Assessing the Value of Unrated Items in Collaborative Filtering

Jérôme Kunegis
Andreas Lommatzsch
Martin Mehlitz
Şahin Albayrak

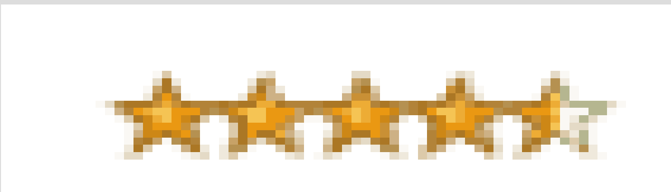
DAI-Labor, Technische Universität Berlin, Germany

2nd Int. Conf. on Digital Information Management

2007-10-29, Lyon

Rating Databases

- Database with user accounts and items: movies, songs, documents, …
- Users rate items, implicitely or explicitly



Sparse Rating Matrix

- A single user cannot rate all items:
  The rating matrix is sparse

	A	B	C	D	E
User 1	5	4	1		
User 2	1		2	5	5
User 3	??	2	1	??	5

Recommendation, Prediction

- Recommender System: Find items a user hasn’t yet seen
- Recommendation must be interesting: Predict the user’s taste
- Predict ratings between user and item

- Collaborative filtering: Based on ratings, not content

Weighted Mean of Ratings

- Rating prediction algorithm for user-item pair *(u,i)*:
- Take ratings of *i* by other users
- Calculate mean → bad because other users may have different taste
- Weight mean by correlation of users

- Use Pearson correlation [Resnick et al. 1994]

Using a Default Value

- Correlation usually calculated on item ratings *both* users have given.
- Variant: Use all items at least one user has rated. Fill missing ratings with a default value ρ
- Default value usually taken to be $\rho = 0$
- Runtime: slightly greater when using a default value due do extra operations when calculating the correlation

- Assumption: Unknown ratings are best represented by a neutral value of $\rho = 0$

Varying the Default Value

- Find out how the accuracy of the prediction varies in function of the default value ρ

Evaluation

Measuring the accuracy of rating prediction:

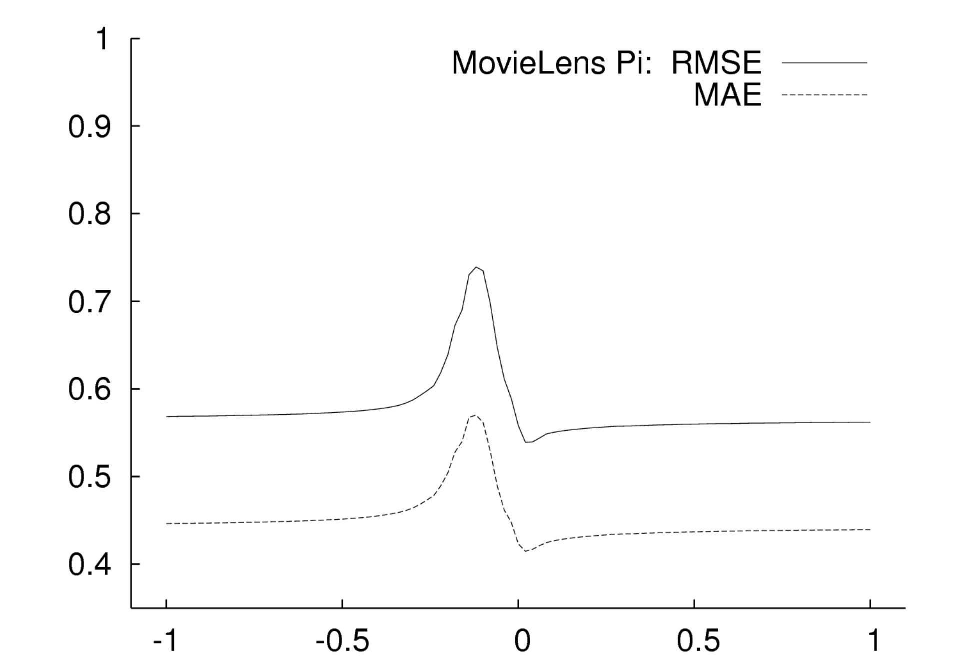
- Mean absolute error, root mean squared error

