# **Hybrid Recommender System**

# **Problem Context and Description**

Recommendation systems are broadly classified into content based recommender systems and collaborative filtering based recommender systems. As far as data requirement for these recommender systems is concerned, content based recommender systems require item descriptions, item attributes and metadata which may not be available in practical cases whereas collaborative filtering based recommender systems just need ratings matrix (matrix of user/item/rating tuple). However, collaborative filtering approach suffers from “cold start” problem which means recommendations cannot be provided to new users or new items that are not part of ratings matrix.

Many techniques/algorithms are used to build collaborative filtering (CF) recommender systems – Singular Value Decomposition (SVD), Alternating Least Squares (ALS), Neighborhood based, etc. We measure the performance of five of these techniques/algorithms, namely – User-based CF, Item-based CF, SVD, ALS and Popular items, on the given dataset and propose a hybrid recommender system which essentially is an ensemble of decent performing techniques/algorithms on the given dataset.

# **Dataset Description**

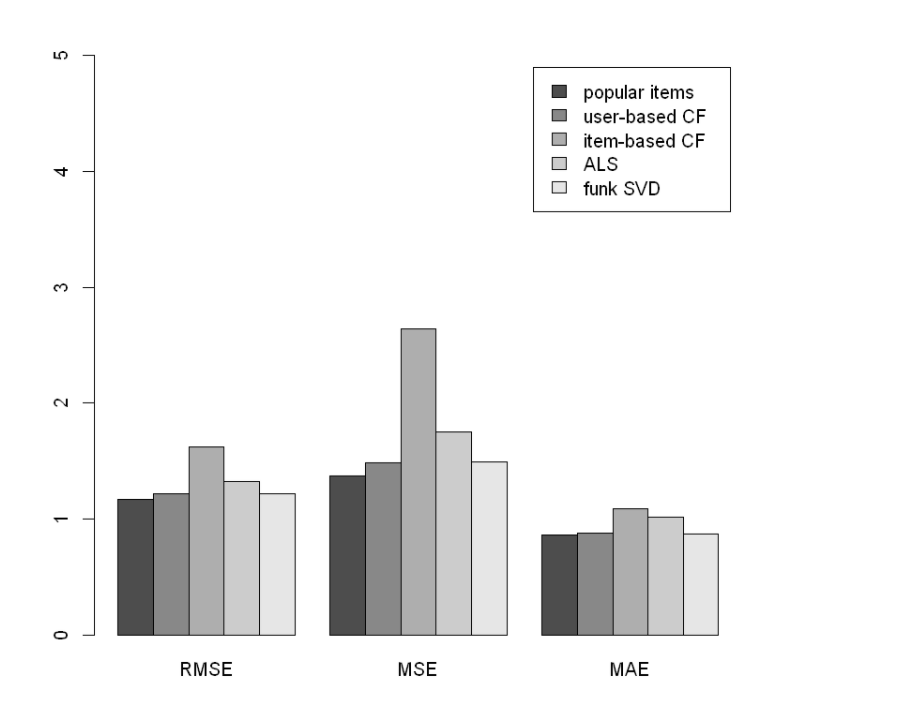
Data can be downloaded from: http://jmcauley.ucsd.edu/data/amazon/

(Review Data under Apps for Android)

Direct download link - http://snap.stanford.edu/data/amazon/productGraph/categoryFiles/ratings\_Apps\_for\_Android.csv

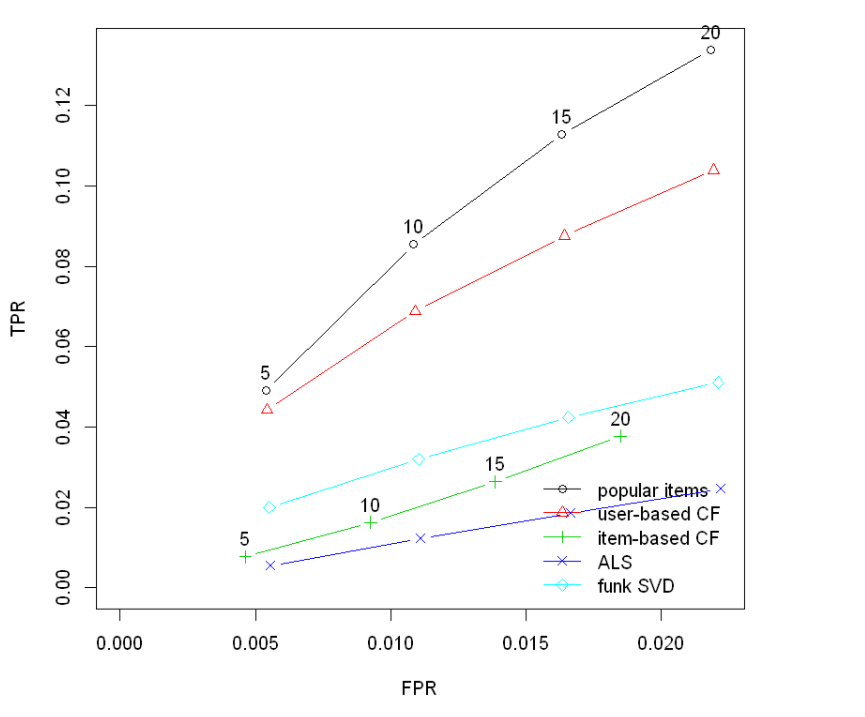
The dataset includes no metadata or reviews, but only (user, item, rating, timestamp) tuples having approx. 2.6+ million ratings of android apps.

