

Task 2

Individual Project

Task Description

- **Goal:** Carry out a small, but complete social media analytics project related to topics covered in the course (APIs, web scraping, network analysis, sentiment analysis, topic modeling, text classification, trend/text mining, etc.). Describe a real or fictitious scenario for your project, explaining its goal and purpose.
- **Data:** Your project must be based on a real-world data set. You can either use an existing data set or engineer your own data set via APIs, web scraping, and possibly annotating the data. The data context can be social media platforms or other forms of text data such as news articles, product reviews, etc. Hint: If you intend to carry out a text classification task, you either need a data set that comes with labels or spend some time on annotating data yourself.
- **Focus:** If your project has a large data engineering part (data acquisition, cleaning, annotation, etc.), then a compact analytical part is sufficient. If your project is based on an existing data set, then I expect a larger scope and/or depth on the analytical side (evaluation of multiple models and parameter choices).
- **Procedure:** Include all relevant steps of your project into the Jupyter notebook and guide the reader through the entire project in words. In particular, make sure to carefully evaluate the results and comment on the insights you gained.
- **Language:** Allowed languages are English and German.
- **Resources:** 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.
- **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.

Grading

Due to the different nature of the projects, the grading will be based on a holistic evaluation of the project. The following aspects will be considered:

- Data engineering efforts
- Correctness of the approach and the code
- Thorough evaluation and correct interpretation of the results
- Scope, complexity and innovativeness of the project
- Storytelling: purpose of the project, insights, reflection on learnings and limitations
- Form: well-structured, concise and clean submission