Ryan Wedoff

Zongsheng Sun

Professor Stewart

4/18/16

# Group 3: Recommendation Systems Report

Companies like Netflix and Amazon use recommendation systems to predict products that the users will use next. The goal of this project was to replicate the algorithms used by these companies.

For our recommendations system, we emulated Netflix and their movie recommendations. Essentially, a user provides ratings for various movies and with that data we predicted which movies they will want to watch and like next.

### Computation

Our data was found on Grouplens.org (citation below). The data was a list with user ids, ratings, and movies ids. With this data, we created a matrix and ran a singular value decomposition on the data. We also calculated the general average rating for each movie and the exceed average rating for each user. With singular values decomposition, we took the orthogonal matrix U, the diagonal of the singular value vector and the transpose of the right singular values and multiplied them together to get the ratings for each person. We then added in the average rating for each movie, and the exceed average rating for the chosen user. We also subtracted out the movies that a user has already rated. This final list was a list of ratings for each movie, where the index of the rating represents the movie. Next we sorted this array and kept the indices of the top ten recommended movies. We used the movie file and obtained the titles of the top ten recommended movies for each user. The results for each user were saved in a separate file for quick access.

### Frontend Work

On the frontend side, we created a Web Application using Python’s Flask library. The web application, if in debug mode, computes the whole SVD and ratings. Otherwise, the web application uses the saved results from the previous computation. This was done to increase web page loading times.

### Testing

For testing, we computed recommendations for every user in the 100,000 ratings test data set. The 100,000 ratings data set was broken up into a base and a test set. We then computed the standard deviation between the two sets. The results!

### Observations

Conclusions

F. Maxwell Harper and Joseph A. Konstan. 2015. The MovieLens Datasets: History and Context. ACM Transactions on Interactive Intelligent Systems (TiiS) 5, 4, Article 19 (December 2015), 19 pages. DOI=http://dx.doi.org/10.1145/2827872