Machine Learning – Introduction

winter 2022/23

Mateusz Krawczyk

mateusz.krawczyk@pwr.edu.pl

WUST, WPPT, Department of Theoretical Physics

503, A1

Organisation

- Consultations : on request
- Grading:
  - Two tests based on the knowledge from the lecture and numerical skills – 1x20pts + 1x25pts
  - Whole semester project – 50pts – groups of 3
  - Activity 20pts – live–coding in front of the entire group.
  - The earned points will be translated into the final grade according to:
    - 51 – 60pts: 3.0
    - 61 – 70pts: 3.5
    - 71 – 80pts: 4.0
    - 81 – 90pts: 4.5
    - 91 – 98pts: 5.0
    - 99 – more?: 5.5
- Possibility of one test revision.

In-Class Tests
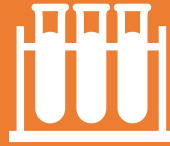
There will be two such tests.

The first one devoted to data preprocessing in numpy and pandas

The second for verifying the knowledge of ML algorithms

Projects

- Groups of maximally 3 people – the worksheet will appear soon ;)

- Basic idea - choose data of your interest and analyze it

- Apply different machine learning algorithms (apart from neural networks – that's the topic of next semester)

Projects Report

Proposal – brief (1pg max!) **report + 5 min presentation** where you describe what are you going to do:

- Motivation: What problem are you tackling and why is it interesting for you? Is it an application or a theoretical research?
- Method: What machine learning techniques are you planning to apply or improve upon?
- Intended experiments: What experiments are you planning to run? How do you plan to evaluate your machine learning algorithm?

Final report + 15 min presentation

- Brief topic description (What? Why?)
- Methods description
- Results
- Conclusion -> What worked best? Why?
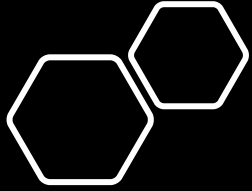
Project criteria and hints

The following criteria will be taken into account when grading:

- Clarity/Relevance of problem statement and description of approach.

- Methods choice and applicability to the given problem

- Discussion of relationship to previous work and references.

- Design and execution of experiments.

- Figures/Tables/Writing: easily readable, properly labelled, informative.

Projects Deadlines

- The one page **project proposal (additional 5 pts.)** – PDF uploaded online

    by noon on **November 17th**.

- The first **presentation (10 pts.) – 5 + 5 min**

    November 17th

- The **project report (20 pts.)** - PDF file or Jupyter notebook

    by noon on **January 19th**.

- The final **presentation (20 pts.) – 15 + 5 min**

    January 19th/26th

Schedule – first draft

1. 2022-10-06 – Introduction
2. 2022-10-13 – Fundamental Python libraries for modern Machine Learning – pandas and numpy
3. 2022-10-20 – Fundamental Python libraries for modern Machine Learning – matplotlib and scikit-learn
4. 2022-10-27 – Test 1 (Python programming – we'll see how it goes)
5. 2022-11-03 – Perceptron and linear regression
6. 2022-11-10 – Logistic regression
7. 2022-11-17 – Project ideas – present your ideas! - last day you can send the project proposal
8. 2022-11-24 – Support Vector Machine
9. 2022-12-01 – Kernelization, Ridge regression
10. 2022-12-08 – Similarity: k–Nearest Neighbors
11. 2022-12-15 – Similarity: k-means clustering
12. 2022-01-22 – Dimensionality reduction
13. 2022-01-12 – Test 2 (Machine Learning methods – up to the progress with the lecture)
14. 2022-01-19 – Presentation day - last day of sending the project report (pdf or Jupyter file with description)
15. 2022-01-26 – Extra time (presentations/revisions)

Workstyle

We need to figure out the way the tutorials will be conducted.
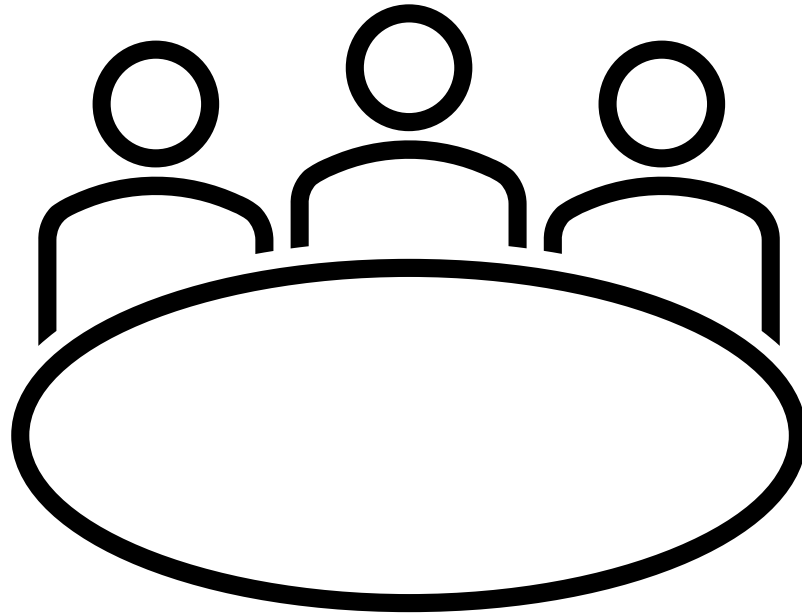
live coding | Jupyters | other???

GitHub | Eportal | other ???