



CS533 – Information Retrieval

Project Proposal

Project Members

Yusuf AYTAŞ	20603649	yusuf.aytas@bilkent.edu.tr
Kemal EROĞLU	20602600	keroglu@bilkent.edu.tr
Mustafa GÜNDOĞAN	20500329	gundogan@cs.bilkent.edu.tr
Fethi Burak Sazoğlu	20601331	fethi.sazoglu@bilkent.edu.tr

1. Introduction

Today, online services gather lots of irrelevant information about users. However, user wants personal related contents from the online service providers, so online service provides deal with the problem of providing 'related' content through the user information. In this project, we will be focusing on a movie recommendation system so called OMRES that uses collaborative filtering approach to recommend movies to users.

2. Problem Statement

Up until recently, recommendations from friends have been the most helpful way for us to find out about new things. In 2000s, the importance of providing personal information and contents to users has grown significantly. Most of the e-commerce companies, social web companies have recommender systems to provide personal information and expectations.

Earlier recommender system methods were developed to solve this kind of problems. Today, we know that there are two major divisions of recommender system methods; content based filtering and collaborative filtering. Content-based recommender systems make recommendations by analyzing the content of textual information and finding regularities in the content [1]. Collaborative filtering systems use the user rating data to calculate the similarity or weight between users or items and make predictions or recommendations according to those calculated similarity values [2].

3. Methodology

In OMRES, we will use the second approach which is collaborative filtering. Collaborative filtering has two different methodologies: Model based, which fits a model to the training data and Memory based, which uses Nearest Neighbor Methods. In our system, we will use k-NN approach that is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification [3].

In order to apply k-NN approach; firstly we will take all query instances such as $q(\text{movieId}, \text{userId})$, then we normalize the data before processing by the help of the algorithm. After that, we will find the distance of the selected instance with all the users who rated selected movie. Then, we will select top k users that are nearest to the query instance as its neighborhood. Lastly, we calculate the averages of ratings of the users form this neighborhood.

4. Expected Results

Netflix company offers a service that makes predictions about its users' movie taste and aims to recommend a movie to them. These predictions are based on the historical ratings of users for the movies they have watched. By 2006, they had a working algorithm called Cinematch, which were giving pretty good results (75% accuracy) for this purpose. But the company decided to make it better. To accomplish this they set up an open competition with a \$1,000,000 prize.[4] An algorithm by BellKor team managed to advance this accuracy by 10% and took the prize.[5] So in the end of our work we are planning to do our experiments and compare them with the

current Netflix solution. Hopefully we are aiming to achieve close or maybe better results than theirs.

5. Conclusion

Recommendation systems are a growing trend that provides content for us by looking at certain factors including what other people are doing as well as what we are doing. The basic algorithm of collaborative filtering is to compare the trends in how we rate movies to those of other people. In this project we are aiming to implement a recommendation system so called OMRES that is based on k-NN approach. Afterwards, we will compare our result with the results of the competition which organized by Netflix. Consequently, we propose a movie recommendation system, OMRES that we will implement and compare the results of it with other recommendation systems.

6. References

- [1] V. Moustakis G. Potamias and M. van Someren. "Using content-based filtering for recommendation." In CML/MLNET Workshop on Machine Learning and the New Information Age, Barcelona, pages 47–56, 2000.
- [2] Xiaoyuan Su and Taghi M. Khoshgoftaar, "A Survey of Collaborative Filtering Techniques," Advances in Artificial Intelligence, vol. 2009, Article ID 421425, 19 pages, 2009.
doi:10.1155/2009/421425
- [3] k-nearest neighbor algorithm, http://en.wikipedia.org/wiki/K-nearest_neighbor_algorithm
- [4] Netflix Prize, <http://www.netflixprize.com/index>
- [5] R. M. Bell, Y. Koren and C. Volinsky, "The BellKor solution to the Netflix grand prize," March 2012. [Online]. Available: <http://www2.research.att.com/~volinsky/netflix/ProgressPrize2007BellKorSolution.pdf>