille / Students

- MyCourses instructions for students
- email: mycourses(at)aalto.fi

#### Opettajille / Teachers

- MyCourses help
- MyTeaching Support form

### Palvelusta

- MyCourses rekisteriseloste
- Tietosuoja-ilmoitus
- Palvelukuvaus
- Saavutettavuusseloste

### About service

- MyCourses protection of privacy
- Privacy notice
- Service description
- Accessibility summary

### Service

- MyCourses registerbeskrivning
- Dataskyddsmeddelande
- Beskrivning av tjänsten
- Sammanfattning av tillgängligheten

