CS 277: The Netflix Prize

Professor Padhraic Smyth Department of Computer Science University of California, Irvine

Netflix

- Movie rentals by DVD (mail) and online (streaming)
- 100k movies, 10 million customers
- Ships 1.9 million disks to customers each day
 - 50 warehouses in the US
 - Complex logistics problem
- Employees: 2000
 - But relatively few in engineering/software
 - And only a few people working on recommender systems
- Moving towards online delivery of content
- Significant interaction of customers with Web site

The \$1 Million Question



Million Dollars Awarded Sept 21st 2009



Background



Training Data

100 million ratings (matrix is 99% sparse)

Rating = [user, movie-id, time-stamp, rating value]

Generated by users between Oct 1998 and Dec 2005

Users randomly chosen among set with at least 20 ratings

– Small perturbations to help with anonymity



Structure of Competition

- Register to enter at Netflix site
- Download training data of 100 million ratings
 - 480k users x 17.7k movies
 - Anonymized
- Submit predictions for 3 million ratings in "test set"
 True ratings are known only to Netflix
- Can submit multiple times (limit of once/day)
- Prize
 - \$1 million dollars if error is 10% lower than Netflix current system
 - Annual progress prize of \$50,000 to leading team each year

Scoring

Minimize root mean square error

Mean square error =
$$1/|R| \sum_{(u,i) \in R} (r_{ui} - \hat{r}_{ui})^2$$

Does not necessarily correlate well with user satisfaction

But is a widely-used well-understood quantitative measure





RMSE Baseline Scores on Test Data

1.054 - just predict the mean user rating for each movie

0.953 - Netflix's own system (Cinematch) as of 2006

0.941 - nearest-neighbor method using correlation

0.857 - required 10% reduction to win \$1 million

Other Aspects of Rules

- Rights
 - Software + non-exclusive license to Netflix
 - Algorithm description to be posted publicly
- Final prize details
 - If public score of any contestant is better than 10%, this triggers a 30-day final competition period
 - Anyone can submit scores in this 30-day period
 - Best score at the end of the 30-day period wins the \$1 million prize
- Competition not open to entrants in North Korea, Iran, Libya, Cuba....and Quebec

Why did Netflix do this?

- Customer satisfaction/retention is key to Netflix they would really like to improve their recommender systems
- Progress with internal system (Cinematch) was slow
- Initial prize idea from CEO Reed Hastings
- \$1 million would likely easily pay for itself
- Potential downsides
 - Negative publicity (e.g., privacy)
 - No-one wins the prize (conspiracy theory)
 - The prize is won within a day or 2
 - Person-hours at Netflix to run the competition
 - Algorithmic solutions are not useful operationally

Key Technical Ideas

Outline

- Focus primarily on techniques used by Koren, Volinsky, Bell team (winners of prize)
 - We will focus on some of the main ideas used in their algorithms
 - Many other details in their papers, and in other papers published on the Netflix data set
- Useful References
 - Y. Koren, Collaborative filtering with temporal dynamics, ACM SIGKDD Conference 2009
 - Koren, Bell, Volinsky, Matrix factorization techniques for recommender systems, IEEE Computer, 2009
 - Y. Koren, Factor in the neighbors: scalable and accurate collaborative filtering, ACM Transactions on Knowledge Discovery in Data, 2010

$R = U \Sigma V^t$

m x n m x n n x n n x n

where: columns of V are eigenvectors of R^tR

 Σ is diagonal (eigenvalues)

rows of U are coefficients in V-space of each row in R

Matrix Approximation with SVD

$R \approx U \Sigma V^{t}$

m x n m x f f x f f x n

where: columns of V are first f eigenvectors of R^tR

 Σ is diagonal with f largest eigenvalues

rows of U are coefficients in reduced dimension V-space

This approximation is the best rank-f approximation to matrix R in a least squares sense (principal components analysis)

Matrix Factorization of Ratings Data



Matrix Factorization of Ratings Data





Figure from Koren, Bell, Volinksy, IEEE Computer, 2009

Computation of Matrix Factors

Problem 1:

Finding the f factors is equivalent to performing a singular value decomposition of a matrix, i.e.,

Let R be an *m x n* matrix

SVD computation has complexity $O(mn^2 + n^3)$

Computation of Matrix Factors

Problem 1:

Finding the f factors is equivalent to performing a singular value decomposition of a matrix, i.e.,

Let R be an *m x n* matrix

SVD computation has complexity $O(mn^2 + n^3)$

Problem 2:

