

→ **Adapting Ratings in Memory-Based Collaborative Filtering using Linear Regression**

**DAI-Labor of
Technische
Universität Berlin**

Jérôme Kunegis
kunegis@dai-labor.de

Agenda

Collaborative Filtering

⇒ **Collaborative Filtering: Rating Prediction**

Baseline Algorithm

⇒ **Baseline Algorithm: Pearson Correlation**

Linear Regression

⇒ **Linear Regression: Adapting Ratings**

Evaluation

⇒ **Evaluation**

The Bipartite Rating Graph

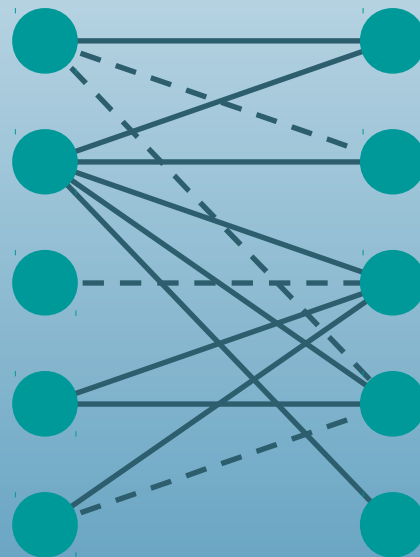
⇒ Database of user and items

⇒ Users rate items

⇒ Rating database



Users **Items**



Rating values

————— > 0

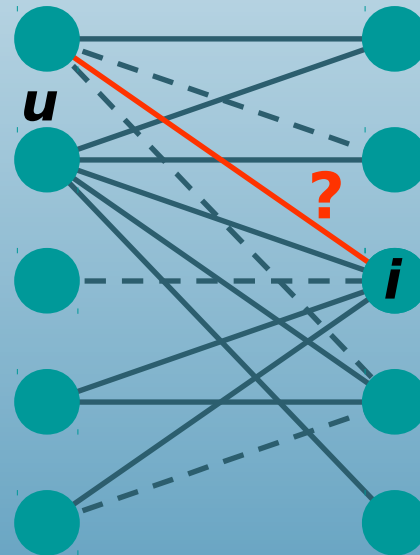
- - - - - < 0

- Collaborative Filtering
- Baseline Algorithm
- Linear Regression
- Evaluation

Collaborative Rating Prediction

- ⇒ Predict a missing rating
- ⇒ Use: Recommend new items (rank unrated items by predicted rating)

Users **Items**



Will user u like item i ?

Collaborative Filtering
Baseline Algorithm
Linear Regression
Evaluation

User Rating Habits and Taste

Collaborative Filtering

Baseline Algorithm

Linear Regression

Evaluation

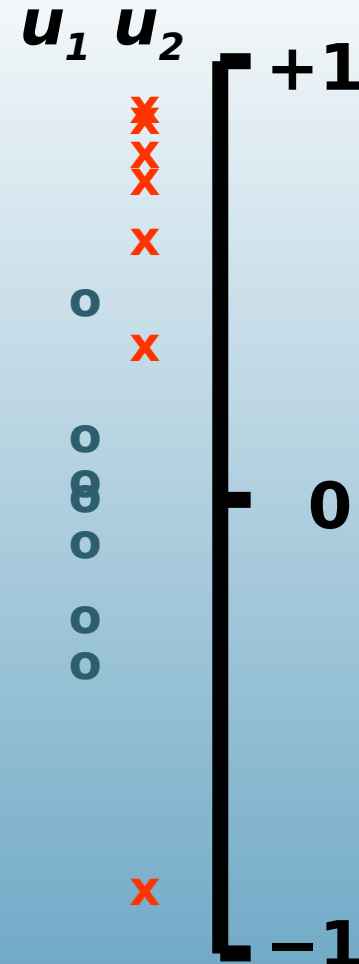
Users have different rating habits:

⇒ **Different mean rating**

⇒ **Different variance**

Weighted mean assumes the same rating scale for all users, although the Pearson correlation takes into account different scales

Solution: Weighted means of ratings scaled according to each user's rating mean and variance



Scaling Ratings

Collaborative Filtering

Baseline Algorithm

Linear Regression

Evaluation

Two Users u and v have common rating vectors U and V using linear regression

- ⇒ Determine factors a and b minimizing $\sum (U_i - (V_i + a + b))^2$
- ⇒ When U and V are negatively correlated, a is negative
- ⇒ Therefore, use absolute value of correlation for weight
- ⇒ Correlation and $(a \ b)$ can be calculated in one pass

