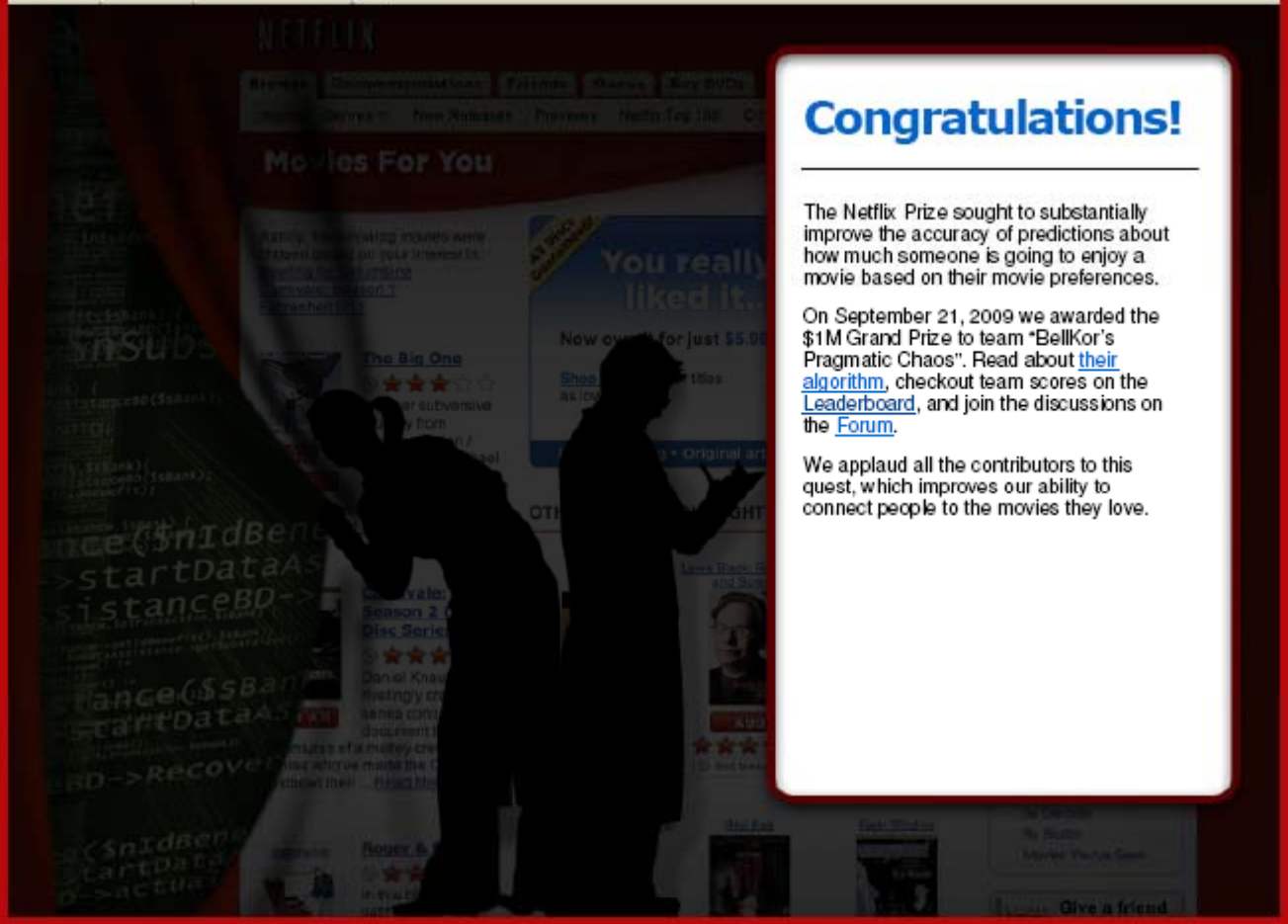
NETFLIX

Netflix Prize

COMPLETED

Home | Rules | Leaderboard | Update



Congratulations!

The Netflix Prize sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences.

On September 21, 2009 we awarded the $1M Grand Prize to team "BellKor's Pragmatic Chaos". Read about their algorithm, checkout team scores on the Leaderboard, and join the discussions on the Forum.

We applaud all the contributors to this quest, which improves our ability to connect people to the movies they love.

