Implementation project
Lab Machine Learning for Data Science
Summer Semester 2024

G. Montavon
Institute of Computer Science
Department of Mathematics and Computer Science
Freie Universität Berlin

Lab ML for Data Science: **Organization of the Course**

1 Structure of the Course

- The course takes place every Fridays from 2:15pm to 3:45pm in room T9/K 036 (Takustr. 9). Further seats are available in room T9/K 044 and T9/K 046.
- The implementation project is divided in three parts: (i) insights from unsupervised data with a use case in customer analysis, (ii) insights from supervised data with a use case in chemistry, and (iii) insights from supervised data and pretrained deep neural networks with a use case in image analysis.
- The project is carried in groups. Your implementation of the different project parts needs to be submitted over the semester (cf. timeline). The project outcomes are then presented at the end of the semester in a 30 minutes presentation (+ questions).
- Software used: python/numpy, Jupyter notebooks, scikit-learn, matplotlib, pytorch, torchvision, various libraries for data loading and visualization.

2 Timeline (tentative)

19 Apr	**Kick-off meeting** (in presence)
26 Apr	**Deadline for group formation**
3 May	Implementation of Part I (in presence)
10 May	Implementation of Part I (in presence)
17 May	Implementation of Part I (in presence)
21 May	**Submission of Part I**
24 May	Implementation of Part II (in presence)
31 May	Implementation of Part II (in presence)
7 Jun	Implementation of Part II (in presence)
11 Jun	**Submission of Part II**
14 Jun	Implementation of Part III (in presence)
21 Jun	Implementation of Part III (in presence)
28 Jun	Implementation of Part III (in presence)
2 Jul	**Submission of Part III**
12 Jul	**Presentation day** (in presence)

3 Group Formation

- Groups can be formed directly at the kick-off meeting, or later using the course's discussion forum.
- Send your group details (names + student ID of members) to `gregoire.montavon@fu-berlin.de`.

4 Submission of Project Parts

- Generate a PDF printout of your Jupyter notebook including the output (use restart & re-run to make sure the results correspond to the actual code).
- Additionally, include a PDF printout of additional Python file you have written and that you have imported from your Jupyter notebook, and place them in a zip file.
- Submit both files via Whiteboard.

5 Submission of the Final Presentation's Slides

- Generate a PDF printout of your slides.
- Submit the PDF file via Whiteboard one day before your presentation.

6 Final Presentation

- The final presentation for a group takes 45 minutes overall and has the following format:
  - 30 minutes: actual presentations (speaking time uniformly split between the group members)
  - 15 minutes: questions

7 Computation of the Grade

- The course's grade is determined solely based on the final presentation. Criteria for determining the grade include the technical quality of the implementation, the scientific quality of the methodology and analysis, and the overall clarity of the presentation.

8 Attendance Requirements

- Attending the kick-off meeting is required. It is recommended that you attend at least two implementation sessions for each part of the project.