Advanced Probabilistic Machine Learning and Applications

Martina Contisciani and Caterina De Bacco

July 9, 2021

Tutorial 11: Matrix Factorization for Recommendation Systems

In this tutorial we will analyse the dataset **MovieLens 100K** with different Matrix Factorization models.

We will use the package *surprise* that can be found at https://surprise.readthedocs.io/en/stable/index.html.

(a) Install the package **surprise**.

(b) Download the dataset **MovieLens 100K** from https://grouplens.org/datasets/movielens/100k/.

(c) Import the dataset and make some explorative analysis.

(d) Pick a model and run the *GridSearchCV* class to choose the best number of factors.

(e) Perform a Cross-Validation routine to choose the best model among SVD, SVD biased, NMF, and NMF biased.

(f) Choose the best model and use it to make predictions and recommendations.

(g) Fill the function *get_top_n* to get the first n recommendations for each user from a set of predictions.