Most of the entries in R are missing, i.e., only 100 x 10⁶ / (480k x 17k) \sim 1% are present

Dealing with Missing Data

$$r_{ui} \approx q_i^t p_u$$

$$\min_{q,p} \sum_{(u,i) \in R} (r_{ui} - q_i^t p_u)^2$$

sum is only over known ratings

Dealing with Missing Data

$$r_{ui} \approx q_i^t p_u$$

$$\min_{q,p} \sum_{(u,i) \in R} (r_{ui} - q_i^t p_u)^2$$

Add regularization

$$\min_{q,p} \sum_{(u,i) \in R} (r_{ui} - q_i^t p_u)^2 + \lambda (|q_i|^2 + |p_u|^2)$$

Stochastic Gradient Descent (SGD)

$$\begin{array}{lll} & \text{goodness of fit} & \text{regularization} \\ & \min_{q,p} \, \sum_{_{(u,i) \, \epsilon \, R}} \left(\begin{array}{cc} r_{ui} & - & q^t{}_i & p_u \end{array} \right)^2 + \, \lambda \, \left(\left| \begin{array}{cc} q_i \right|^2 + \, \left| \begin{array}{cc} p_u \right|^2 \right) \end{array} \right) \end{array} \right. \end{array} \right.$$

Online ("stochastic") gradient update equations:

$$\begin{aligned} \varepsilon_{ui} &= r_{ui} - q_i^t p_u \\ q_i &<= q_i + \gamma (\varepsilon_{ui} p_u - \lambda q_i) \\ etc.... \end{aligned}$$

 μ = overall mean rating

 $b_u = mean rating for user u$

 b_i = mean rating for movie i

Components of a rating predictor



Baseline predictor

- Separates users and movies
- Often overlooked
- Benefits from insights into users' behavior
- Among the main practical contributions of the competition

User-movie interaction

- Characterizes the matching between users and movies
- Attracts most research in the field
- Benefits from algorithmic and mathematical innovations

(slide from Yehuda Koren)

A baseline predictor

 We have expectations on the rating by user u of movie i, even without estimating u's attitude towards movies like i





 Values of other ratings user gave recently (day-specific mood, anchoring, multi-user accounts)



- (Recent) popularity of movie i
- Selection bias; related to number of ratings user gave on the same day ("frequency")

(slide from Yehuda Koren)

Modeling Systematic Biases



Example:

Mean rating $\mu = 3.7$

You are a critical reviewer: your ratings are 1 lower than the mean $-> b_u = -1$

Star Wars gets a mean rating of 0.5 higher than average movie: $b_i = +0.5$

Predicted rating for you on Star Wars = 3.7 - 1 + 0.5 = 3.2

$$\min_{q,p} \ \left\{ \ \sum_{(u,i) \ \epsilon \ R} (\ r_{ui} \ - \ (\mu \ + \ b_u \ + \ b_i + \ q^t_{\ i} \ p_u) \)^2 \right. \label{eq:goodness}$$

+
$$\lambda \left(|q_i|^2 + |p_u|^2 + |b_u|^2 + |b_i|^2 \right)$$

 \uparrow regularization

Typically selected via grid-search on a validation set



Figure from Koren, Bell, Volinksy, IEEE Computer, 2009

Adding Implicit Information





Figure from Koren, Bell, Volinksy, IEEE Computer, 2009


Rating by movie age





Time-dependence parametrized by linear trends, binning, and other methods

For details see

Y. Koren, Collaborative filtering with temporal dynamics, ACM SIGKDD Conference 2009

Models the fact that user's interests over "genres" (the q's) may change over time



Figure from Koren, Bell, Volinksy, IEEE Computer, 2009

The Kitchen Sink Approach....

- Many options for modeling
 - Variants of the ideas we have seen so far
 - Different numbers of factors
 - Different ways to model time
 - Different ways to handle implicit information

•

- Other models (not described here)
 - Nearest-neighbor models
 - Restricted Boltzmann machines
- Model averaging was useful....
 - Linear model combining
 - Neural network combining
 - Gradient boosted decision tree combining
 - Note: combining weights learned on validation set ("stacking")

Ensemble NNBlend

Train many small NN's (>1000) on a random subset Per net: 20..40 weights Combine them linearly predictors about 1000 nets #2 #1000 #1 linear combiner → our best probe blend: PB-101: RMSE=0.8584 prediction Michael Jahrer / Andreas Töscher - Team BigChaos - September 21, 2009

CS 277: Data Mining

Other Aspects of Model Building

- Automated parameter tuning
 - Using a validation set, and grid search, various parameters such as learning rates, regularization parameters, etc., can be optimized
- Memory requirements
 - Memory: can fit within roughly 1 Gbyte of RAM
- Training time
 - Order of days: but achievable on commodity hardware rather than a supercomputer
 - Some parallelization used

Matrix factorization vs Near Neighbor?

