

Advanced Probabilistic Machine Learning and Applications

Diego Baptista and Caterina De Bacco

January 21, 2022

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 *Singular Value Decomposition* (SVD), SVD biased, *Nonnegative Matrix Factorization* (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.