Computer Vision & Natural Language Processing

Technikum Wien

Bachelor Computer Science

Python exercise sheet C

This exercise sheet uses the corpus of plot summaries provided on the Moodle page of the course.

Exercise C.1: Preprocessing

- 1. Preprocess the data.
- 2. Use LDA on the data.
- 3. Perform hyperparameter tuning for the LDA model by varying the number of topics.

Hint: The Python package 'gensim.models' provides the 'LdaModel' function.