Tutorial on Recent Progress in Collaborative Filtering

Yehuda Koren

Collaborative filtering is a relatively young algorithmic approach, which already found its way into many commercial applications and established itself as a prime component of recommender systems. The field enjoyed a surge of interest thanks to the Netflix Prize competition that began in October 2006. Impact on research was significant in several ways. For the first time, public research community could access a large-scale industrial strong dataset. In addition, thousands of scientists, students, engineers and enthusiasts were attracted to the field. Importantly, all are judged by the same rigid yardstick, what makes comparisons and experiments much more meaningful. The nature of the competition encourages a rapid development pace, where first generation techniques were quickly replaced by second and third generation innovations.

Pushing techniques to be more and more accurate requires deepening their foundations, while reducing reliance on arbitrary decisions. An interesting outcome is forming surprising links among seemingly different techniques. For example, at their limit, user-user and item-item neighborhood models may converge to a single model. Furthermore, at that point both become equivalent to a simple matrix factorization model.

Some previous distinctions become less relevant. A major example is the traditional broad categorization of techniques into "model based" and "memory based", which is no longer appropriate. Traditional heuristic, memory-based techniques evolved to rely on rigorous models. On the other hand, some techniques which were considered purely model-based, are now improving their accuracy by also directly accessing all past ratings, a memory-based habit.

The quest for more accurate models does not stop at pushing the foundations of the models to their limits. At least as important is the identification of which kinds of signal, or features, are extractable from the data. Conventional techniques address the sparse data of user-item preferences (or, ratings). Accuracy significantly improves by also addressing less obvious sources of information. One prime example includes all kinds of temporal effects reflecting the dynamic, time-series nature of the data. Not less important is listening to hidden feedback that users provide. One kind of such feedback is which items they chose to rate (regardless of rating values). Rated items are not selected at random, but rather reveal interesting aspects of user preferences, going far and beyond the numerical values of the ratings.

There are many different ways to measure model accuracy. Commonly used measures, such as RMSE (or, MAE) tend to seemingly diminish differences among models. However, a new study finds that a small difference in RMSE can lead to a very significant improvement when ordering items by their predicted preferences. This way, quality of the top-K items presented to a user can vastly improve. For example, we have found that replacing traditional Pearson-coefficient item-item models with more rigorous item-item models triples the probability of placing a top-rated product at the top of a recommendation list.

However improved accuracy is not a panacea. There are several other challenges for collaborative filtering recommenders; we will deal with some of them. Obviously, as in all fields of computer science, computational efficiency is a key consideration. Some recommender systems need to scale to millions of users and items placing a high premium on efficiency. Explainability is another key aspect of recommenders. Users will likely benefit more from recommendations when they are accompanied with an intuitive reasoning such as "you will like this movie because you liked those movies". Good explainability is not going hand in hand with accuracy. In reality, it may be challenging to explain the reasoning beyond the more sophisticated approaches. Yet, much can be done to bridge these two important goals and devise accurate and explainable techniques. Another area that attracts new developments is the analysis of implicit feedback. While most research is centered on *explicit feedback*, where users tell us about their preferences, in many scenarios such feedback is hard to collect and we need to work with the more prevalent *implicit feedback*. Then, unobtrusive observations on users' behavior hint us on their preferences. Beyond the obvious difficulties in collecting and interpreting implicit feedback, it also requires special treatment through revamped models. To begin with, reliable negative feedback is no longer available. The sparse set of observed actions corresponds to only positive feedback, making it necessary to look at the broader full dense set of user-item relations.

To summarize, we believe that collaborative filtering is still a young field with many open challenges, which we, researchers, like to view as opportunities. Recent progress, mainly thanks to the widely popular Netflix Prize competition led to some innovations that will hopefully mark a significant impact on the field and find their way into the growing number of real life recommender systems.