Joshua Harkness

CAP5610 – Machine Learning

3/28/20

Homework 7

Link to code used for assignment: <https://github.com/darkhark/RecommenderSystems>

10. Fold 1

|  |  |  |
| --- | --- | --- |
|  | RMSE | MAE |
| SVD | 0.9480 | 0.7480 |
| PMF | 0.9599 | 0.7569 |
| NMF | 0.9729 | 0.7649 |
| User based Collaborative Filtering | 0.9862 | 0.7785 |
| Item based Collaborative Filtering | 0.9863 | 0.7835 |

The algorithm with the lowest RMSE value in the first fold was SVD, while the highest was Item based Collaborative Filtering. The same results applied to the MAE. Altogether, for RMSE we get SVD < PMF < NMF < User based < Item Based. For MAE, it is the same. This means that SVD possesses the least error.

11. Fold 2

|  |  |  |
| --- | --- | --- |
|  | RMSE | MAE |
| SVD | 0.9392 | 0.7420 |
| PMF | 0.9747 | 0.7675 |
| NMF | 0.9756 | 0.7657 |
| User based Collaborative Filtering | 0.9936 | 0.7859 |
| Item based Collaborative Filtering | 0.9891 | 0.7804 |

The results for fold 2 differ from fold 1, but only between User and Item based Collaborative Filtering for RMSE. MAE differs a bit more.

RMSE: SVD < PMF < NMF < Item based < User Based.

MAE: SVD < NMF < PMF < Item based < User based

12. Fold 3

|  |  |  |
| --- | --- | --- |
|  | RMSE | MAE |
| SVD | 0.9518 | 0.7510 |
| PMF | 0.9643 | 0.7632 |
| NMF | 0.9725 | 0.7631 |
| User based Collaborative Filtering | 0.9883 | 0.7817 |
| Item based Collaborative Filtering | 0.9814 | 0.7790 |

RMSE: SVD < PMF < NMF < Item based < User Based

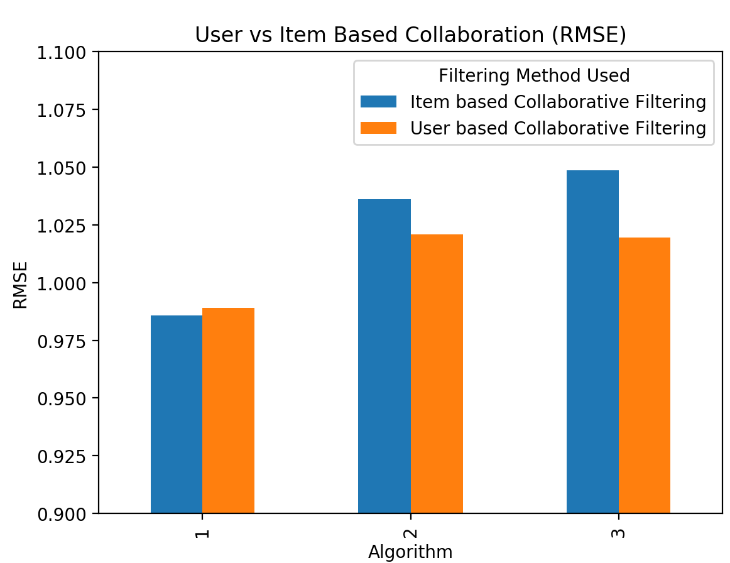
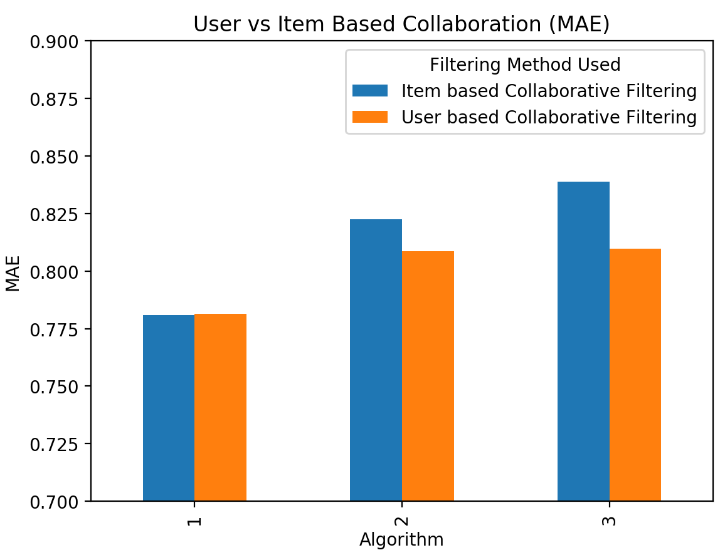
MAE: SVD < NMF < PMF < Item based < User based