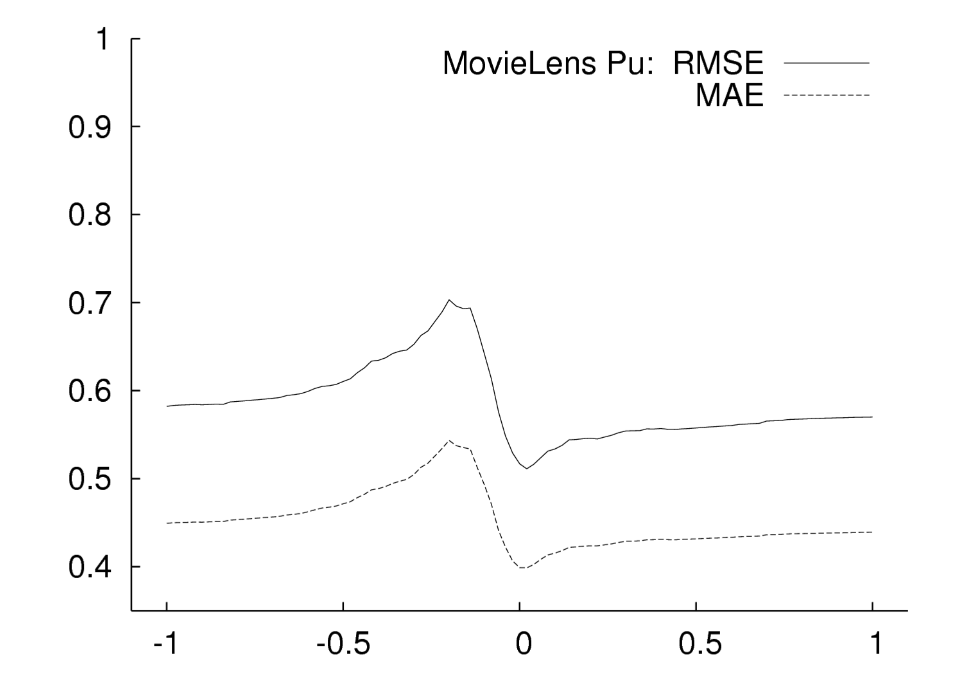
$$MAE = \frac{1}{n} \sum_i |r(u,i) - \tilde{r}(u,i)|$$

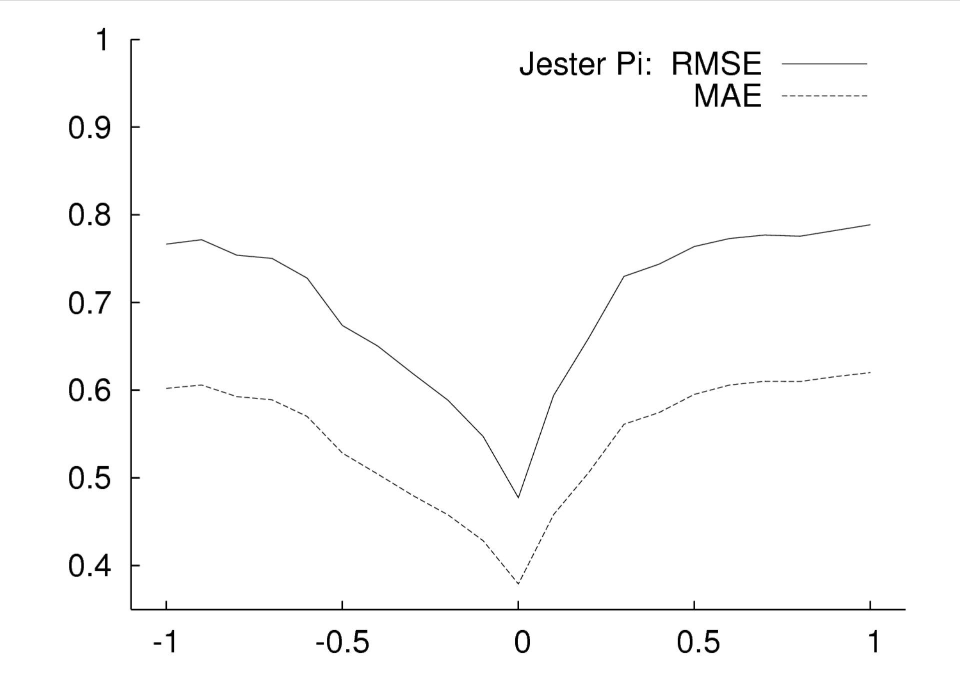
$$RMSE = \sqrt{\frac{1}{n} \sum_i (r(u,i) - \tilde{r}(u,i))^2}$$