From Koren, ACM Transactions on Knowledge Discovery, 2010

"Latent factor models such as SVD face real difficulties when needed to explain predictions. ...Thus, we believe that for practical applications neighborhood models are still expected to be a common choice."

The Competition: 2006-2009

Setting up and Launching...

- Summer 2006
 - Email from Netflix about large monetary award
 - Is this real?
 - Apparently so: serious and well-organized
 - Spent summer carefully designing data set and rules

- Official Launch, Oct 2nd 2006
 - Email lists, conferences, press releases, etc
 - Significant initial interest in research community, blogs, etc
 - 40,000 teams (eventually) from over 150 countries.
 - Number of initial registrants significantly exceeded expectations

Progress in first 3 months

Oct 2, 2006	Launch of competition
Oct 8, 2006	WXY Consulting already better than Cinematch score
Oct 15, 2006	3 teams above Cinematch, one with 1.06% improvement (qualifying for \$50k progress prize)
Dec, 2006:	Jim Bennett from Netflix describes progress so far during an invited talk at NIPS

Prize Progress



Prize Submissions



Prize Submissions







Home Program KDD Cup Registration Travel Organizers

Co-organized by ACM SIGKDD and Netflix

KDD Cup is the first and the oldest data mining competition, and is an integral part of the annual ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). This year's KDD Cup will be related to (but different from) the current Netflix Prize competition (http://www.netflixprize.com/). There will also be a workshop at the KDD-07 conference, where the participants of both the KDD Cup and the current Netflix Prize competition will present their papers and exchange ideas.

There are 2 parallel options for participating:*

- 1. The KDD Cup competition (open to all)
- 2. Workshop paper submissions (open to Netflix prize participants only)

Full details are provided at the KDD Cup and Workshop 2007 website:

http://www.cs.uic.edu/Netflix-KDD-Cup-2007

If you have any comments and suggestions, please email *liub@cs.uic.edu*.

Organizing Committee

KDD Cup

- Jim Bennett, Neflix, USA
- · Charles Elkan, University of California, San Diego, USA
- Bing Liu (Chair), University of Illinois at Chicago, USA
- Padhraic Smyth, University of California, Irvine, USA
- Domonkos Tikk, Budapest University of Technology and Economics, Hungary



Links

Papers Workshops Tutorials Panels Demos Awards Exhibits BOF

Organizational Sponsor:



Platinum Supporter:



Gold Supporters:



First Progress Prize, October 2007

Progress prize: \$50k annually awarded to leading team provided there is at least 1% improvement over previous year

Sept 2nd First progress prize "30 day" last call

Oct 2nd Leaders were BellKor, 8.4% improvement (Yehuda Koren, Bob Bell, Chris Volinksy, AT&T Research)

First Progress Prize, October 2007

Progress prize: \$50k annually awarded to leading team provided there is at least 1% improvement over previous year

Sept 2nd First progress prize "30 day" last call

Oct 2nd Leaders were BellKor, 8.4% improvement (Yehuda Koren, Bob Bell, Chris Volinksy, AT&T Research)

Oct/Nov Code and documentation submitted for judging

Complicated methods: primarily relying on factor models

Nov 13Winners officially declared and BellKor documentationpublished on Netflix Web site

Progress slows down...improvements are incremental

Many of the leading prize contenders publishing their methods and techniques at academic conferences (2nd KDD workshop in August)

Much speculation on whether the prize would ever be won – is 10% even attainable?

Many initial participants had dropped out – too much time and effort to seriously compete

But leaderboard and forum still very active

Sept 2 nd year	Only 3 teams qualify for 1% improvement over previous
Oct 2 nd	Leading team has 9.4% overall improvement
Oct/Nov	Code/documentation reviewed and judged

Sept 2 nd	Only 3 teams qualify for 1% improvement over previous
year	

Oct 2nd Leading team has 9.4% overall improvement

Oct/Nov Code/documentation reviewed and judged

Progress prize (\$50,000) awarded to BellKor team of 3 AT&T researchers (same as before) plus 2 Austrian graduate students, Andreas Toscher and Martin Jahrer

Key winning strategy: clever "blending" of predictions from models used by both teams

Speculation that 10% would be attained by mid-2009

Example of Predictor Specifications....

