

Cmpe 462 Project Proposal

Movie Recommendation

Using Latent Factor Models

Mustafa Onur Eken

March 21, 2016

1 Introduction

In building recommendation systems, nowadays, there are two mainstream approaches: *Content-Based Filtering (CBF)* and *Collaborative Filtering (CF)*.

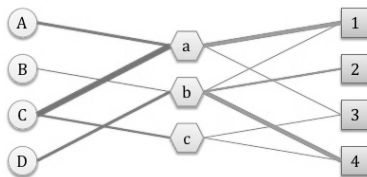
The key idea in the former method is that we are given some structured knowledge both about users and items themselves (Items are movies in our case) and we try to discover user, item pairs that have similar feature values.

The latter method does not require any intrinsic knowledge about users and items, but instead the only information needed is that a matrix A (usually sparse), where $a_{ij} \in A$ is the rating given to item j by user i . We can attempt to fill empty entries in the matrix by two different approaches. (1) *Neighborhood models* (2) *Latent factor models*.

2 Goals & Approach

In this project, we have obtained a dataset (provided by *MovieLens*) containing 100,000 ratings (1-5) from 943 users on 1682 movies. Our objective is to discover missing entries for each user and display recommendations accordingly. We are going to assume that there are a number (*say* k) of hidden factors in which people consider when rating a movie. Then by taking actual ratings into account, we discover the association between users, factors and as well as movies, factors. We will use matrix factorization methods [1, 2] such as *alternating least squares*, *stochastic gradient descent*, *probabilistic factorization* in order to build our factor matrices and compare their performances.

- Users and items are connect by latent features.



References

- [1] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems, 2009.
- [2] Ruslan Salakhutdinov and Andriy Mnih. Probabilistic matrix factorization. Citeseer, 2011.