The Netflix Prize

• • •

Ahmed El-Razi

The Setup

- Improve upon Netflix's Cinematch Algorithm
- Start date: October 2, 2006
- $1,000,000 Grand Prize for best collaborative filtering algorithm
- 10% Improvement Goal (open ended)
- Final prize awarded on September 18, 2009

The Scoring

- Training set: 100,480,507, Test set: 1,408,789
- RMSE of Trivial Algorithm: **1.0540**
- RMSE of Cinematch: **0.9514**
- Grand Prize for a RMSE of **0.8572**
- Progress prizes of $50,000 each year Grand Prize not won

The Contest

- 44,014 submissions, 5,169 teams
- October 8, 2006 WXYZ Consulting
- October 15, 2006 Cinematch beaten by 1.06%
- Frontrunners (WXYZ, ML@UTor, Gravity, BellKor (AT&T)
- BellKor 10.09% and Ensemble 10.10%
- Grand Prize: BellKor September 2009

The Aftermath

- Sequel scrapped due to lawsuit (UT researchers)
- Algorithm improvement at a discount
- Kaggle launch 2010

kaggle

The Solution

- Paper: http://www.netflixprize.com/assets/GrandPrize2009_BPC_BellKor.pdf
- Model was a blend of 3 models (GBDT, RBM, Neighborhood)
- User Bias: bui = μ +bu +bi
- Time Change: Popularity over time, Change in baseline rating
- 24 total predictors

Questions?