OB-42 rmse=0.8998

MovieKNNV3, Residual: OB-39, Pearson correlation, K = 13, $\alpha = 658$, $\beta = 2480$, $\gamma = -2.6$, $\delta = 7.5$

OB-43 rmse=0.8971

SVD-AUF, Residual: OB-30, adaptiveUserFactorMode=KRR, kernelType=extended-polynomial, $\lambda = 4.78376$, $\alpha = 0.657533$, $\gamma = 0.720031$, $\beta = 3.27554$

OB-44 rmse=0.9245

GE, Residual: OB-35, 16 effects, $\alpha_1 = 374.977$, $\alpha_2 = 8.90702e - 05$, $\alpha_3 = 2535.9$, $\alpha_4 = 900.414$, $\alpha_5 = 1.04115e - 05$, $\alpha_6 = 2087.92$, $\alpha_7 = 131.291$, $\alpha_8 = 3173.84$, $\alpha_9 = 1.45471e - 06$, $\alpha_{10} = 6.40823e - 08$, $\alpha_{11} = 4451.15$, $\alpha_{12} = 274.423$, $\alpha_{13} = 1020.64$, $\alpha_{14} = 0.00758424$, $\alpha_{15} = 3858.57$, $\alpha_{16} = 0.00346888$

OB-45 rmse=0.8998

MovieKNNV3, Residual: OB-40, Spearman's rank correlation, K = 38, $\alpha = 667.6$, $\beta = 255.5$, $\gamma = -1.39$, $\delta = 3.3$

OB-46 rmse=0.8958

MovieKNNV3, Residual: OB-41, Pearson correlation, K = 60, $\alpha = 804$, $\beta = 231$, $\gamma = -2.6$, $\delta = 17$

OB-47 rmse=0.9777

NSVD2, Residual: no, k = 50, $\eta_i = 2e - 3$, $\eta_u = 2e - 3$, $\eta_{\mu_i} = 2e - 3$, $\eta_{\mu_u} = 2e - 3$, $\lambda_i = 5e - 4$, $\lambda_u = 5e - 4$, $\lambda_{\mu_i} = 1e - 4$, $\lambda_{\mu_u} = 1e - 4$, 3 epochs

The big picture Solution of BellKor's Pragmatic Chaos



Michael Jahrer / Andreas Töscher – Team BigChaos – September 21, 2009

The End-Game

June 26th 2009: after 1000 Days and nights...

TFU	X			
NC Rule	s Leaderboard Registe	20 r Update Sut	Domit Download	
Lea	aderboard	1	Display top	20 leaders.
Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	BellKor's Pragmatic Chaos	0.8558	10.05	2009-06-26 18:42:37
Grand	Prize - RMSE <= 0.8563			
2	PragmaticTheory	0.8582	9.80	2009-06-25 22:15:51
3	BellKor in BigChaos	0.8590	9.71	2009-05-13 08:14:09
4	Grand Prize Team	0.8593	9.68	2009-06-12 08:20:24
5	Dace	0.8604	9.56	2009-04-22 05:57:03
6	BigChaos	0.8613	9.47	2009-06-23 23:06:52
Progra	ess Prize 2008 - RMSE = 0.1	8616 - Winning Te	am: BellKor in BigC	haos
7	BellKor	0.8620	9.40	2009-06-24 07:16:02
8	Gravity	0.8634	9.25	2009-04-22 18:31:32
9	Opera Solutions	0.8638	9.21	2009-06-26 23:18:13
10	BruceDengDaoCiYiYou	0.8638	9.21	2009-06-27 00:55:55
11	pengpengzhou	0.8638	9.21	2009-06-27 01:06:43
12	xivector	0.8639	9.20	2009-06-26 13:49:04
13	xiangliang	0.8639	9.20	2009-06-26 07:47:34
14	Feeds2	0.8641	9.18	2009-06-26 22:51:55
	0	0.0000	0.47	0000 06 04 14:04-14

The Leading Team

- BellKorPragmaticChaos
 - BellKor:
 - Yehuda Koren (now Yahoo!), Bob Bell, Chris Volinsky, AT&T
 - BigChaos:
 - Michael Jahrer, Andreas Toscher, 2 grad students from Austria
 - Pragmatic Theory
 - Martin Chabert, Martin Piotte, 2 engineers from Montreal (Quebec)
- June 26th submission triggers 30-day "last call"
- Submission timed purposely to coincide with vacation schedules

