

Hybrid content-collaborative recommender for personalized multimedia item retrieval

Matevž Kunaver, mentor: Jurij Tasič, Andrej Košir

*Faculty of Electrical Engineering, University of Ljubljana
E-mail: matevz.kunaver@fe.uni-lj.si*

Abstract. User modeling is a procedure used to filter the available content in order to present the user with a selection of items which are deemed to be interesting for this particular user. In this thesis we developed and tested two recommender systems with two different datasets. The first dataset was almost empty (it included only a small percent of all the possible ratings) and was therefore useful for creating an early-stage scenario to simulate users having only recently begun to use the system. The second dataset which was almost full (almost all the possible ratings were inputted) enabled us to simulate a scenario where the users were using the system for a longer period of time.

We evaluated both systems using the F-measure which is frequently used in the field of user modeling. In addition to the F-measure metric we also performed statistical testing on all our samples in order to determine whether they are significantly different or not.

During the development of our first system, our focus was on collaborative recommenders that are based on the nearest-neighbor search. We tested two different metrics for the nearest-neighbor selection to determine which of them provides better performance. Our results show that a combination of both methods performs best. When testing methods for calculating predicted ratings, we tested the Pearson's weighted sum and the True Bayesian's Estimate. Based on our results, we developed a new method – adjusted weighted sum.

In our second recommender system we expanded the existing user profiles by adding genre preferences, thus creating a hybrid recommender system. These preferences reflected the user's preferences for each genre type. We used these preferences to select the nearest neighbors for the active user in order to select the most appropriate items for him/her.

The hybrid system was working noticeably faster while still maintaining a high level of efficiency (high values of the F-measure). This also demonstrates that hybrid recommenders perform better due to a combination of different approaches.