Evaluation

Collaborative
Filtering

Baseline
Algorithm

Linear
Regression

Evaluation

Linear Regression adaptation can be used in variants of weighted mean algorithm:

- ⇒ **User-based, item-based**
- ⇒ **With/out normalized ratings**
- ⇒ **Fill missing ratings with default parameter**
- ⇒ **Weight users by number of common ratings ($1, n, n^2$)**

Evaluation Results

Collaborative Filtering

Baseline Algorithm

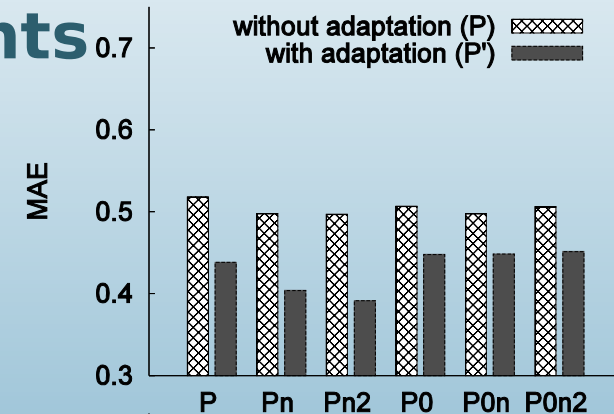
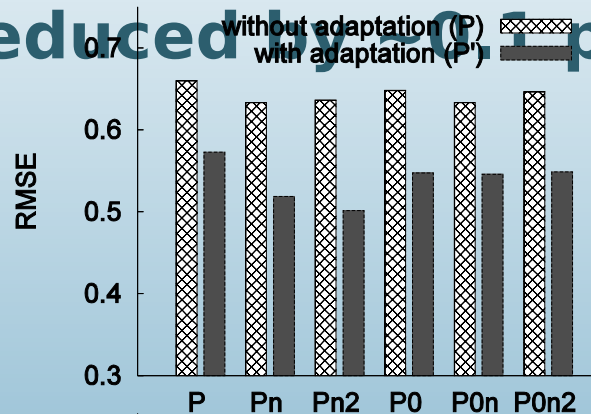
Linear Regression

Evaluation

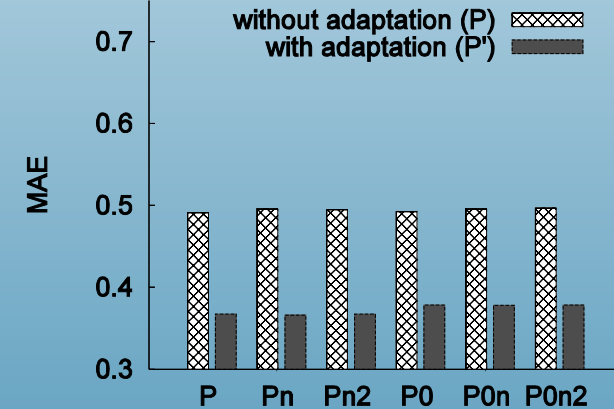
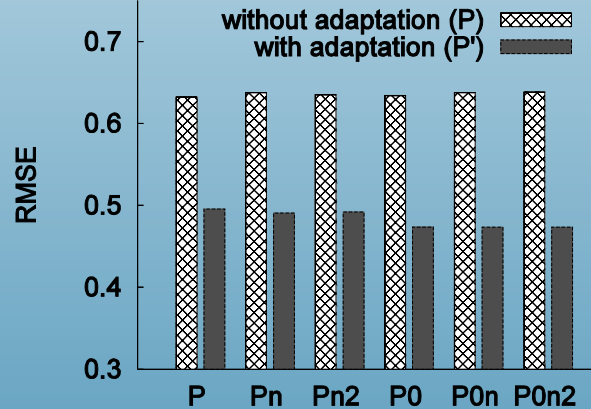
Both error measures (mean average error, mean root squared error)

reduced by 0.1 points

MovieLens



Jester



The End

Collaborative
Filtering

Baseline
Algorithm

Linear
Regression

Evaluation

Thank you!

Questions?

Comments?

Backup

Collaborative
Filtering

Baseline
Algorithm

Linear
Regression

Evaluation

- ⇒ **A: Why doesn't normalization make this obsolete?**
- ⇒ **Q: Normalization corrects differences in rating habits (scales), whereas regression maps different tastes to each other**