The Last 30 Days

- Ensemble team formed
 - Group of other teams on leaderboard forms a new team
 - Relies on combining their models
 - Quickly also get a qualifying score over 10%
- BellKor
 - Continue to eke out small improvements in their scores
 - Realize that they are in direct competition with Ensemble
- Strategy
 - Both teams carefully monitoring the leaderboard
 - Only sure way to check for improvement is to submit a set of predictions
 - This alerts the other team of your latest score

24 Hours from the Deadline

- Submissions limited to 1 a day
 - So only 1 final submission could be made by either team in the last 24 hours
- 24 hours before deadline...
 - BellKor team member in Austria notices (by chance) that Ensemble posts a score that is slightly better than BellKor's
 - Leaderboard score disappears after a few minutes (rule loophole)
- Frantic last 24 hours for both teams
 - Much computer time on final optimization
 - run times carefully calibrated to end about an hour before deadline
- Final submissions
 - BellKor submits a little early (on purpose), 40 mins before deadline
 - Ensemble submits their final entry 20 mins later
 - …and everyone waits….

TF	I F L I X				
N	Ruli	es Leaderboard Register Upc	late Subr	nit Download	
Le	28	aderboard		Display top	20 📦 leaders.
Rar	ık	Team Name	Best Scon	s % Improvem	ent ast Submit Time
1	:	The Ensemble	0.8553	10.10	009-07-26 18:38:22
2	1	BellKor's Pragmatic Chaos	0.8554	10.09	009-07-26 18:18:28
Gr	and	Prize - RMSE <= 0.8563			
3	÷	Grand Prize Team	0.8571	9.91	2009-07-24 13:07:49
4	1	Opera Solutions and Vandelay United	0.8573	9.89	2009-07-25 20:05:52
5	÷	Vandelay Industries !	0.8579	9.83	2009-07-26 02:49:53
6	÷	PragmaticTheory	0.8582	9.80	2009-07-12 15:09:53
7	÷	BellKor in BigChaos	0.8590	9.71	2009-07-26 12:57:25
8	:	Dace	0.8603	9.58	2009-07-24 17:18:43
9	÷	Opera Solutions	0.8611	9.49	2009-07-26 18:02:08
10	÷	BellKor	0.8612	9.48	2009-07-26 17:19:11
11	÷	BigChaos	0.8613	9.47	2009-06-23 23:06:52
12	÷	Feeds2	0.8613	9.47	2009-07-24 20:06:46
Pre	ogn	ess Prize 2008 - RMSE = 0.8616 - Wi	inning Team	: BellKor in Big	Chaos
13	:	xiangliang	0.8633	9.26	2009-07-21 02:04:40
14	-	Gravity	0.8634	9.25	2009-07-26 15:58:34
15	1	Ces	0.8642	9.17	2009-07-25 17:42:38
16	1	Invisible Ideas	0.8644	9.14	2009-07-20 03:26:12
17	1	Just a guy in a garage	0.8650	9.08	2009-07-22 14:10:42
18	1	Craig Carmichael	0.8656	9.02	2009-07-25 16:00:54
		J Dennis Su	0.8658	9.00	2009-03-11 09:41:54
19					

TF	LIX							
N	C Rules	Flix Prize	odate	Subr	mit Do	wnload	K	
Le	ea	derboard			Displ	ay top 20	🗘 leader:	5.
Ran	k	Team Name	Best	Scon	e % Impi	rovement	Last Subm	nit Time
1	: D	he Ensemble	0.8	553	: 1	0.10	2009-07-26	18:38:22
2	: 8	ellKor's Pragmatic Chaos	0.8	554	1	0.09	2009-07-26	18:18:28
Gra	and Pri	<u>ize</u> - RMSE <= 0.8563						
3	: 9	rand Prize Team	0.8	571	: 9	9.91	2009-07-24	13:07:49
4	: 0	pera Solutions and Vandelay United	0.8	573	: 9	9.89	2009-07-25	20:05:52
5	<u> </u>	andelay Industries !	8.0	579	: 9	9.83	2009-07-26	02:49:53
6	: 8	ragmaticTheory	0.8	582	: 9	9.80	2009-07-12	15:09:53
7	: <u>B</u>	ellKor in BigChaos	0.8	590	: 9	3.71	2009-07-26	12:57:25
8	: 0	808	0.8	603	: 9	9.58	2009-07-24	17:18:43
9	: 0	pera Solutions	0.8	611	: 9	1.49	2009-07-26	18:02:08
10	: B	ellKor	0.8	612	: 9	.48	2009-07-26	17:19:11
11	8	igChaos	0.8	613	: 9	47	2009-06-23	23:06:52
12	E	eeds2	0.8	613	: 9	9.47	2009-07-24	20:06:46
Pro	ogress	Prize 2008 - RMSE = 0.8616 - W	linning	Team	1: BellKor	in BigCha	ios	
13	: xi	angliang	8.0	633	: 9	9.26	2009-07-21	02:04:40
14	G	ravity	8.0	634	: 9	9.25	2009-07-26	15:58:34
15	: 0	es	8.0	642	: 9	9.17	2009-07-25	17:42:38
16	: In	visible Ideas	8.0	644	: :	3.14	2009-07-20	03:26:12
17	J	ust a guy in a garage	0.8	650	: 9	8.08	2009-07-22	14:10:42
	: 0	raig Carmichael	0.8	656	: 9	0.02	2009-07-25	16:00:54
18								00.44-54
18 19	: J	Dennis Su	0.8	959		1.00	2009-03-11	09/41/54