13. Average

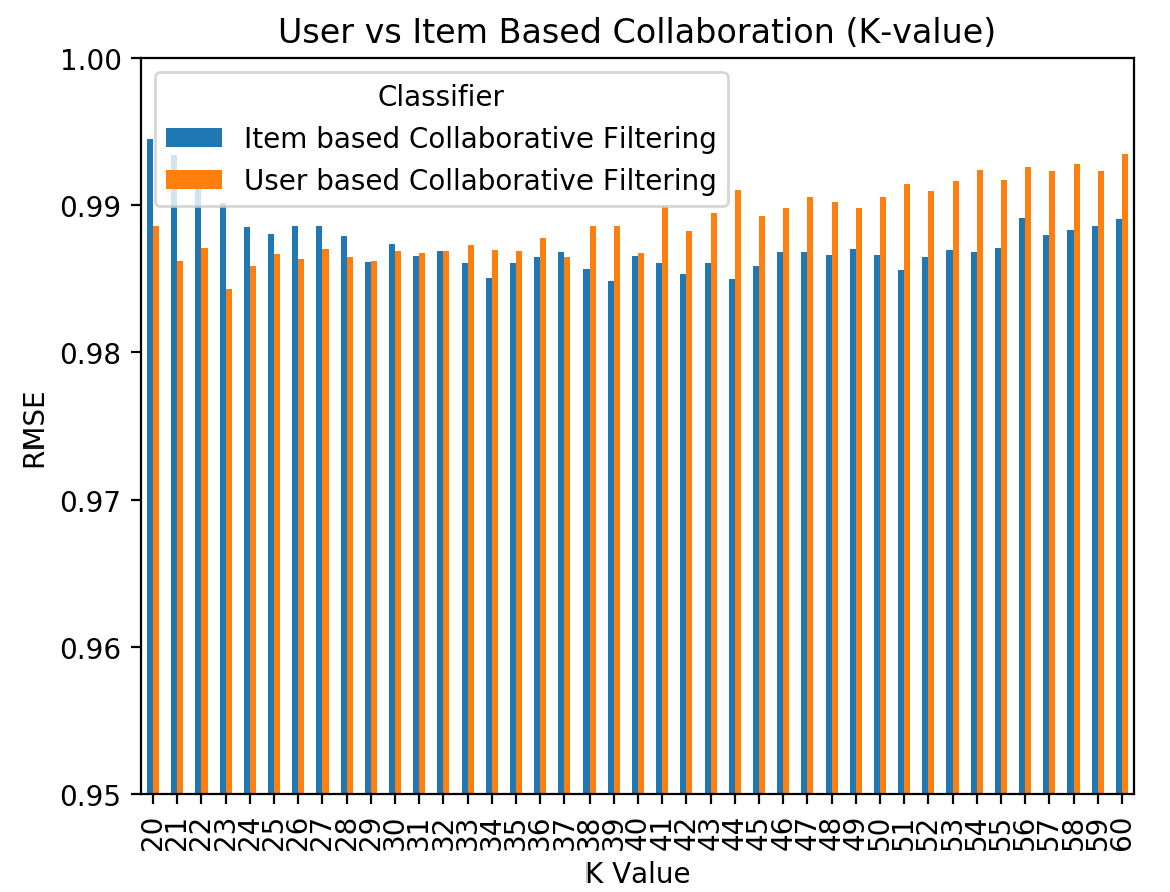
|  |  |  |
| --- | --- | --- |
|  | RMSE | MAE |
| SVD | 0.9463 | 0.7470 |
| PMF | 0.9663 | 0.7626 |
| NMF | 0.9737 | 0.7646 |
| User based Collaborative Filtering | 0.9894 | 0.7821 |
| Item based Collaborative Filtering | 0.9856 | 0.7810 |

