Assignment: Recommendation engine

Image: FRacco, via Wikimedia Commons, under CC BY-SA 3.0

Learning goals

In this assignment, you:

1. learn to apply a recommendation engine to a data set.
2. explore the properties and limits of a recommendation engine implementation.

Assignment

Load the public domain Anime dataset either from the original location (https://www.kaggle.com/CooperUnion/anime-recommendations-database/version/1) or from the **Methods/Data/Anime** folder in the course's Oma workspace.

In either case, get acquainted with the data description at the website.

This assignment is of exploratory nature. Your task is to explore the applicability of **scikit-surprise** in building a recommendation engine for the Anime dataset.

The questions of interest include:

1. What kind of preprocessing is necessary for the ratings dataset?
2. How do the recommendation algorithms (e.g. KNN and SVD) perform with a data set of this magnitude? Do you encounter hardware limitations? If yes, how can you circumvent some of the limitations to be able to carry on with the experiment?
3. Can you combine the information in the two files in a meaningful way to have the recommender display the titles of the recommended movies?

Deliverables

In this assignment, aim at sharing your observations and experiences from **scikit-surprise**.

Your deliverable should include both the Python codes and the results needed to verify the conclusions.

Submit your work preferably in pdf format. The deliverable should contain the information specified in the points 1 to 3 above.