

# Task II: Social Media Analytics Project

Prof. Dr. Tillmann Schwörer

## 1. Project Overview

- **Task:** Carry out a small but complete social media analytics project related to topics covered in the course. Focus on 1 (!) core analytical goal, which may well consist of several sub goals or questions. Students are allowed to exchange ideas, but this is not a teamwork exercise.
- **Submissions:** Submit all that is needed to fully reproduce your work (notebook, scripts, data, etc.) on Moodle, or submit a link to a public GitHub repository. Note that you must make an authorship statement upon submission saying that the project is your own work and that you have cited external sources. Submission deadline is 10.01.2023, 23:59.

## 2. Detailed Task Description

Conduct a Social Media Analytics project on a real-world dataset:

- **Goal:** Describe a real or fictitious scenario for your project, explaining its goal and purpose.
- **Data:** You can either use a ready-made data set or assemble your own data set via APIs or web scraping. If you are using a ready-made data set, I expect a larger scope or depth on the analytical side.
- **Pipeline:** Include all necessary steps of your project: data acquisition, exploration, preprocessing, modeling, evaluation, conclusion. Start from a trivial baseline. Then optimize your results by using e.g. alternative preprocessing steps, machine or deep learning algorithms, parameter settings.
- **Explanation:** Explain your approach and your steps, and interpret your results. If your main analytical task is unsupervised, you must be particularly clear about the criteria according to which you evaluate the results, and demonstrate the usefulness and/or limitations of your results using appropriate examples.
- **Language:** Allowed languages are English and German.
- **Code:** You may use all the code from the lectures. Copying and adapting from other sources is allowed in small quantities. Copying code in large quantities will be treated as intent to deceive and result in a score of zero points. Cite all relevant resources on which your project is based, or from which you draw inspiration.

## 3. Grading

- **Data acquisition and analytical content (60 points):** Here I will evaluate whether your overall approach and your steps are adequate and correct, and whether you draw the right conclusions from your results. If you are using a ready-made dataset, I expect a larger scope or depth on the analytical side. Also, I will evaluate the overall complexity and innovativeness of your project. Innovativeness and complexity can arise e.g. from the acquisition of data, a novel and insightful combination of known techniques, or the use of techniques not explicitly covered in class. Particularly innovative or complex projects may be awarded up to 10 bonus points.
- **Story (25 points):** Guide through your project using a clear language. Explain the main goal of the project, and what the value of achieving this goal is. Explain your overall approach. Describe your data and their properties. Explain what your steps are, why they are important, and what insights you draw from them. In your conclusion, summarise your key results in terms of the

stated goal, and reflect on your personal learnings, limitations of your project and used techniques, and possible directions for future work. Note that an honest assessment of existing limitations is considered a strength and will be honoured.

- **Form (15 points):** Make sure that your submission is well structured and clean: remove non-essential files, code, and explanations; make sure that your code is free of debug output or error messages; check spelling. Provide structure to your notebook via section headers, paragraphs, text highlighting, etc. Respect [Python's style guide](#) and avoid redundant code via e.g. functions or loops. It may make sense to move function definitions to a separate script.