Training Data

100 million ratings



Netflix Scoring and Judging

- Leaders <u>on test set</u> are contacted and submit their code and documentation (mid-August)
- Judges review documentation and inform winners that they have won \$1 million prize (late August)
- Considerable speculation in press and blogs about which team has actually won
- News conference scheduled for Sept 21st in New York to announce winner and present \$1 million check

NETFLIX

Rules

Netflix Prize

Home

Leaderboard

Download

Update

Leaderboard

Showing Test Score. Click here to show quiz score

COMPLETED

Display top 20 \$ leaders.

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time			
<u>Grand Prize</u> - RMSE = 0.8567 - Winning Team: BellKorle Pragmatic Chaos							
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28			
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22			
3	Grand Prize Team	0.0032	9.90	2009-07-10 21:24:40			
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31			
5	Vandelay Industries !	0.8591	9.81	2009-07-10 00:32:20			
6	PragmaticTheory	0.8594	9.77	2009-06-24 12:06:56			
7	BellKor in BigChaos	0.8601	9.70	2009-05-13 08:14:09			
8	Dace_	0.8612	9.59	2009-07-24 17:18:43			
9	Feeds2	0.8622	9.48	2009-07-12 13:11:51			
10	BigChaos	0.8623	9.47	2009-04-07 12:33:59			
11	Opera Solutions	0.8623	9.47	2009-07-24 00:34:07			
12	BellKor	0.8624	9.46	2009-07-26 17:19:11			
Progr	<u>ess Prize 2008</u> - RMSE = 0.8627 - Wi	nning Team: BellKo	r in BigChaos				
13	xiangliang	0.8642	9.27	2009-07-15 14:53:22			
14	Gravity	0.8643	9.26	2009-04-22 18:31:32			
15	Ces	0.8651	9.18	2009-06-21 19:24:53			
16	Invisible Ideas	0.8653	9.15	2009-07-15 15:53:04			
17	Just a guy in a garage	0.8662	9.06	2009-05-24 10:02:54			
18	<u>J Dennis Su</u>	0.8666	9.02	2009-03-07 17:16:17			
19	Craig Carmichael	0.8666	9.02	2009-07-25 16:00:54			
20	acmehill	0.8668	9.00	2009-03-21 16:20:50			

Progress Prize 2007 - RMSE = 0.8723 - Winning Team: KorBell

NETFLIX

Rules

Netflix Prize

Home

Leaderboard

Download

Update

Leaderboard

Showing Test Score. Click here to show quiz score

COMPLETED

Display top 20 \$ leaders.

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time			
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos							
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28			
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22			
3	Grand Prize Team	0.0082	0.00	071:1			
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31			
5	Vandelay Industries !	0.8591	9.81	2009-07-10 00:32:20			
6	PragmaticTheory	0.8594	9.77	2009-06-24 12:06:56			
7	BellKor in BigChaos	0.8601	9.70	2009-05-13 08:14:09			
8	Dace_	0.8612	9.59	2009-07-24 17:18:43			
9	Feeds2	0.8622	9.48	2009-07-12 13:11:51			
10	BigChaos	0.8623	9.47	2009-04-07 12:33:59			
11	Opera Solutions	0.8623	9.47	2009-07-24 00:34:07			
12	BellKor	0.8624	9.46	2009-07-26 17:19:11			
Progress Prize 2008 - RMSE = 0.8627 - Winning Team: BellKor in BigChaos							
13	xiangliang	0.8642	9.27	2009-07-15 14:53:22			
14	Gravity	0.8643	9.26	2009-04-22 18:31:32			
15	Ces	0.8651	9.18	2009-06-21 19:24:53			
16	Invisible Ideas	0.8653	9.15	2009-07-15 15:53:04			
17	Just a guy in a garage	0.8662	9.06	2009-05-24 10:02:54			
18	<u>J Dennis Su</u>	0.8666	9.02	2009-03-07 17:16:17			
19	Craig Carmichael	0.8666	9.02	2009-07-25 16:00:54			
20	acmehill	0.8668	9.00	2009-03-21 16:20:50			