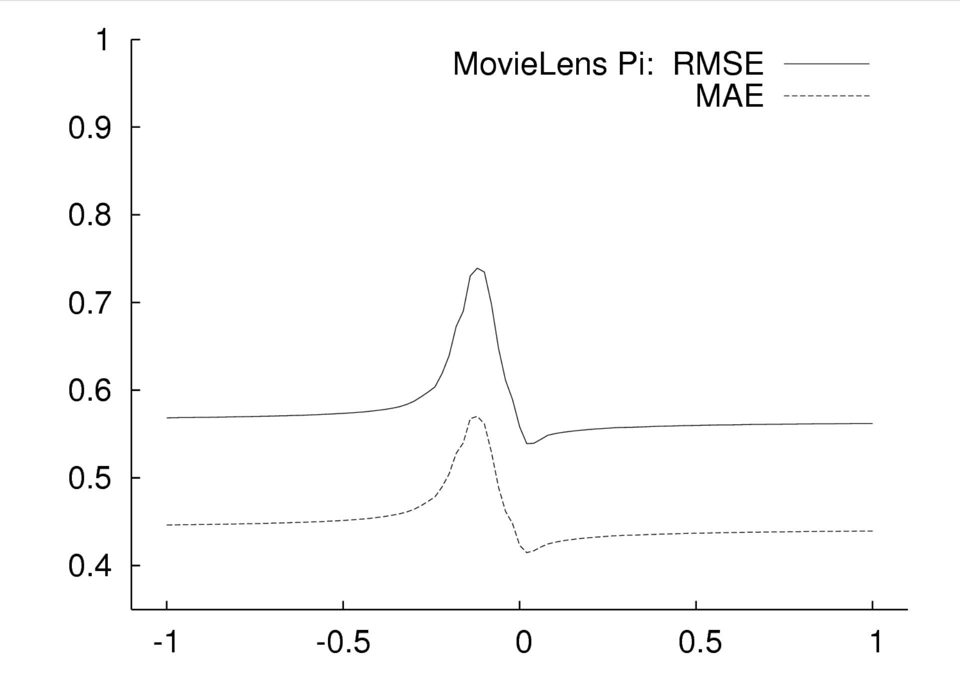
- MovieLens: 943 × 1,543; 75,000 ratings
- Jester: 24,900 × 100; 617,000 ratings

