MACHINE LEARNING
DEPARTMENT

Data Analysis Project Presentation

Jointly Modeling Aspects, Ratings and Sentiments for Movie Recommendation

Abstract:

Recommendation and review sites offer a wealth of information beyond ratings. For instance, on IMDb users leave reviews, commenting on different aspects of a movie (e.g. actors, plot, visual effects), and expressing their sentiments (positive or negative) on these aspects in their reviews. This suggests that uncovering aspects and sentiments will allow us to gain a better understanding of users, movies, and the process involved in generating ratings. The ability to answer questions such as "Does this user care more about the plot or about the special effects?" or "What is the quality of the movie in terms of acting?" helps us to understand why certain ratings are generated. This can be used to provide more meaningful recommendations.

In this work we propose a probabilistic model based on collaborative filtering and topic modeling. It allows us to capture the interest distribution of users and the content distribution for movies; it provides a link between interest and relevance on a per-aspect basis and it allows us to differentiate between positive and negative sentiments on a per-aspect basis. Unlike prior work our approach is entirely unsupervised and does not require knowledge of the aspect specific ratings or genres for inference.

We evaluate our model on a live copy crawled from IMDb. Our model offers superior performance by joint modeling. Moreover, we are able to address the cold start problem — by utilizing the information inherent in reviews our model demonstrates improvement for new users and movies.

DAP Committee: Alex Smola, Roy Maxion

Speaker:
Chao-Yuan Wu

Date: 4/21/15
Time: 3:00pm
Place: 4303 GHC