

## Submission #2: Matrix Factorization And End-To-End Evaluation

Total Points: 40 (30 mandatory, 10 additional)

Submission Structure:

- link to the repository
- branch and commit hash for evaluation
- pdf report [optional in case of repository is self-explanatory with clear readme files]

### Mandatory Tasks (30 points)

#### Matrix Factorization Algorithm (21 points)

##### Objective

Implement and evaluate three different matrix factorization techniques. Evaluate each model using an offline evaluation framework developed during previous homework. If the framework needs changes or adaptations, all models should be reevaluated.

##### Deliverables

- Alternating Least Squares (7 points)
- FunkSVD (7 points)
- Choice of one (7 points) - SVD Feature or any Deep Learning solution

#### A/B Testing Framework (9 points)

##### Objective

Outline the methodology for A/B testing of recommender algorithms on your dataset. You should make assumptions about the online system characteristics in case

##### Deliverables

Markdown or pdf file with a defined framework for comparing recommender algorithms on your dataset. This document should go over metrics selection, statistical testing approach, experiment design (control/treatment split, sample size calculation based on the effect size, statistical power, and significance level), and the methodology for decision-making. In addition to the framework definition, you should provide at least two detailed walkthrough examples.

### Additional Task (10 points)

#### Multi-Armed Bandit Real-Time Simulation (10 points)

##### Objective

Formulate a problem for your dataset and solve it using multi-armed bandits. You can use any type of external data and you are not limited to only creating a recommender algorithm

using a multi-armed bandit approach. The only limitation is that the system should be applicable and somehow connected to the dataset you chose.

#### Deliverables

Jupyter Notebook or Python script, which allows to run the simulation, track the results online, and compare strategies. (e.g., epsilon-greedy vs UCB and/or Thompson Sampling)