Progress Prize 2007 - RMSE = 0.8723 - Winning Team: KorBell

Million Dollars Awarded Sept 21st 2009

The second secon		
		2009
	NETFLIX	DATE 09.21.09
	PAY TO THE BOUND'S Promotio Chaos	\$ 1 000 000 ⁽¹⁰⁾
ALL COLLECTION OF	cente de La Chinar à réglitaire cristos	
Statistics of the	AMOUNT ONE MILLION	00/100
	EDA The Notflix Prize Red 7	factings

Lessons Learned

Who were the Real Winners?

- Winning team
 - BellKor: 2 statisticians + computer scientist (AT&T, Yahoo!; US)
 - BigChaos: 2 computer science masters students (Austria)
 - Pragmatic Theory: 2 electrical engineers (Canada)
 - Division of prize money within team not revealed
- Netflix
 - Publicity
 - New algorithms
 - More research on recommender systems
- Machine learning/data mining research community
 - Increased interest in the field
 - Large new data set
 - Interest in more competitions
Lessons Learned

- Scale is important
 - e.g., stochastic gradient descent on sparse matrices
- Latent factor models work well on this problem
 - Previously had not been explored for recommender systems
- Understanding your data is important, e.g., time-effects
- Combining models works surprisingly well
 - But final 10% improvement can probably be achieved by judiciously combining about 10 models rather than 1000's
 - This is likely what Netflix will do in practice
- Surprising amount of collaboration among participants

The New York Times

Netflix Competitors Learn the Power of Teamwork

By STEVE LOHR Published: July 27, 2009

A contest set up by <u>Netflix</u>, which offered a <u>\$1 million prize</u> to anyone who could significantly improve its movie recommendation system, ended on Sunday with two teams in a virtual dead heat, and no winner to be declared until September.

Enlarge This Image



Ozier Muhammad/The New York Times

Chris Volinsky, a scientist at AT&T Research, left, is on a high-ranking team in a Netflix contest. With him is Robert Bell.

Related

The Screens Issue: If You Liked This, You're Sure to Love That (November 23, 2008)

Times Topics: Netflix Inc.

But the contest, which began in October 2006, has already produced an impressive legacy. It has shaped careers, spawned at least one start-up company and inspired research papers. It has also changed conventional wisdom about the best way to build the automated systems that increasingly help people make online choices about movies, books, clothing, restaurants, news and other goods and services.

These so-called recommendation engines are computing models that predict what a person might enjoy based on statistical scoring of that person's stated preferences, past consumption patterns and similar choices made by many others — all made possible by the ease of data collection and tracking on the Web. Openness of competition structure

- Rules stated that winning solutions would be published
 - Non-exclusive license of winning software to Netflix
 - "Description of algorithm to be posted on site"
- Research workshops sponsored by Netflix
- Leaderboard was publicly visible: "it was addictive...."

Netflix Prize

Home Rules Leaderboard Update Download

Netflix Prize: Forum

Forum for discussion about the Netflix Prize and dataset.

Index Userlist Rules Search Register Login

You are not logged in.

Announcement

Congratulations to team "BellKor's Pragmatic Chaos" for being awarded the \$1M Grand Prize on September 21, 2009. Stay tuned for details of the next contest, Netflix Prize 2.