RMSE: SVD < PMF < NMF < Item based < User Based

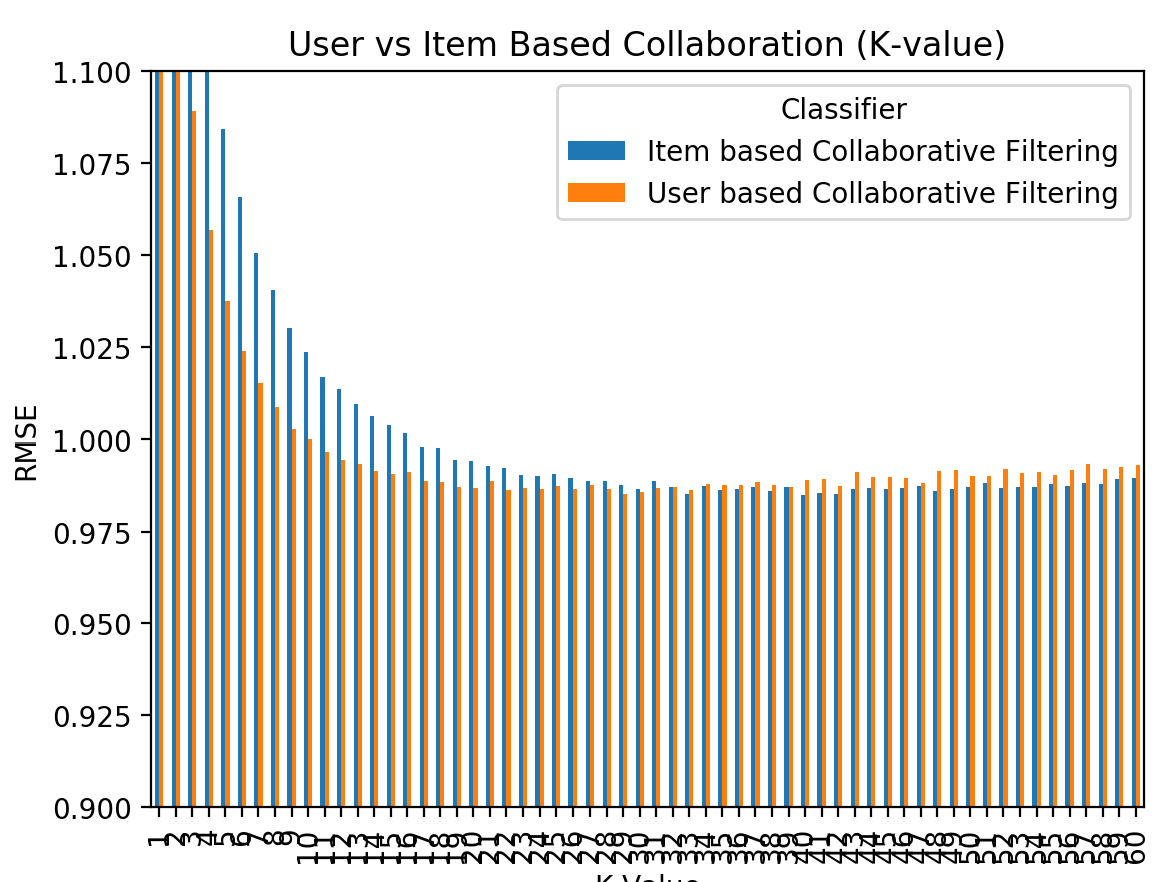
MAE: SVD < PMF < NMF < Item based < User Based

14. 1 – MSD, 2 – Cosine, 3 – Pearson

The RMSE reported above is the mean RMSE over 3 folds. MSD reported the lowest RMSE and MAE values. Cosine and Pearson were almost identical for User Based Collaboration, but for Item Based, Cosine reported less error then Pearson.

15.

The RMSE reported above is the mean RMSE over 3 folds. Item based collaboration appears to decrease in error as the number of neighbors increase, whereas user-based collaboration error increases.



Item based collaboration appears to drop off in error much slower than user based. User based least lowest error appears to be in the range of 18-22, while item based has the least error in the range of 36 – 42.