Basic Weighted Mean

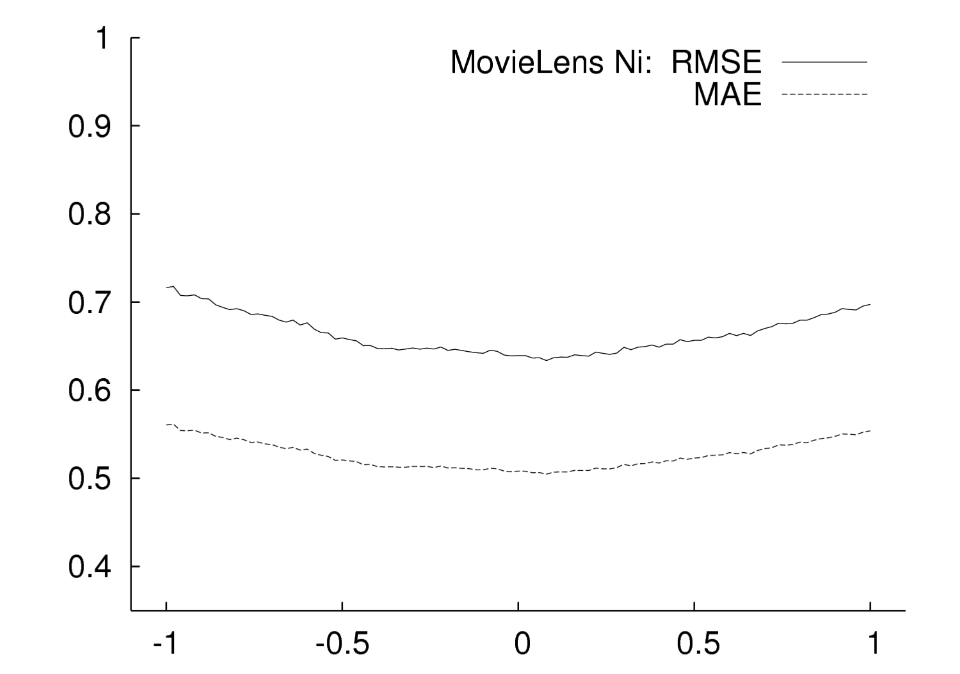


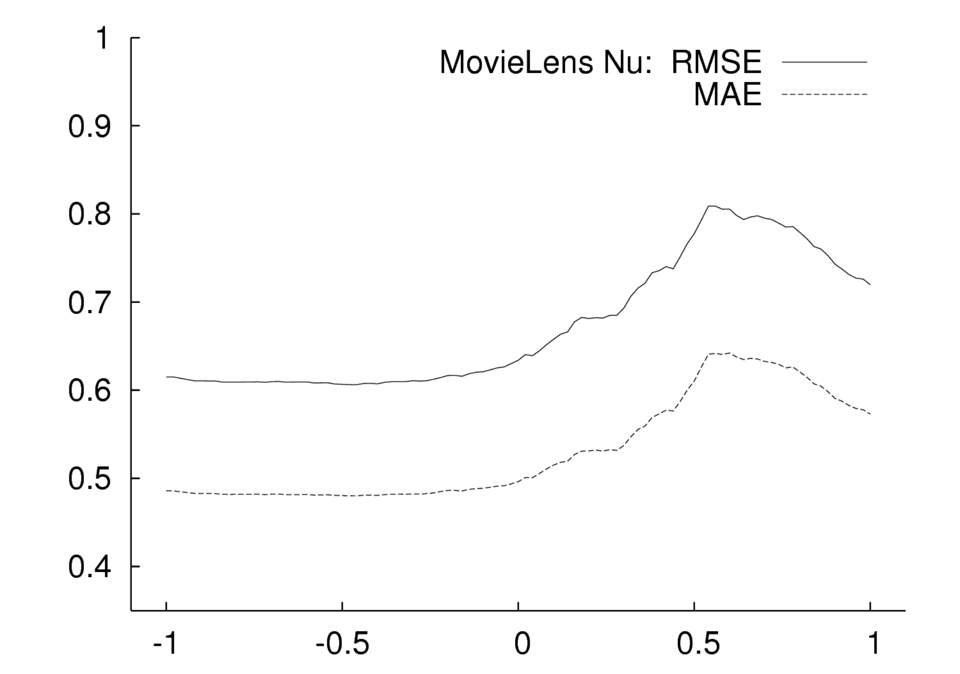


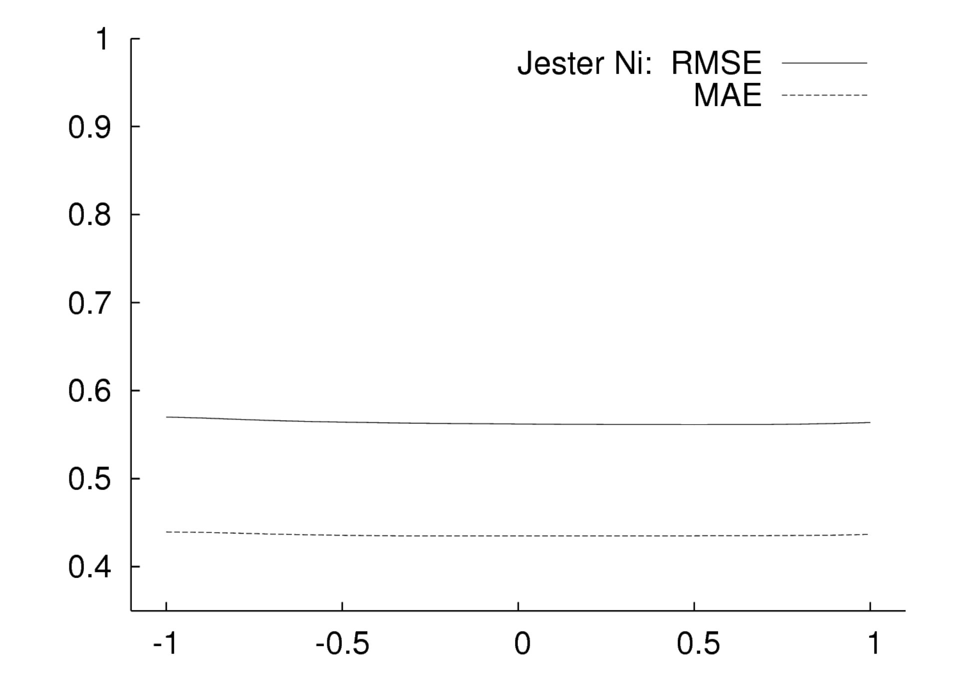


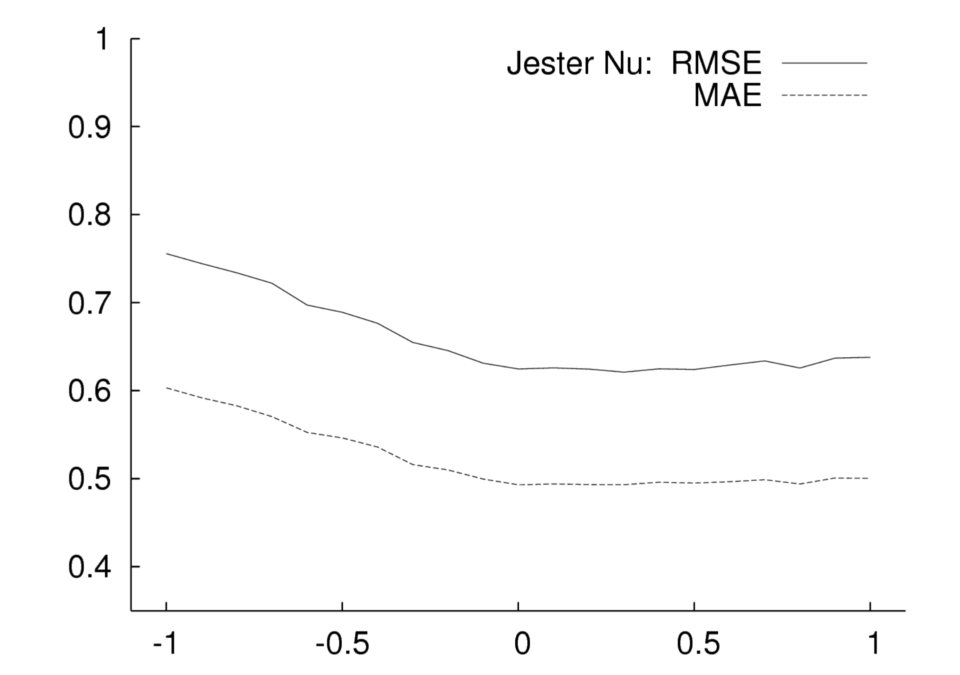


Normalized Weighted Mean

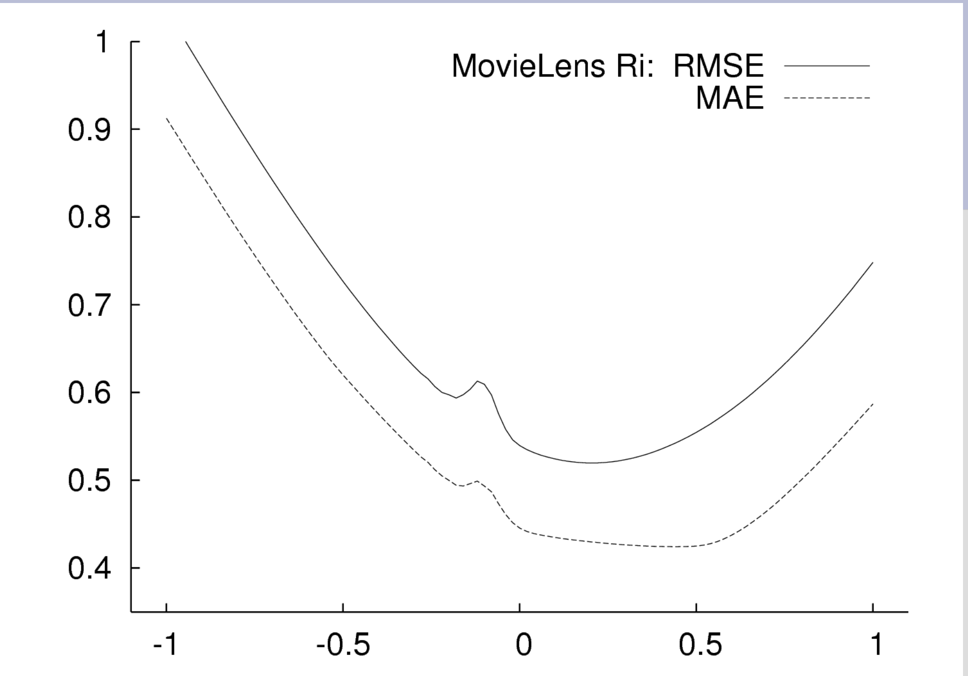


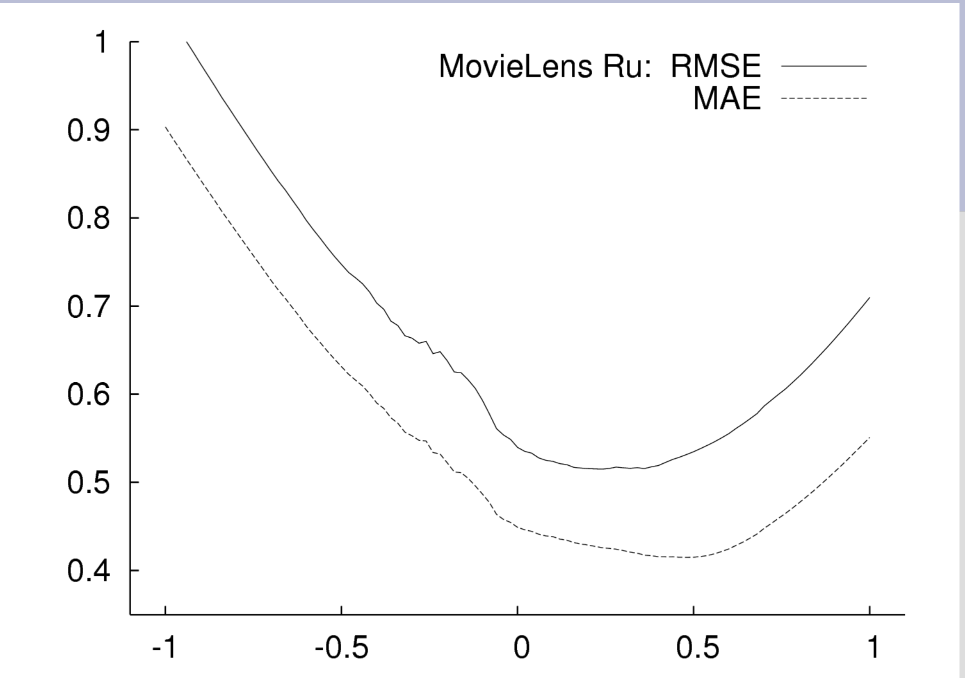


