

# Introduction to Artificial Intelligence

Summer 2022

## Exercise 4: Regression and Classification

Consultations: 20th April 2022

Assessment: 27th April 2022 (Room No. 009)

Teacher: Khanday Owais Mujtaba, [owais.khanday@pw.edu.pl](mailto:owais.khanday@pw.edu.pl) Group 101

Exercise should be done in pairs.

### 1. Exercise details

Write a program that predicts whether a packet is a DDoS attack packet or normal packet. The dataset can be downloaded from the link below:

<https://www.unb.ca/cic/datasets/nsl.html>

Use at least two separate methods and compare them with each other. Prepare a report with comparison between the methods used and explanation which one is the best method.

Delivered solutions will be assessed by results of chosen models and correctness of research methods.

- 1) The data should be properly divided and preprocessed
- 2) Train your model on the train+ file and test it on test+ file.
- 3) Training process should contain hyperparameter optimization
- 4) Interpretation of achieved results and final conclusions will also be assessed.
- 5) Remember to adhere to basic standards of clean coding.

### 2. Technical details

- a. The preferred language to write your solutions is Python.

### 3. Handing-in guidelines

- a. A brief report with methods comparison should be sent to my e-mail before the classes. Please include "[EARIN] Exercise 4" in the title, and do not forget about adding names and emails of both team members in the e-mail content.
- b. You may get 0-8 pts for this assignment.

**Exercise 4: Regression and Classification**  
**Consultations: 22th April 2022**  
**Assessment: 29th April 2022 (Room No. 009)**

**Teacher: Khanday Owais Mujtaba, [owais.khanday@pw.edu.pl](mailto:owais.khanday@pw.edu.pl) Group 102**

Exercise should be done in pairs.

**1. Exercise details**

Write a program that predicts whether a packet is a DDoS attack packet or normal packet. The dataset can be downloaded from the link below:

<http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>

Use at least two separate methods and compare them with each other. Prepare a report with comparison between the methods used and explanation which one is the best method.

Delivered solutions will be assessed by results of chosen models and correctness of research methods.

- 1) The data should be properly divided and preprocessed
- 2) Train your model on the kddcup.data file and test it on corrected file
- 3) Training process should contain hyperparameter optimization
- 4) Interpretation of achieved results and final conclusions will also be assessed.
- 5) Remember to adhere to basic standards of clean coding.

**2. Technical details**

- a. The preferred language to write your solutions is Python.

**3. Handing-in guidelines**

- a. A brief report with methods comparison should be sent to my e-mail before the classes. Please include "[E4RIN] Exercise 4" in the title, and do not forget about adding names and emails of both team members in the e-mail content.
- b. You may get 0-8 pts for this assignment.