FAQ | Forum | Netflix Home

The Netflix Prize Contest

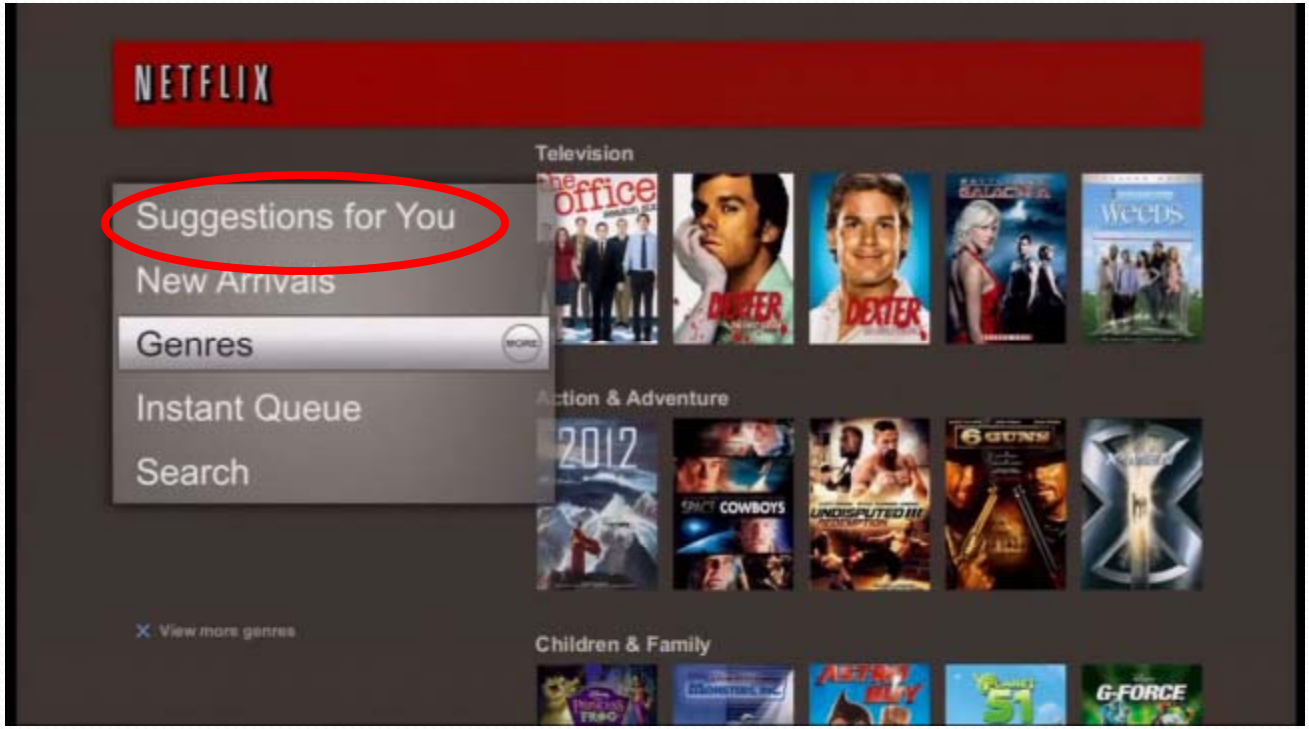
1) New Paths to New Machine Learning Science

2) How an Unruly Mob Almost Stole the Grand Prize at the Last Moment

Jeff Howbert

February 6, 2012

Netflix Viewing Recommendations



Recommender Systems

DOMAIN: some field of activity where *users* buy, view, consume, or otherwise experience *items*

PROCESS:

1. *users* provide *ratings* on *items* they have experienced
2. Take all < *user*, *item*, *rating* > data and build a predictive model
3. For a *user* who hasn't experienced a particular *item*, use model to *predict* how well they will like it (i.e. *predict rating*)

Roles of Recommender Systems

- Help users deal with *paradox of choice*
- Allow online sites to:
  - Increase likelihood of sales
  - Retain customers by providing positive search experience
- Considered essential in operation of:
  - Online retailing, e.g. Amazon, Netflix, etc.
  - Social networking sites

Amazon.com Product Recommendations

Customers Who Bought This Item Also Bought

OtterBox Impact Case for iPhone 3G, 3GS (White)

★★★★☆ (218)

Click to see price

5-Pack Premium Reusable LCD Screen Protector with Lint Cleaning...

★★★½☆ (258)

$1.18

5-Pack Premium Reusable LCD Mirror Screen Protector with Lint Cl...

★★★½☆ (91)

$2.27

Car Charger for Apple 3G iPhone, Black

★★★☆☆ (179)

$2.67

Social Network Recommendations

- Recommendations on essentially every category of interest known to mankind
  - Friends
  - Groups
  - Activities
  - Media (TV shows, movies, music, books)
  - News stories
  - Ad placements
- All based on connections in underlying social network graph and your expressed ‘likes’ and ‘dislikes’

Types of Recommender Systems

Base predictions on either:

- content-based approach
  - *explicit* characteristics of users and items
- collaborative filtering approach
  - *implicit* characteristics based on similarity of users' preferences to those of other users

The Netflix Prize Contest

- *GOAL*: use *training data* to build a recommender system, which, when applied to *qualifying data*, improves error rate by 10% relative to Netflix's existing system

- *PRIZE*: first team to 10% wins $1,000,000
  - Annual Progress Prizes of $50,000 also possible

The Netflix Prize Contest

- *CONDITIONS*:
  - Open to public
  - Compete as individual or group
  - Submit predictions no more than once a day
  - Prize winners must publish results and license code to Netflix (non-exclusive)
- *SCHEDULE*:
  - Started Oct. 2, 2006
  - To end after 5 years

