A Comparison of How Demographic Data Affects Recommendation

Alan Said, Till Plumbaum, Ernesto W. De Luca, and Sahin Albayrak

TU Berlin,
DAI-LAB
{alan.said, till.plumbaum, ernesto.deluca, sahin.albayrak}@dai-lab.de

Abstract. Recommender systems attempt to find relevant data for their users. As the amount of data available on the Web grows, this task becomes increasingly harder. In this paper we present a comparison of recommendation results when using different demographic features (age, location and gender) commonly available in online communities. We assume that demographic data holds implicit information about users' taste and interests, and present results of a simple method that extends standard collaborative filtering algorithms to include one or several of these features. We evaluate our assumption in a movie recommendation scenario and compare results from different features to standard collaborative filtering.

Keywords: recommender systems, experimentation, human factors, demographics

1 Introduction

During the last 20 years, the amount of time that recommender systems have been researched, the de facto standard has been *Collaborative Filtering* (CF). However, current systems contain much more information about users than their counterparts twenty years ago did. One type of information commonly available in current systems is the *age*, *gender* and *location* of the users. Research has shown that these and similar features are of importance when attempting to increase the quality of recommenders [2,3]. In this paper, we apply the implicit relation brought by these features in a movie recommendation scenario by a simple extension to the k-Nearest Neighbor algorithm and show that even a very simple approach utilizing this sort of data brings significant improvements in terms of recommendation quality.

It has been previously shown that demographic data increases the quality of different information retrieval tasks. Weber and Castillo [3] used demographic information like average income, race, etc. to find difference between groups in a search engine scenario. Said et al. [2] showed that different social groups have difference in taste when it comes to movies as well.

In our work, we use a model that employs these features to create higher similarity scores between users from the same demographic groups. We present early stage results of experiments performed on a dataset containing demographic data.

The main contribution of this paper is a comparison of basic demographic features and their effect on recommendation quality in a Collaborative Filtering-based system.

2 Dataset and Experiments

In this paper we use a dataset provided by Moviepilot[1], which is Germany's largest online movie recommendation community. The snapshot used in our experiments contains the ratings of 10,000 randomly selected users who have rated at least one movie. In addition to the ratings, the dataset also contains the age for 1,292 users, gender for 6,583 users and city for 4,400 users. Table 1a shows the percentage of ratings performed by the users for which we have demographic data. The total number of ratings in our subset is 1,539,393 spread over four years (2006 to 2010). This corresponds to roughly 20% of the full dataset.

2.1 Experimental Setup

For the experiments, 50 training and evaluation sets each for every demographic feature were created. The evaluation sets consisted of 5000 ratings for 500 randomly selected users. The selected users had to have rated at least 30 movies. Out of these, 10 movies having been rated with a value above the user's average rating value were extracted (i.e. the set of true positive recommendations). The rest of the data was used for training. Users were assumed to belong to the same demographic group if they a) lived in the same city, b) were of the same gender, or c) were born in the same decade. For each of the demographic features, our recommender was run once with the similarity of users within the same demographic group multiplied by a factor set to 10,000, which is the the same as the number of users in our dataset, in order to significantly heighten the similarities. The recommendation algorithm used in our experiments was a slightly modified version of the *K-Nearest Neighbor* using the Pearson Correlation Coefficient as the neighbor similarity measure. Additionally, for comparison the recommender was run once for each training and test set without multiplying the similarities, i.e. the assigned similarities were solely based on the users' rating behavior. The results presented are the averages of all runs for each demographic group.

2.2 Results

We evaluate the results with Mean Average Precision (MAP) and Precision at 10. These measures were chosen since they are well-known and widely used in the field of Recommender Systems and Information Retrieval, providing a statistically sound estimate of the recommendation quality [1].

[1] http://moviepilot.de

	Number	%
City	$991,845$	64%
Age	$1,144,761$	74%
Sex	$1,398,732$	91%

(a) The number and percentages of ratings assigned by users who have stated either the city they live in , their age or their gender.

	P@10^{10K}	P@10	%
City	$2.79E-4$	$2.66E-4$	4.7%
Age	$2.45E-4$	$2.45E-4$	0.0%
Sex	$2.86E-4$	$2.33E-4$	22.9%

(b) The Precision@10 values for the demographic-aware recommender and a regular CF one.

	MAP^{10K}	MAP	%
City	$3.89E-3$	$3.81E-3$	2.2%
Age	$3.93E-e$	$3.93E-3$	0.0%
Sex	$4.22E-3$	$3.82E-3$	10.4%

(c) The Mean Average Precision values for the demographic-aware recommender and a regular CF one.

Table 1: Data statistics and results

The initial tests showed that our assumption, "demographic data has an impact on CF", is true. Gender, especially, seems to have a large impact with a resulting increase of 10% for MAP (and 22% Precision at 10). Age seems, however, not to have a very high impact. This could be due to the way we define age, i.e. born in the 60's, 70s, 80s, etc. We intend to treat age with a more dynamic approach using time slices, e.g. a sliding window of +/- a number of years.

3 Conclusion and Future Work

Our early stage results show that demographic data does matter even in a movie recommendation scenario. We expect that finer grained data, similar to that used by Weber and Castillo [3] will most likely affect the quality of a recommender system even more, especially if used in a more elaborate way than this simple CF extension. Our current work focuses on collecting and connecting this data to what we already have, as well as finding subgroups based on several features, e.g. age and city combined.

References

1. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. 22 (2004)
2. Said, A., De Luca, E.W., Albayrak, S.: How social relationships affect user similarities. In: IUI Workshop on SRS (2010)
3. Weber, I., Castillo, C.: The demographics of web search. In: SIGIR'10. ACM (2010)