The Netflix Prize and Recommender Systems

Andreas Töscher and Michael Jahrer

25. Juni 2008
1. Who are we

2. Netflix

3. Models in Recommender Systems

4. Outlook
Andreas Töscher

- Telematik student since 2003
- BA 2006
- Computational Intelligence / Computer Vision
Michael Jahrer

- Telematik student since 2003
- BA 2006
- System on Chip Design / Computer Vision / Computational Intelligence
Netflix as Company

The Best Way to Rent Movies

FREE TRIAL (Offer Details)
- You'll get free shipping both ways
- Watch classics to new releases to TV series
- Cancel anytime

Plans from only $4.99 a month

Start Now
Netflix is a US online movie rental service
- Lend movies over mail
- over 100,000 titles
- 55 million DVDs total
- Productive start at 1997
- Have own recommendation system called “Cinematch”
  - Based on linear neighborhood model with a lot of data conditioning
- Approximately 60% of Netflix members select their movies based on movie recommendations
Grand Prize, 1 Mio. US-Dollar for 10% improvement in prediction accuracy
Progress Prize, 50.000 US-Dollar, October every year
Starts Oct, 2 2006
End Oct, 2 2011
or someone reach 10% improvement in RMSE
www.netflixprize.com
Overview Recommendation Systems

- Extraction of user's taste
- Top-k recommendations
  - List on online account
  - Recommendations as personal email

Examples
- Netflix
- Amazon.com
- last.fm
Dataset consists of 100 Mio. entries

Quadruples of \( <\text{movie}, \text{user}, \text{rating}, \text{date} > \)

Integer ratings from 1 to 5

Error measure: RMSE (root mean square error)

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (r_i - t_i)^2} \quad (r_i = \text{prediction} \quad t_i = \text{target})
\]

Over 30,000 contestants from 170 countries
Dataset Details

- Training Set: ~100Mio.
- Probe Set: ~2.8Mio.
- Qualifying Set: ~2.8Mio.

Feedback from 50% subset

Histogram of user ratings in the training set:

- Count
- Ratings per user

Andreas Töscher and Michael Jahrer
Average error as function of user votes

- Grand Prize level at $\sim 100$ votes/user
RMSE Scores

- 0.8563 (10.0%) Grand Prize
- 0.8643 (9.15%) Leader
- 0.8667 (8.90%) Our current progress
- 0.8712 (8.43%) Progress Prize Winner 2007
- 0.9514 (0.0%) Netflix Cinematch
- 1.0540 (-10.78%) Movie Average
Overview of Models

- Similarity based
  - Item based
  - User based
- Latent factor model
- Neural Networks
- Restricted Boltzmann Machines
- Hybrid approaches
- Target: Rating prediction of any user/item combination
Similarity between users

\[ U_0 = [1, 4, 3] \]
\[ U_1 = [1, 5, 2] \]
**Prediction from user similarity**

- Pearson correlation between users
- Rating prediction for user 5 on item 0:
  - select k-best correlating users which voted item 0
  - weight the ratings of the k-best users with their correlation to user 5

---

**Item 0**

User 0 -> User 1 -> User 2 -> User 3 -> User 4 -> User 5 -> User 6 -> User 7

Rating:
- User 0: 2
- User 1: 5
- User 2: 5
- User 3: 5
- User 4: ?
- User 5: 3
- User 6: ?
- User 7: ?

---

Andreas Tösscher and Michael Jahrer

Netflix Prize
Prediction from user similarity

- $N(u, i)$ ... k-best correlating users
- $c_{uv}$ ... Pearson correlation between users $u$ and $v$

\[
\hat{r}_{ui} = \frac{\sum_{v \in N(u, i)} c_{uv} r_{vi}}{\sum_{v \in N(u, i)} c_{vi}}
\] (1)
Notes on user similarity models

- In general two users have very few common ratings
- 500,000 users in the Netflix dataset
- Precomputation of all user/user correlations is not possible (takes over 1TByte storage)

**Item/Item Correlations**
- In general well defined
- 18,000 movies in the Netflix dataset
- Precomputable
- Predictions are more accurate
Rating Matrix factorization

- Low-rank approximation of the rating matrix
- Prediction given by inner product of item and user feature
- Fast and good performance
- Problem: $R$ is very sparse, 1% filled in the Netflix dataset
Rating Matrix factorization: Rating Prediction

- Rating prediction is given by a dot product from an item and an user feature.
Notes on Matrix Factorization

- Fast and good performance
- Training can be done with gradient descent
- Regularization is important
- A single matrix factorization can achieve an improvement of more than 5%
- Very accurate model for users with many votes
- A restriction to non-negative features can also be used
Combination of Predictions

- Combination with linear blending
- Simple and efficient
- $\alpha_n$ can be calculated with pseudo-inverse

Training Set | Probe Set | Qualifying Set
---|---|---
~ 100Mio. | ~ 2.8Mio. | Feedback from 50% subset

P1 P2 P3 P4 P5
$\alpha_1 \alpha_2 \alpha_3 \alpha_4 \alpha_5$
Recommender systems can be used in

- Webshops
- Search engines
- Social networks
- Personalized advertisement
- Everywhere a back-channel exists
The secret behind our success

- Knowledge in machine learning
- Independent exploration
- Different viewpoints
- Communication is “all”
  - Many skype-hours of talking
- Good intuition
  - How can we squeeze out max. information out of the dataset?
- Good programming skills to generate fast code
- Computing power
  - 8x DualCore >3GHz, 8GB RAM
Leading team from AT&T Research
We are currently on 2\textsuperscript{nd} place (team BigChaos)
Thank you for your attention.