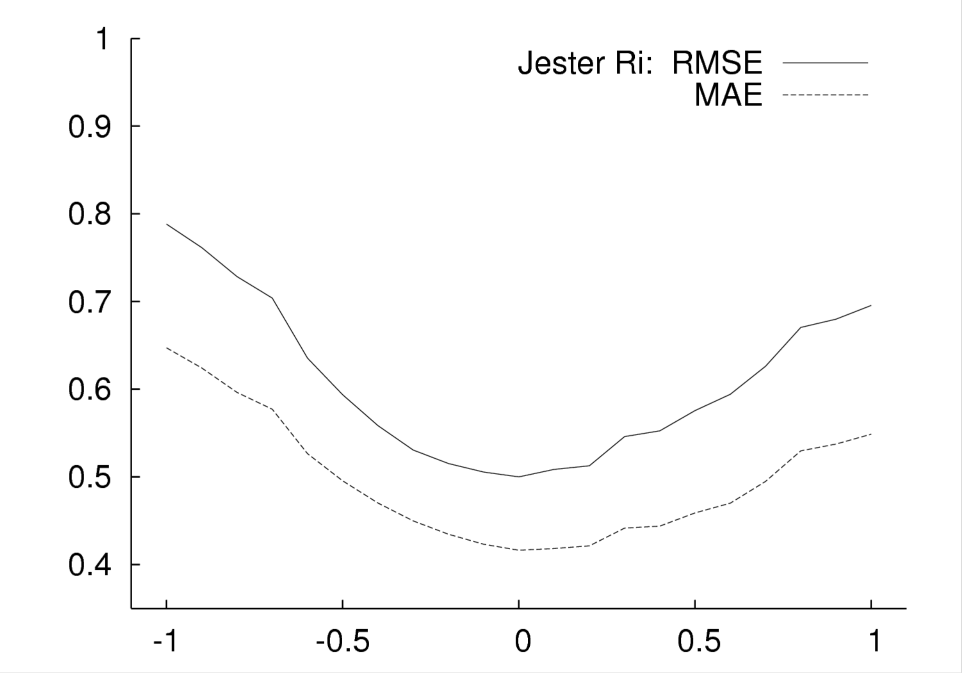


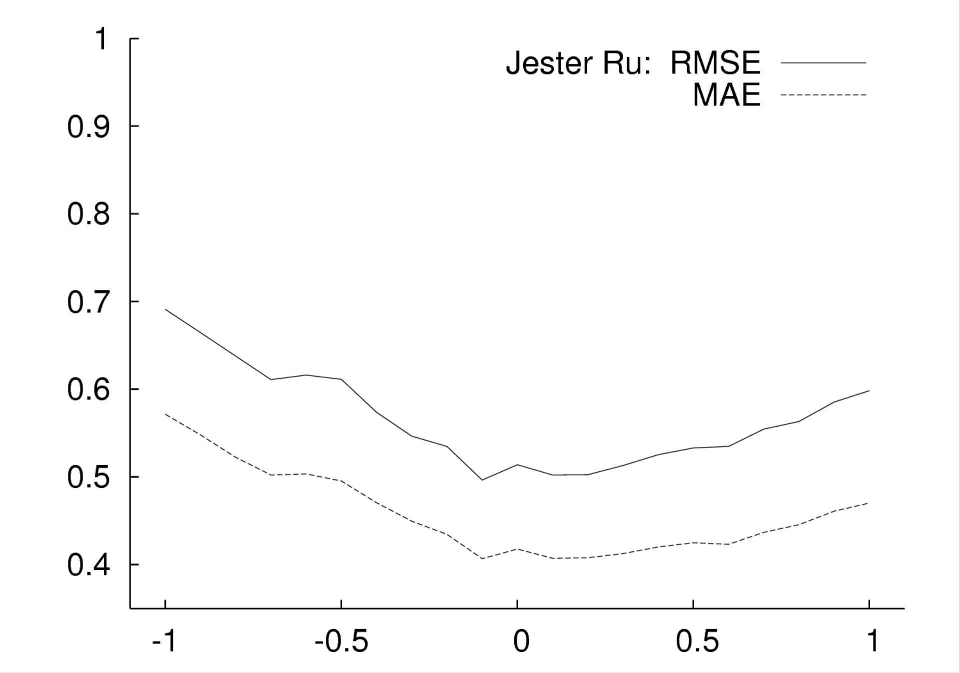


Normalized Weighted Mean









Special: Netflix Results

- Predict ratings in the Netflix corpus: ~100,000,000 Movie ratings by ~480,000 users for ~17,000 movies
- Numbers on a subset of the data
- Use normalized weight mean algorithm

ρ	−0.5	−0.4	−0.3	−0.2	−0.1	„0	„+0.1	„+0.2	„+0.3	„+0.4	„+0.5
RMSE	1.15	1.18	1.17	1.15	1.14	1.09	1.09	1.16	1.19	1.14	1.12
MAE	0.81	0.82	0.82	0.81	0.79	0.75	0.75	0.81	0.83	0.80	0.78

Conclusion

- Optimal default value depends on corpus
- Slight tendency towards positive values
- Contradicts the suggestion in [Herlocker 1999] to use slightly negative values
- Zero is a good value generally
- Using a default value better than not using one

Todo

- Add rating matrix example
- Add rating graph example
- Add formulas