

Marcin Szpyrka

Katedra Informatyki Stosowanej AGH

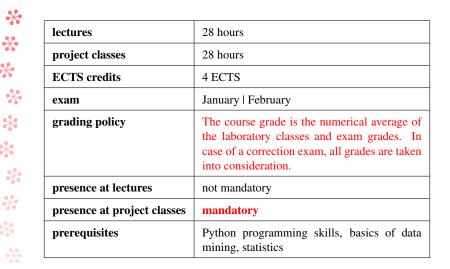
2024/2025

*

Selected information from syllabus

*

*



General project implementation requirements

Group	Number of students	Number of project teams
AID	29	9-10 (at most 2 teams with 4 students)
GSI	16	5-6 (at most 1 team with 4 students)
IOP	13	4-5 (at most 1 team with 4 students)

Teams should be defined before the first project classes!

General area of project topics: text mining.

Expected milestones:

**

- 1. Data acquisition (text data) Don't use datasets downloaded from the internet!
- 2. Data cleaning
- 3. Feature generation
- 4. Model(s) training
- 5. Results analysis
- 6. Final report preparation

Project classes – general rules

*

×

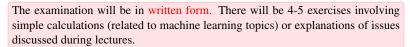
*

- The dataset must have at least 5000 rows or 10000 if you want a grade higher than 4.
- The work with data is realized using the Python language. A Python notebook must be included as a part of the final report.
- Each team must present a progress report at least once every two weeks.
- The final report must be prepared using LaTeX or as a Python notebook (markdown cells included into code cells). The notebook (either an attachment or the final report) must be provided as the ipynb file and exported to the HTML format.
- The final report must be provided at the latest during the last project classes (last week of January). Sending your project a few days in advance would be appreciated
- All team members are expected to be familiar with the entire contents of the final report. During the delivery of the report there will be a discussion!

Exam – competition

**

**



The main examination date is scheduled for the first week of the session. Admission to the examination is subject to the delivery of a successfully assessed project.

Projects competition

During the last lecture there will be presentations of the top 3 projects (20 minutes for presentation + 10 minutes for discussion). The members of these teams will obtain a passing grade in the exam based on this presentation. The competition will be open to projects submitted to me by mid-January.