

# Study of Recommender Systems

Ronghao Yang

ID:20511820

David R. Cheriton School of Computer Science  
University of Waterloo

October 17, 2017

## Abstract

This project will focus on the the study of recommendation algorithms on specific recommendation problems. To have deeper insights of different algorithms, I am interested in testing their performance, applying them on specific datasets, and furthermore, having a comparison between them.

## 1 Introduction

In 2006, Netflix held a \$100,0000 competition[1] to call for people to improve their movie recommender system. Not only is a good recommender system important to Netflix, but also crucial in many real world applications such as book recommendation, hotel booking, event planning, etc. These algorithms usually involve player modelling, dimensionality reduction, and clustering, etc. Therefore, testing different recommendation algorithms' performances, analysing the capability of them, etc, become an interesting problem here.

## 2 Related work

For *The Netflix Prize*, in some of the winning papers[2][3][4][5]. Algorithms like  $K - NN$ , collaborative filtering, regression, matrix factorization were tested on the given dataset, the winning method has achieved  $RMSE = 0.8712$ .

In another of the related works proposed in 2015, Yu and Riedl from Georgia Tech have shown that  $NMF$ (Non-negative matrix factorization) is successful in recommending personalized interactive narratives.

## 3 Proposed work

### 3.1 Goal, challenges and plans

The overall goal here is try to achieve reasonably good results on the selected datasets and have an analysis of different recommendation algorithms and the results (including

why the algorithms (not) work on (certain) datasets).

This project will tackle one or two specific problems, apply different recommendation algorithms to them. The challenges may include algorithms performing poorly on the dataset, tuning the parameters of the algorithms, choosing the capable algorithms for specific problems.

To achieve the goal, I will first do a literature survey about different recommendation algorithms, learn about their advantages and disadvantages. And then apply the chosen algorithms on the selected dataset. At the end, there will be an analysis of algorithms, datasets and results.

### 3.2 Datasets

Datasets will possibly be selected from Kaggle Competition and UCL Machine Learning Repository, however, they have not been decided yet.

One of the interesting datasets is the Last.fm Dataset for music recommendation. Another interesting datasets is Expedia dataset for hotel recommendation.

### References

- [1] Netflix Prize homepage, Website, 2006. <http://www.netflixprize.com>
- [2] Andreas Toscher, Michael Jahrer *The BigChaos Solution to the Netflix Grand Prize* September 5, 2009
- [3] Andreas Toscher, Michael Jahrer *The BigChaos Solution to the Netflix Prize 2008*
- [4] Robert M. Bell, Yehuda Koren and Chris Volinsky *The BellKor solution to the Netflix Prize*
- [5] Robert M. Bell, Yehuda Koren and Chris Volinsky *The BellKor 2008 Solution to the Netflix Prize*
- [6] Hong Yu, Mark O. Riedl. *Personalized Interactive Narratives via Sequential Recommendation of Plot Points*