The Netflix Prize Contest

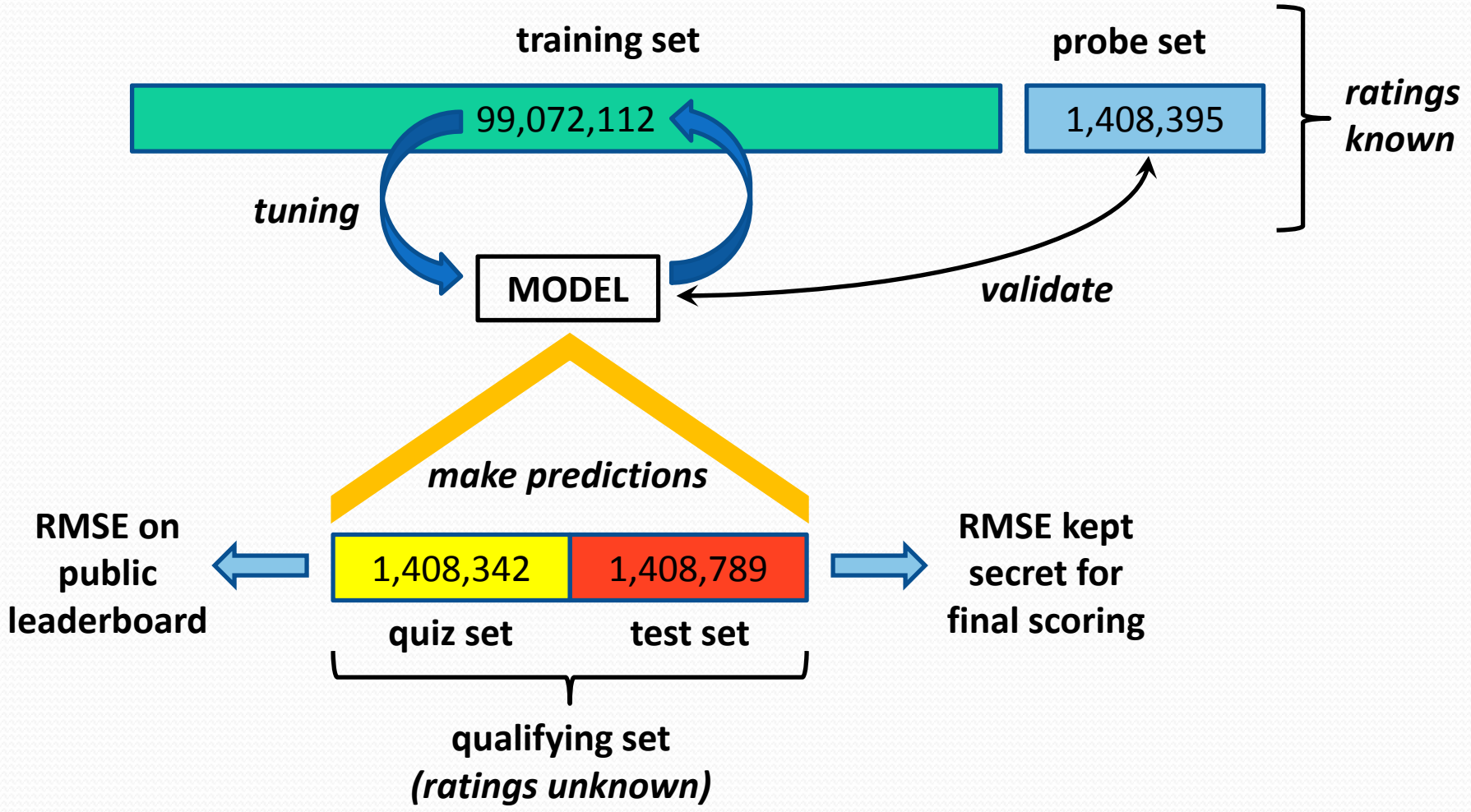
- *PARTICIPATION*:
  - 51051 contestants on 41305 teams from 186 different countries
  - 44014 valid submissions from 5169 different teams

The Netflix Prize Data

- Netflix released three datasets
  - 480,189 *users* (anonymous)
  - 17,770 *movies*
  - *ratings* on integer scale 1 to 5

- *Training set*: 99,072,112 < *user, movie* > pairs with *ratings*
- *Probe set*: 1,408,395 < *user, movie* > pairs with *ratings*
- *Qualifying set* of 2,817,131 < *user*, *movie* > pairs with *no ratings*

Model Building and Submission Process



Netflix Prize

Home Rules Leaderboard Register Update Submit Download

Leaderboard

Display top 100 leaders.

Rank	Team Name	Best Score	% Improvement	Last Submit Time
--	No Grand Prize candidates yet	-	--	--
	Grand Prize - RMSE <= 0.8563			
1	BellKor in BigChaos	0.8604	9.56	2008-12-03 16:46:15
	Progress Prize - RMSE <= 0.8625			
2	BigChaos	0.8626	9.33	2008-12-04 19:18:27
3	BellKor	0.8630	9.29	2008-12-04 19:25:59
4	PragmaticTheory	0.8638	9.21	2008-11-28 11:46:23
5	Gravity	0.8654	9.04	2008-11-27 21:18:37
6	My Brain and His Chain	0.8668	8.89	2008-09-30 02:19:47
7	Just a guy in a garage	0.8672	8.85	2008-12-07 06:51:12
8	When Gravity and Dinosaurs Unite	0.8675	8.82	2008-10-05 14:16:53
9	Opera Solutions	0.8676	8.81	2008-12-02 22:08:45
10	acmehill	0.8677	8.80	2008-12-05 08:01:00
11	scientist	0.8677	8.80	2008-12-02 01:10:13
12	Ces	0.8711	8.44	2008-08-25 05:00:23
13	Dace	0.8711	8.44	2008-12-07 03:46:04
	Progress Prize 2007 - RMSE = 0.8712 - Winning Team: KorBell			
14	KorBell	0.8712	8.43	2007-10-01 23:25:23
15	basho	0.8714	8.41	2008-05-21 22:06:00
16	pengpengzhou	0.8714	8.41	2008-11-05 01:11:13
17	blednotik	0.8717	8.38	2008-11-26 00:12:12