COMPLETED

Administrivia					
	Forum	Topics	Posts	Last post	
	Important Announcements	5	151	Today 04:29:38 by YehudaKoren	
11	Registration Problems	1	1	2006-10-05 08:37:53 by prizemaster	
11	Administrivia Administrative notes from the maintainers	з	43	2009-06-22 09:23:04 by dale5351	
11	Prize and Forum FAQ	15	18	2009-03-24 10:18:36 by prizemaster	
11	Request for new Category or Forum Want to add a new high-level Category or Forum? This is the place to ask or comment.	18	40	2008-04-29 20:50:19 by filmmakershelp	
Awarded Prizes					
Aw	arded Prizes				
Aw	arded Prizes Forum	Topics	Posts	Last post	
Aw	arded Prizes Forum Grand Prize	Topics 1	Posts 14	Last post 2009-10-09 12:18:23 by statistician	
Aw	arded Prizes Forum Grand Prize Progress Prize 2008	Topics 1 2	Posts 14 17	Last post 2009-10-09 12:18:23 by statistician 2009-03-18 02:40:53 by CS1	
Aw	arded Prizes Forum Grand Prize Progress Prize 2008 Progress Prize 2007	Topics 1 2 5	Posts 14 17 29	Last post 2009-10-09 12:18:23 by statistician 2009-03-18 02:40:53 by CS1 2008-10-06 06:51:51 by dinc3r	
Aw	arded Prizes Forum Grand Prize Progress Prize 2008 Progress Prize 2007 estions (and answers)	Topics 1 2 5	Posts 14 17 29	Last post 2009-10-09 12:18:23 by statistician 2009-03-18 02:40:53 by CS1 2008-10-06 06:51:51 by dinc3r	

Development of Online Community

- Active Netflix prize forum + other blogs
- Quickly acquired "buzz"
- Forum was well-moderated by Netflix
- Attracted discussion from novices and experts alike
- Early posting of code and solutions
- Early self-identification (links via leaderboard)

Academic/Research Culture

- Nature of competition was technical/mathematical
- Attracted students, hobbyists, researchers
- Many motivated by fundamental interest in producing better algorithms - \$1 million would be a nice bonus
- History in academic circles of being open, publishing, sharing

Technical Reasons

 Realization that combining many different models and techniques always produced small but systematic improvements

(Statistical theory supports this....)

- "Teaming" was strategically attractive
- Particularly for the "end-game" (summer 2009), teaming was quite critical in terms of who won the competition

Questions

 Does reduction in squared error metric correlate with real improvements in user satisfaction?

- Are these competitions good for scientific research?
 - Should researchers be solving other more important problems?
- Are competitions a good strategy for companies?

Where does a 5-star movie get ranked on average? From Y. Koren, ACM SIGKDD 2008 1 0.9 0.8 0.7 Probability MovieAvg (RMSE=1.0534) of 0.6 CorNgbr (RMSE=0.9406) Rank 0.5 WgtNgbr (RMSE=0.9107) SVD (RMSE=0.9025) 0.4 integrated (RMSE=0.8870) 0.3 0.2 0.1 0 0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00%

Rank of best recommendation

Where does a 5-star movie get ranked on average? From Y. Koren, ACM SIGKDD 2008 1 0.9 0.8 0.7 Probability MovieAvg (RMSE=1.0534) of 0.6 CorNgbr (RMSE=0.9406) Rank 0.5 WgtNgbr (RMSE=0.9107) SVD (RMSE=0.9025) 0.4 integrated (RMSE=0.8870) 0.3 0.2 0.1 0 0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00% ~0.00%

Rank of best recommendation



Conclusions

- Was the Netflix prize a success?
 - For Netflix?
 - Publicity
 - Algorithms
 - For Participants?
 - For Research Community?
 - For recommender systems in general?

THE WALL STREET JOURNAL. BLOGS



DECEMBER 18, 2009, 4:32 PM ET

Did Netflix Violate Subscribers' Privacy? Lawsuit Says Yes

Article Comments (26)		LAV
🖾 Email 🗏 Printer Friendly 🥜 Permalink Share: 📑 facebook	▼ - Text Size ◆	Thora
By Ashby Jones A potentially very interesting lawsuit was filed in San Jose, Calif., on Thursday concerning a closeted lesbian's privacy and Netflix, the movie rental company.		one thing
The woman, an Ohio resident who, for purposes of the lawsuit has requested anonymity, sued Netflix alongside others hoping to sue on behalf of a potential class, claiming that the company divulged personal information which allowed thousands to discern her sexual preference. Click <u>here</u> for the story, from Wired; <u>here</u> for the complaint.	NETFLIX	agree on.

Here's the backstory. In 2006, Netflix initiated a contest in which it would pay \$1 million to the first person or team of people to make certain improvements to its recommendation system —

Links to additional information

 Netflix prize page (FAQs, rules, forum, etc) http://www.netflixprize.com/

 Page with links to articles, blogs, etc http://www.research.att.com/~volinsky/netflix/bpc.html