Walking in Netflix: A Case Study of Collaborative Filtering for Social Media Recommendation System

Group Member: Wei BI, Wei WANG

Dataset

Netflix provided a training data set of 100,480,507 ratings that 480,189 users gave to 17,770 movies. Each training rating is a quadruplet of the form <user, movie, date of grade, grade>. The user and movie fields are integer IDs, while grades are from 1 to 5 (integral) stars. The qualifying data set contains over 2,817,131 triplets of the form <user, movie, date of grade>, with grades known only to the jury. A participating team's algorithm must predict grades on the entire qualifying set, but they are only informed of the score for half of the data, the quiz set of 1,408,342 ratings. The other half is the test set of 1,408,789, and performance on this is used by the jury to determine potential prize winners. Submitted predictions are scored against the true grades in terms of root mean squared error (RMSE), and the goal is to reduce this error as much as possible.

Suggested Approach

Recommendation systems suggest items of interest and enjoyment to people based on their preferences, which play an important role in Netflix. Netflix, an on-line movie subscription rental service, allows people to rent movies for a fixed monthly fee, maintaining a prioritized list of movies they wish to view (their "queue"). Movies are mailed to them or delivered electronically over the Internet. In the case of DVDs, when they are finished watching the movie they can return it by post and the next DVD is automatically mailed, postage free.

The Cinematch recommendation system automatically analyzes the accumulated movie ratings weekly using a variant of Pearson's correlation with all other movies to determine a list of "similar" movies that are predictive of enjoyment for the movie. The performance of Cinematch is measured in several ways. In addition to various system throughput requirements, the accuracy of the system is determined by computing the root mean squared error (RMSE) [1] of the system's prediction against the actual rating that a subscriber provides.

In our project, we will study the state-of-art Collaborative Filtering approaches in case study of Netflix data sets and evaluate the pros and cons of these approaches in extensive experiments.

Related Works

In the past few years, much research was devoted to the Netflix dataset. Many works were published in the two KDD workshops dedicated to that dataset. Best reported results were obtained by integrating the factorization and neighborhood models. Results reported in this paper by pure factorization are more accurate, in a sense showing that addressing temporal dynamics is not less important than algorithmic sophistication created by integration of two different models.

Evaluation Metrics

To evaluate recommendation quality we use a combination of different metrics. The primary metrics we consider include click through rate (CTR), long CTR (only counting clicks that led to watches of a substantial fraction of the video), session length, time until first long watch, and recommendation coverage (the fraction of logged in users with recommendations). We use these metrics to both track performance of the system at an ongoing basis as well as for evaluating system changes on live traffic.

Reference

[1] Herlocker, J, Konstan, J., Terveen, L., and Riedl, J. Evaluating Collaborative Filtering Recommender Systems. ACM Transactions on Information Systems 22 (2004), ACM Press, 5-53.

[2] Koren, Yehuda. “Factorization meets the neighborhood: a multifaceted collaborative filtering model.” In Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, 426–434. ACM, 2008.

[3] Koren, Yehuda. “Collaborative filtering with temporal dynamics.” Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’09 (2009): 447.

[4] Das, A.S., M. Datar, A. Garg, and S. Rajaram. “Google news personalization: scalable online collaborative filtering.” In Proceedings of the 16th international conference on World Wide Web, 271–280. ACM New York, NY, USA, 2007.

[5] Linden, G., B. Smith, and J. York. “Amazon.com recommendations: item-to-item collaborative filtering.” IEEE Internet Computing 7, no. 1 (January 2003): 76-80.

[6] Davidson, James, Benjamin Liebald, and Taylor Van Vleet. “The YouTube Video Recommendation System.” Design (2010): 293-296.