Why the Netflix Prize Was Hard

- Massive dataset
- Very sparse – matrix only 1.2% occupied
- Extreme variation in number of ratings per user
- Statistical properties of qualifying and probe sets different from training set

	movie 1	movie 2	movie 3	movie 4	movie 5	movie 6	movie 7	movie 8	movie 9	movie 10	…	movie 17770
user 1			1		2							3
user 2		2		3	3			4				
user 3							5	3		4		
user 4	2				3			2				2
user 5		4				5			3			4
user 6			2									
user 7			2					4	2	3		
user 8	3	4				4						
user 9									3			
user 10			1		2							2
…												
user 480189		4			3			3				

Dealing with Size of the Data

- MEMORY:
  - 2 GB bare minimum for common algorithms
  - 4+ GB required for some algorithms
  - need 64-bit machine with 4+ GB RAM if serious
- SPEED:
  - Program in languages that compile to fast machine code
  - 64-bit processor
  - Exploit low-level parallelism in code (SIMD on Intel x86/x64)

Common Types of Algorithms

- Global effects
- Nearest neighbors
- Matrix factorization
- Restricted Boltzmann machine
- Clustering
- Etc.

Nearest Neighbors in Action

	movie 1	movie 2	movie 3	movie 4	movie 5	movie 6	movie 7	movie 8	movie 9	movie 10	...	movie 17770
user 1			1		2							3
user 2		2		3	3			4		?		
user 3							5	3				
user 4	2				3			2				2
user 5		2		3		5		4		2		4
user 6			2									
user 7			2					4	2			
user 8	3	1			3	4		5		4		
user 9									3			
user 10			1		2							2
...												
user 480189		4			3			3				

Identical preferences – strong weight

Similar preferences – moderate weight

Matrix Factorization in Action

	movie 1	movie 2	movie 3	movie 4	movie 5	movie 6	movie 7	movie 8	movie 9	movie 10	..	movie 17770
user 1			1		2							3
user 2		2		3	3			4				
user 3							5	3		4		
user 4	2				3			2				2
user 5		4				5			3			4
user 6			2									
user 7			2					4	2	3		
user 8	3	4				4						
user 9									3			
user 10			1		2							2
...												
user 480189		4			3			3				

reduced-rank singular value decomposition (sort of)

	movie 1	movie 2	movie 3	movie 4	movie 5	movie 6	movie 7	movie 8	movie 9	movie 10	..	movie 17770
feature 1												
feature 2												
feature 3					*< a bunch of numbers >*							
feature 4												
feature 5												

+

	feature 1	feature 2	feature 3	feature 4	feature 5
user 1					
user 2					
user 3					
user 4					
user 5					
user 6					
user 7					
user 8			*< a bunch of numbers >*		
user 9					
user 10					
...					
user 480189					

Matrix Factorization in Action

	movie 1	movie 2	movie 3	movie 4	movie 5	movie 6	movie 7	movie 8	movie 9	movie 10	...	movie 17770
feature 1												
feature 2												
feature 3												
feature 4												
feature 5												

+

	feature 1	feature 2	feature 3	feature 4	feature 5
user 1					
user 2					
user 3					
user 4					
user 5					
user 6					
user 7					
user 8					
user 9					
user 10					
...					
user 480189					

multiply and add features (dot product) for desired < *user*, *movie* > prediction

	movie 1	movie 2	movie 3	movie 4	movie 5	movie 6	movie 7	movie 8	movie 9	movie 10	...	movie 17770
user 1			1		2							3
user 2		2		3	3			4				
user 3							5	3		4		
user 4	2				3			2				2
user 5		4				5			3			4
user 6			2									
user 7			2					4	2	3		
user 8	3	4				4	?					
user 9									3			
user 10			1		2							2
...												
user 480189		4			3			3				

The Power of Blending

- Error function (RMSE) is convex, so linear combinations of models should have lower error
- Find blending coefficients with simple least squares fit of model predictions to true values of probe set
- Example from my experience:
  - blended 89 diverse models
    - RMSE range = 0.8859 – 0.9959
  - blended model had RMSE = 0.8736
    - Improvement of 0.0123 over best single model
    - 13% of progress needed to win

Algorithms: Other Things That Mattered

- Overfitting
  - Models typically had millions or even billions of parameters
  - Control with aggressive *regularization*
- Time-related effects
  - Netflix data included date of movie release, dates of ratings
  - Most of progress in final two years of contest was from incorporating temporal information

The Netflix Prize: Social Phenomena

- Competition intense, but sharing and collaboration were equally so
  - Lots of publications and presentations at meetings while contest still active
  - Lots of sharing on contest forums of ideas and implementation details
- Vast majority of teams:
  - Not machine learning professionals
  - Not competing to win (until very end)
  - Mostly used algorithms published by others

One Algorithm from Winning Team

(time-dependent matrix factorization)

This leads to the prediction rule

$$\hat{r}_{ui} = \mu + b_i(t_{ui}) + b_u(t_{ui}) + |\mathrm{R}(u)|^{-\frac{1}{2}} \sum_{j \in \mathrm{R}(u)} e^{-\beta_u \cdot |t_{ui} - t_{uj}|}((r_{uj} - b_{uj})w_{ij} + c_{ij}). \quad (15)$$

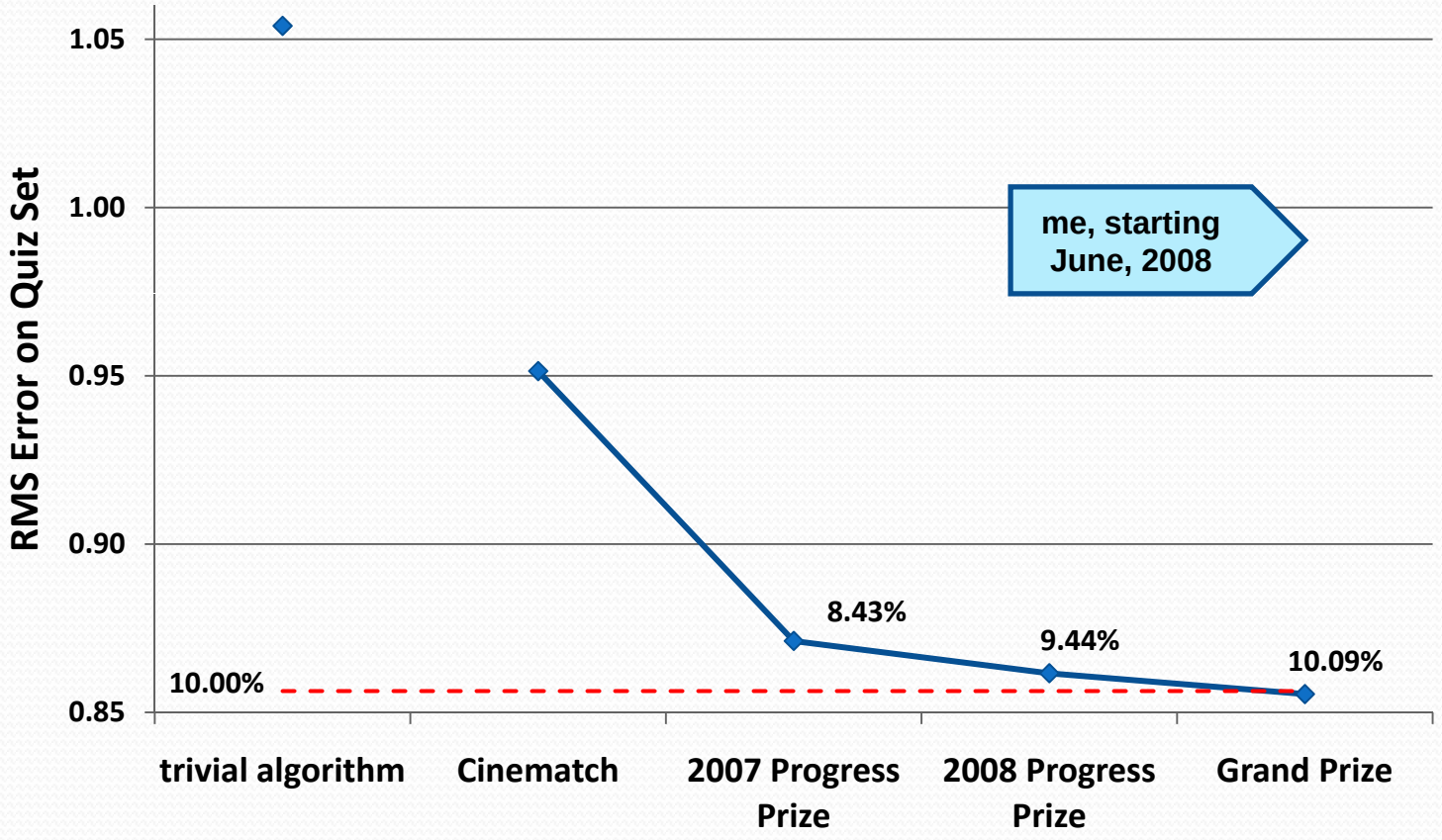
The involved parameters, $b_i(t_{ui}) = b_i + b_{i,\mathrm{Bin}(t_{ui})}$, $b_u(t_{ui}) = b_u + \alpha_u \cdot \mathrm{dev}_u(t_{ui}) + b_{u,t_{ui}}$, β_u, w_{ij} and c_{ij}, are learned by minimizing the associated regularized squared error

$$\sum_{(u,i) \in \kappa} \Big(r_{ui} - \mu - b_i - b_{i,\mathrm{Bin}(t_{ui})} - b_u - \alpha_u \mathrm{dev}_u(t_{ui}) - b_{u,t_{ui}} - |\mathrm{R}(u)|^{-\frac{1}{2}} \sum_{j \in \mathrm{R}(u)} e^{-\beta_u \cdot |t_{ui} - t_{uj}|}((r_{uj} - b_{uj})w_{ij} + c_{ij})\Big)^2 + \lambda_{12}(b_i^2 + b_{i,\mathrm{Bin}(t_{ui})}^2 + b_u^2 + \alpha_u^2 + b_{ut}^2 + w_{ij}^2 + c_{ij}^2). \quad (16)$$

Minimization is performed by stochastic gradient descent.

Yehuda Koren, *Comm. ACM*, **53**, 89 (2010)

Netflix Prize Progress: Major Milestones



DATE:	*Oct. 2007*	*Oct. 2008*	*July 2009*
WINNER:	*BellKor*	*BellKor in BigChaos*	*???*

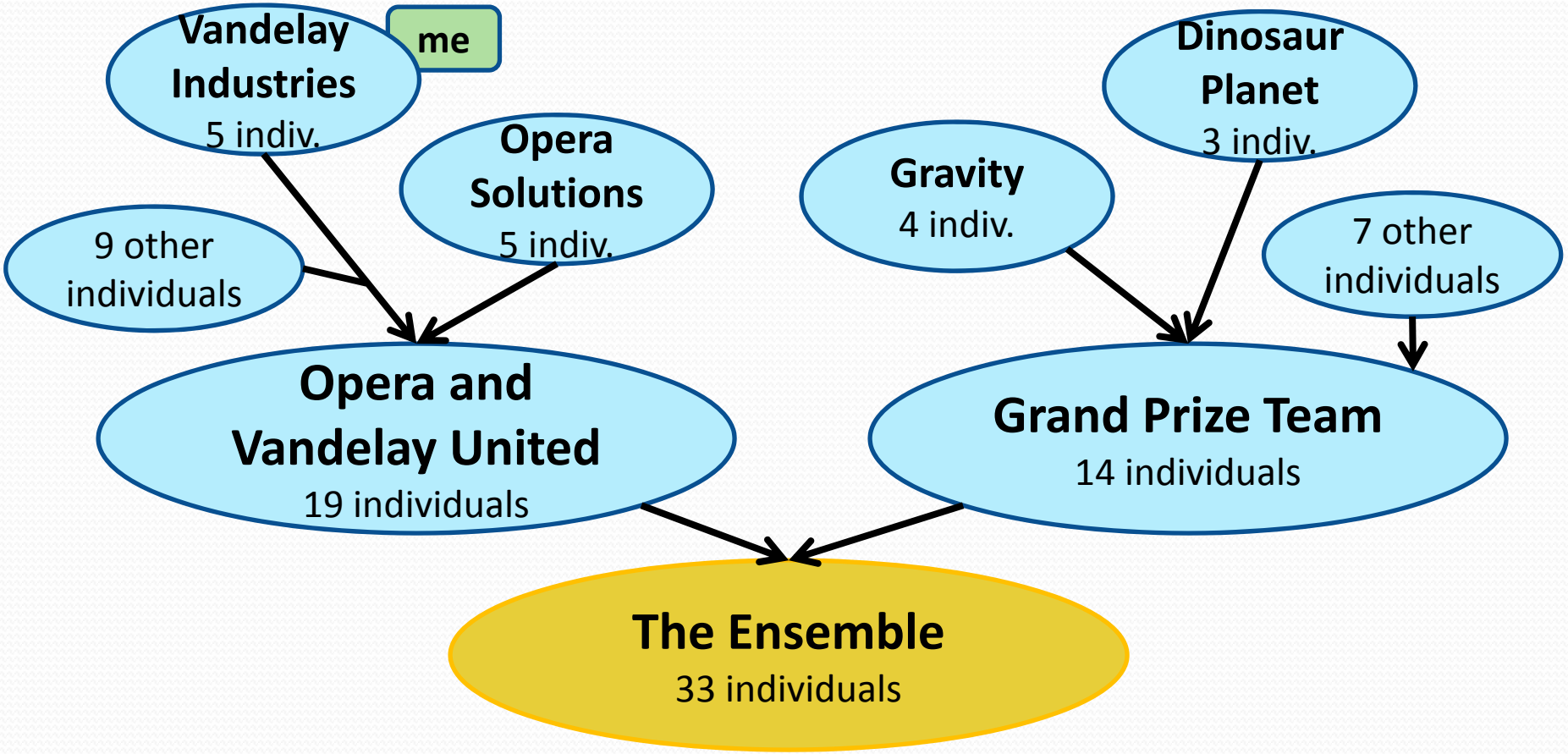
June 25, 2009 20:28 GMT

Rank	Team Name	Best Score	% Improvement	Last Submit Time
--	No Grand Prize candidates yet	--	--	--
	Grand Prize - RMSE <= 0.8563			
1	PragmaticTheory	0.8582	9.80	2009-06-24 12:06:56
2	BellKor in BigChaos	0.8590	9.71	2009-05-13 08:14:09
3	Grand Prize Team	0.8593	9.68	2009-06-12 08:20:24
4	Dace	0.8604	9.56	2009-04-22 05:57:03
5	BigChaos	0.8613	9.47	2009-06-23 23:06:52
	Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos			
6	BellKor	0.8620	9.40	2009-06-24 07:16:02
7	Gravity	0.8634	9.25	2009-04-22 18:31:32
8	Opera Solutions	0.8638	9.21	2009-06-22 05:53:30
9	xlvector	0.8639	9.20	2009-06-25 11:53:49
10	xiangliang	0.8639	9.20	2009-06-23 14:22:37
11	BruceDengDaoCiYiYou	0.8641	9.18	2009-06-02 17:08:31
12	Ces	0.8642	9.17	2009-06-24 14:34:14
13	majia2	0.8642	9.17	2009-06-23 08:07:50
14	Feeds2	0.8647	9.11	2009-06-16 22:21:19
15	Just a guy in a garage	0.8650	9.08	2009-05-24 10:02:54
16	Team ESP	0.8651	9.07	2009-06-25 16:29:45
17	pengpengzhou	0.8654	9.04	2009-05-05 18:18:03
18	NewNetflixTeam	0.8657	9.01	2009-05-31 07:30:22
19	J Dennis Su	0.8658	9.00	2009-03-11 09:41:54
20	Vandelay Industries !	0.8658	9.00	2009-05-11 00:43:14

June 26, 18:42 GMT – BPC Team Breaks 10%

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	0.8592	9.69	2009-06-27 20:02:29
5	Dace	0.8604	9.56	2009-06-27 22:59:46
6	BigChaos	0.8613	9.47	2009-06-23 23:06:52
	Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos			
7	BellKor	0.8620	9.40	2009-06-24 07:16:02
8	Solar System	0.8629	9.30	2009-06-27 22:55:42
9	xlvector	0.8630	9.29	2009-06-27 14:08:39
10	Gravity	0.8634	9.25	2009-04-22 18:31:32
11	Opera Solutions	0.8638	9.21	2009-06-26 23:18:13
12	BruceDengDaoCiYiYou	0.8638	9.21	2009-06-27 00:55:55
13	pengpengzhou	0.8638	9.21	2009-06-27 01:06:43
14	Feeds2	0.8641	9.18	2009-06-26 22:51:55
15	Ces	0.8642	9.17	2009-06-24 14:34:14
16	majia2	0.8642	9.17	2009-06-23 08:07:50
17	Just a guy in a garage	0.8650	9.08	2009-05-24 10:02:54
18	Team ESP	0.8650	9.08	2009-06-27 17:01:43
19	NewNetflixTeam	0.8657	9.01	2009-05-31 07:30:22
20	J Dennis Su	0.8658	9.00	2009-03-11 09:41:54

Genesis of The Ensemble



www.the-ensemble.com

June 30, 16:44 GMT

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	Grand Prize Team	0.8578	9.84	2009-06-30 15:26:35
3	PragmaticTheory	0.8582	9.80	2009-06-25 22:15:51
4	BellKor in BigChaos	0.8590	9.71	2009-05-13 08:14:09
5	Dace	0.8604	9.56	2009-06-30 01:28:47
6	xlvector	0.8609	9.51	2009-06-29 14:32:33
7	BigChaos	0.8613	9.47	2009-06-23 23:06:52
	Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos			
8	BellKor	0.8620	9.40	2009-06-24 07:16:02
9	Vandelay Industries !	0.8623	9.37	2009-06-30 02:15:06
10	Gravity	0.8634	9.25	2009-04-22 18:31:32
11	Opera Solutions	0.8638	9.21	2009-06-26 23:18:13
12	BruceDengDaoCiYiYou	0.8638	9.21	2009-06-27 00:55:55
13	pengpengzhou	0.8638	9.21	2009-06-27 01:06:43
14	Team ESP	0.8639	9.20	2009-06-30 06:20:30
15	Feeds2	0.8640	9.19	2009-06-29 10:57:13
16	Ces	0.8642	9.17	2009-06-29 03:36:56
17	majia2	0.8642	9.17	2009-06-29 03:25:01
18	Just a guy in a garage	0.8650	9.08	2009-05-24 10:02:54
19	NewNetflixTeam	0.8657	9.01	2009-05-31 07:30:22
20	J Dennis Su	0.8658	9.00	2009-03-11 09:41:54

July 8, 14:22 GMT

Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	BellKor's Pragmatic Chaos	0.8558	10.05	2009-07-07 18:26:16
	Grand Prize - RMSE <= 0.8563			
2	Grand Prize Team	0.8572	9.90	2009-07-07 21:37:25
3	Opera Solutions and Vandelay United	0.8576	9.86	2009-07-07 22:49:58
4	xlvector	0.8579	9.83	2009-07-08 08:36:52
5	Three Kingdoms	0.8581	9.81	2009-07-07 14:53:41
6	PragmaticTheory	0.8582	9.80	2009-07-05 11:02:53
7	Peloponnesian League	0.8582	9.80	2009-07-07 16:02:42
8	Vandelay Industries !	0.8584	9.78	2009-07-08 12:15:35
9	BellKor in BigChaos	0.8590	9.71	2009-07-08 06:55:44
10	Team ESP	0.8598	9.63	2009-07-08 08:03:14
11	BigChaos	0.8613	9.47	2009-06-23 23:06:52
12	Opera Solutions	0.8614	9.46	2009-07-02 17:32:37
13	BellKor	0.8615	9.45	2009-07-06 21:37:37
	Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos			
14	space drop	0.8621	9.39	2009-07-07 18:27:20
15	Feeds2	0.8624	9.35	2009-07-07 23:24:03
16	Gravity	0.8634	9.25	2009-04-22 18:31:32
17	BruceDengDaoCiYiYou	0.8638	9.21	2009-06-27 00:55:55
18	pengpengzhou	0.8638	9.21	2009-06-27 01:06:43
19	majia2	0.8638	9.21	2009-07-07 07:13:18
20	Ces	0.8642	9.17	2009-07-07 03:14:03

July 17, 16:01 GMT

Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	BellKor's Pragmatic Chaos	0.8556	10.07	2009-07-17 10:24:54
	Grand Prize - RMSE <= 0.8563			
2	Grand Prize Team	0.8571	9.91	2009-07-17 12:35:45
3	Opera Solutions and Vandelay United	0.8573	9.89	2009-07-16 23:29:47
4	Vandelay Industries !	0.8579	9.83	2009-07-16 16:42:26
5	PragmaticTheory	0.8582	9.80	2009-07-12 15:09:53
6	BellKor in BigChaos	0.8590	9.71	2009-07-15 13:14:59
7	Dace	0.8605	9.55	2009-07-16 07:28:44
8	BigChaos	0.8613	9.47	2009-06-23 23:06:52
9	Feeds2	0.8613	9.47	2009-07-13 17:17:28
10	BellKor	0.8613	9.47	2009-07-17 09:20:57
11	Opera Solutions	0.8614	9.46	2009-07-02 17:32:37
	Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos			
12	xiangliang	0.8633	9.26	2009-07-17 04:54:04
13	Gravity	0.8634	9.25	2009-07-12 14:11:18
14	Ces	0.8642	9.17	2009-07-11 23:33:26
15	Invisible Ideas	0.8644	9.14	2009-07-17 00:07:55
16	Just a guy in a garage	0.8650	9.08	2009-07-06 16:12:33
17	J Dennis Su	0.8658	9.00	2009-03-11 09:41:54
18	acmehill	0.8659	8.99	2009-04-16 06:29:35
19	MonteCarlo	0.8661	8.97	2009-03-25 15:00:05
20	IDEA2	0.8661	8.97	2009-03-25 15:37:59

July 25, 18:32 GMT – The Ensemble First Appears!

Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	The Ensemble	0.8554	10.09	2009-07-25 18:32:29
2	BellKor's Pragmatic Chaos	0.8555	10.08	2009-07-25 15:53:34
	Grand Prize - RMSE <= 0.8563			
3	Grand Prize Team	0.8571	9.91	2009-07-24 13:07:49
4	Opera Solutions and Vandelay United	0.8573	9.89	2009-07-25 20:05:52
5	Vandelay Industries !	0.8579	9.83	2009-07-24 05:26:13
6	PragmaticTheory	0.8582	9.80	2009-07-12 15:09:53
7	BellKor in BigChaos	0.8590	9.71	2009-07-20 19:09:34
8	Dace	0.8603	9.58	2009-07-24 17:18:43
9	Opera Soluti		9	2009-07-25 00:48:38
10	BellKor		8	2009-07-22 20:30:30
11	BigChaos		7	2009-06-23 23:06:52
12	Feeds2		7	2009-07-24 20:06:46
	Progress Prize 200		BigChaos	
13	xiangliang		6	2009-07-21 02:04:40
14	Gravity		5	2009-07-24 16:45:05
15	Ces		7	2009-07-25 17:42:38
16	Invisible Idea		4	2009-07-20 03:26:12
17	Just a guy in		8	2009-07-22 14:10:42
18	Craig Carmichael		9.02	2009-07-25 16:00:54
19	J Dennis Su	0.8658	9.00	2009-03-11 09:41:54
20	acmehill	0.8659	8.99	2009-04-16 06:29:35

24 hours, 10 min before contest ends

#1 and #2 teams each have one more submission !

July 26, 18:18 GMT
BPC Makes Their Final Submission

Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	The Ensemble	0.8554	10.09	2009-07-25 18:32:29
2	BellKor's Pragmatic Chaos	0.8554	10.09	2009-07-26 18:18:28

24 minutes before contest ends

The Ensemble can make one more submission – window opens 10 minutes before contest ends

July 26, 18:43 GMT – Contest Over!

Rank	Team Name	Best Score	% Improvement	Last Submit Time
1	The Ensemble	0.8553	10.10	2009-07-26 18:38:22
2	BellKor's Pragmatic Chaos	0.8554	10.09	2009-07-26 18:18:28
	Grand Prize - RMSE <= 0.8563			
3	Grand Prize Team	0.8571	9.91	2009-07-24 13:07:49
4	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
	Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos			
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	Ces	0.8642	9.17	2009-07-25 17:42:38
16	Invisible Ideas	0.8644	9.14	2009-07-20 03:26:12
17	Just a guy in a garage	0.8650	9.08	2009-07-22 14:10:42
18	Craig Carmichael	0.8656	9.02	2009-07-25 16:00:54
19	J Dennis Su	0.8658	9.00	2009-03-11 09:41:54
20	acmehill	0.8659	8.99	2009-04-16 06:29:35

Final Test Scores

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	0.8582	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
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	J Dennis Su	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				
Cinematch score - RMSE = 0.9525				

Final Test Scores

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	0.8582	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
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	J Dennis Su	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
21	MonteCarlo	0.8669	8.99	2009-03-24 10:45:14
22	IDEA2	0.8669	8.99	2009-03-25 15:37:59
23	just_a_student	0.8675	8.92	2009-07-17 08:37:11
24	Howbert	0.8677	8.90	2009-07-26 07:13:00
25	My Brain and His Chain	0.8678	8.89	2008-09-30 02:19:47

Netflix Prize: What Did We Learn?

- Significantly advanced science of recommender systems
  - Properly tuned and regularized matrix factorization is a powerful approach to collaborative filtering
  - Ensemble methods (blending) can markedly enhance predictive power of recommender systems
- Crowdsourcing via a contest can unleash amazing amounts of sustained effort and creativity
  - Netflix made out like a bandit
  - But probably would not be successful in most problems

Netflix Prize: What Did I Learn?

- Several new machine learning algorithms
- A lot about optimizing predictive models
  - Stochastic gradient descent
  - Regularization
- A lot about optimizing code for speed and memory usage
- Some linear algebra and a little PDQ
- Enough to come up with one original approach that actually worked

- Money makes people crazy, in both good ways and bad

COST: about 1000 hours of my free time over 13 months