# **Comparison of Prediction Accuracy**



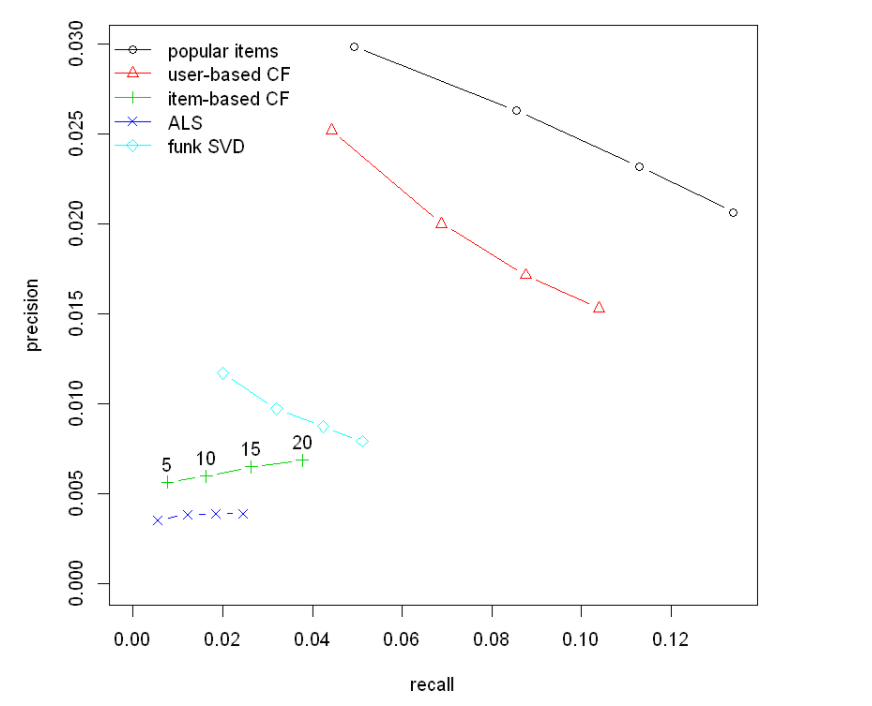
We can observe that User-based CF, Popular Items and Funk SVD perform better than ALS and Item-based CF.

# **Comparison of topN ROC Curve (N = 5, 10, 15, 20)**



We can again observe from the ROC curve that User-based CF, Popular Items and Funk SVD perform better than ALS and Item-based CF.

# **Comparison of topN Precision-Recall Curve (N = 5, 10, 15, 20)**

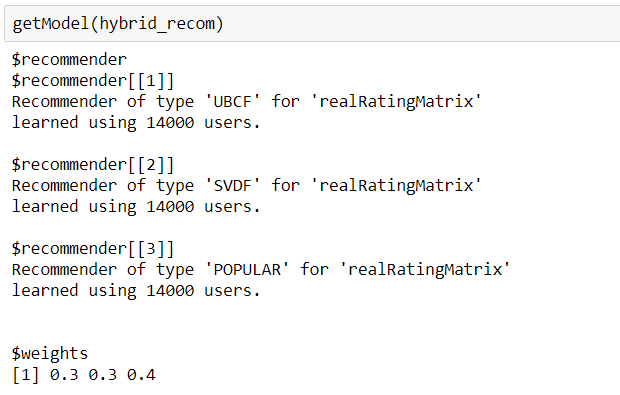


We can again observe from the precision-recall curve that User-based CF, Popular Items and Funk SVD perform better than ALS and Item-based CF.

# **Hybrid Recommender System**

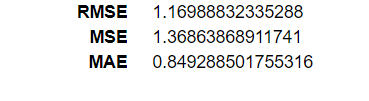
We select User-based CF, Popular Items and Funk SVD to create an ensemble for the hybrid recommender system.

## Model



## Accuracy

The accuracy of the hybrid recommender



## Top 5 Recommendations to 10 users

The top 5 recommendations given by the hybrid recommender to a sample of 10 users are as follows:

