Data Mining Techniques

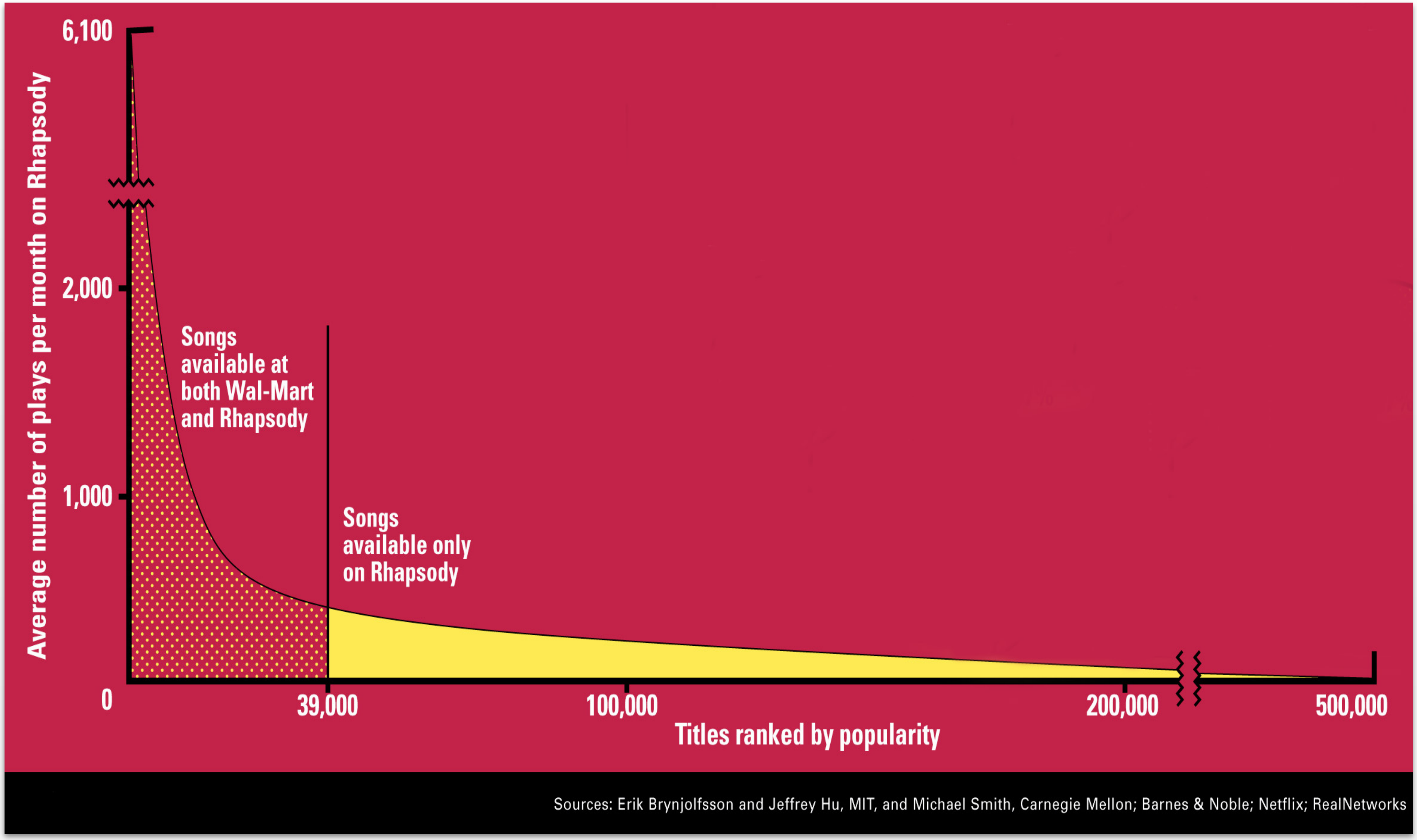
CS 6220 - Section 3 - Fall 2016

Lecture 14

Jan-Willem van de Meent
(credit: Andrew Ng, Alex Smola, Yehuda Koren, Stanford CS246)

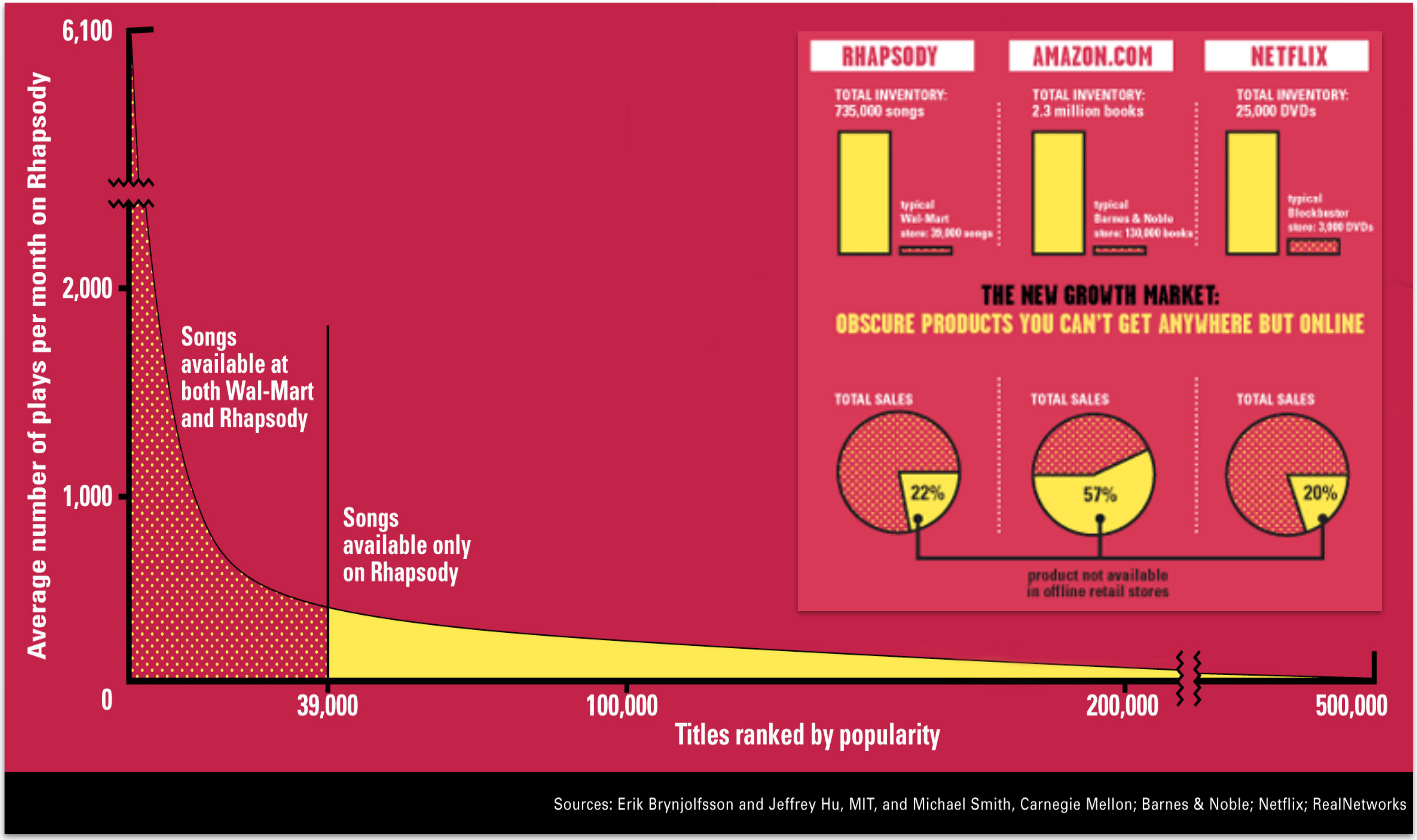
Recommender Systems

The Long Tail



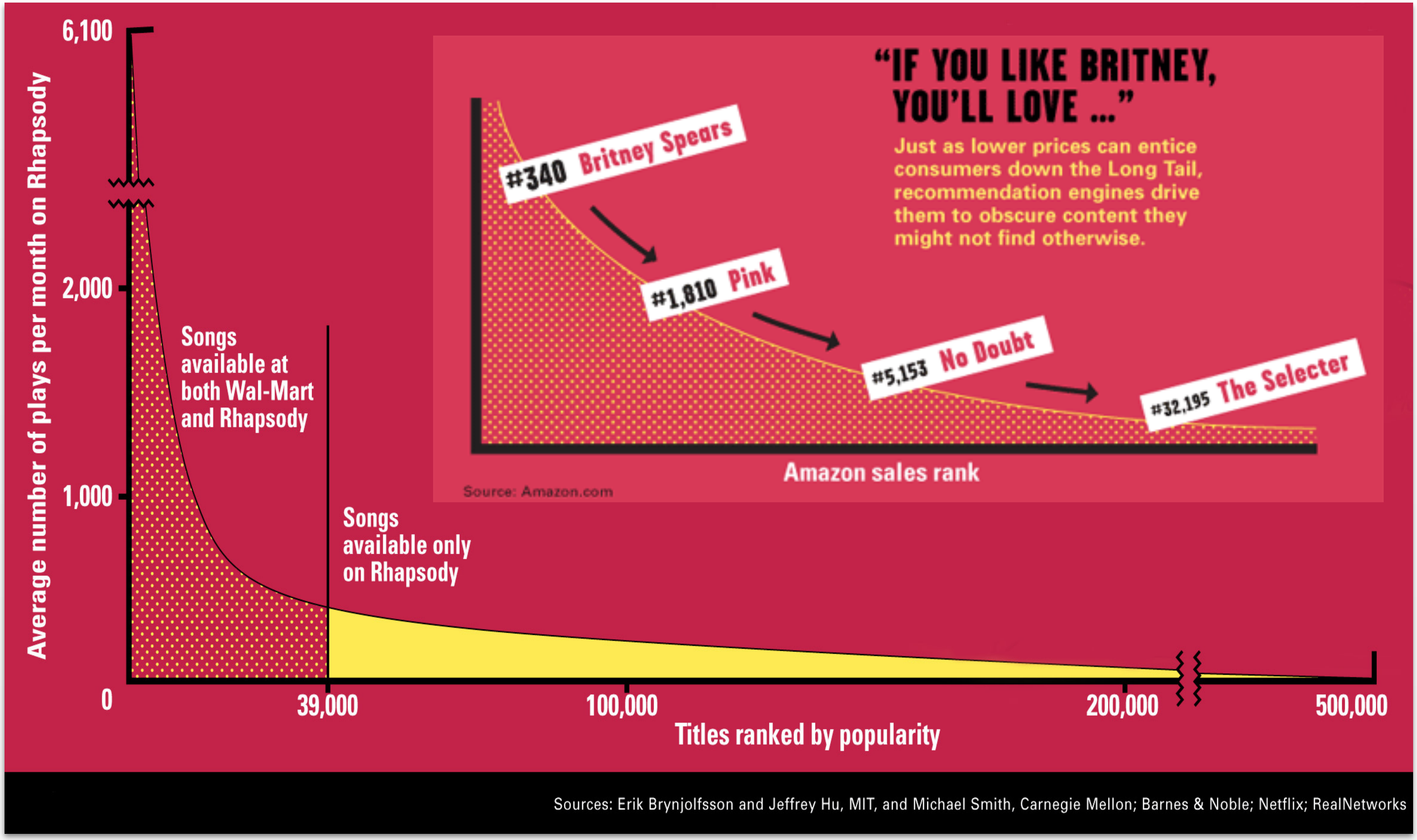
(from: https://www.wired.com/2004/10/tail/)

The Long Tail



(from: https://www.wired.com/2004/10/tail/)

The Long Tail



(from: https://www.wired.com/2004/10/tail/)

Problem Setting

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

Problem Setting

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

Problem Setting

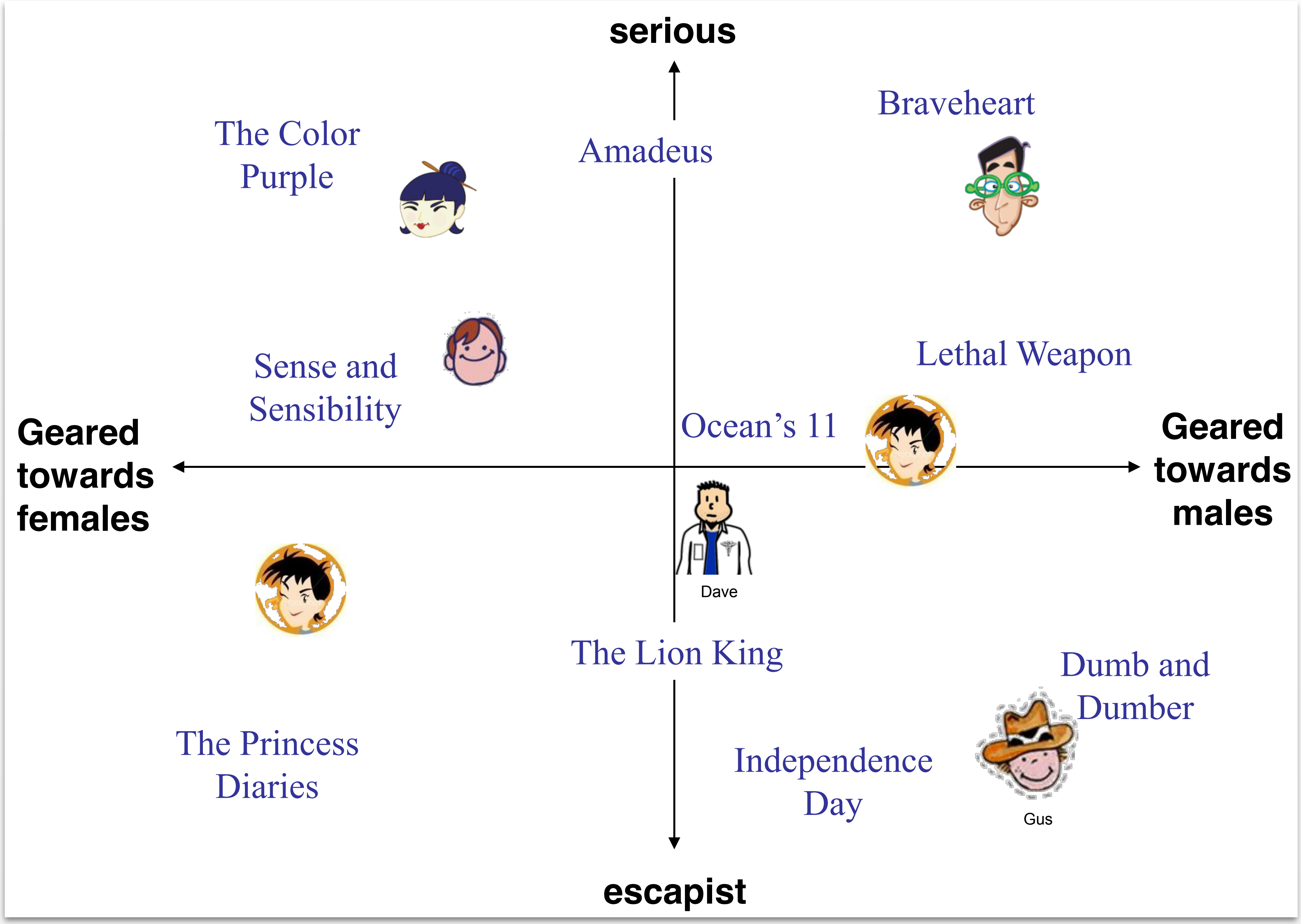
Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

Problem Setting

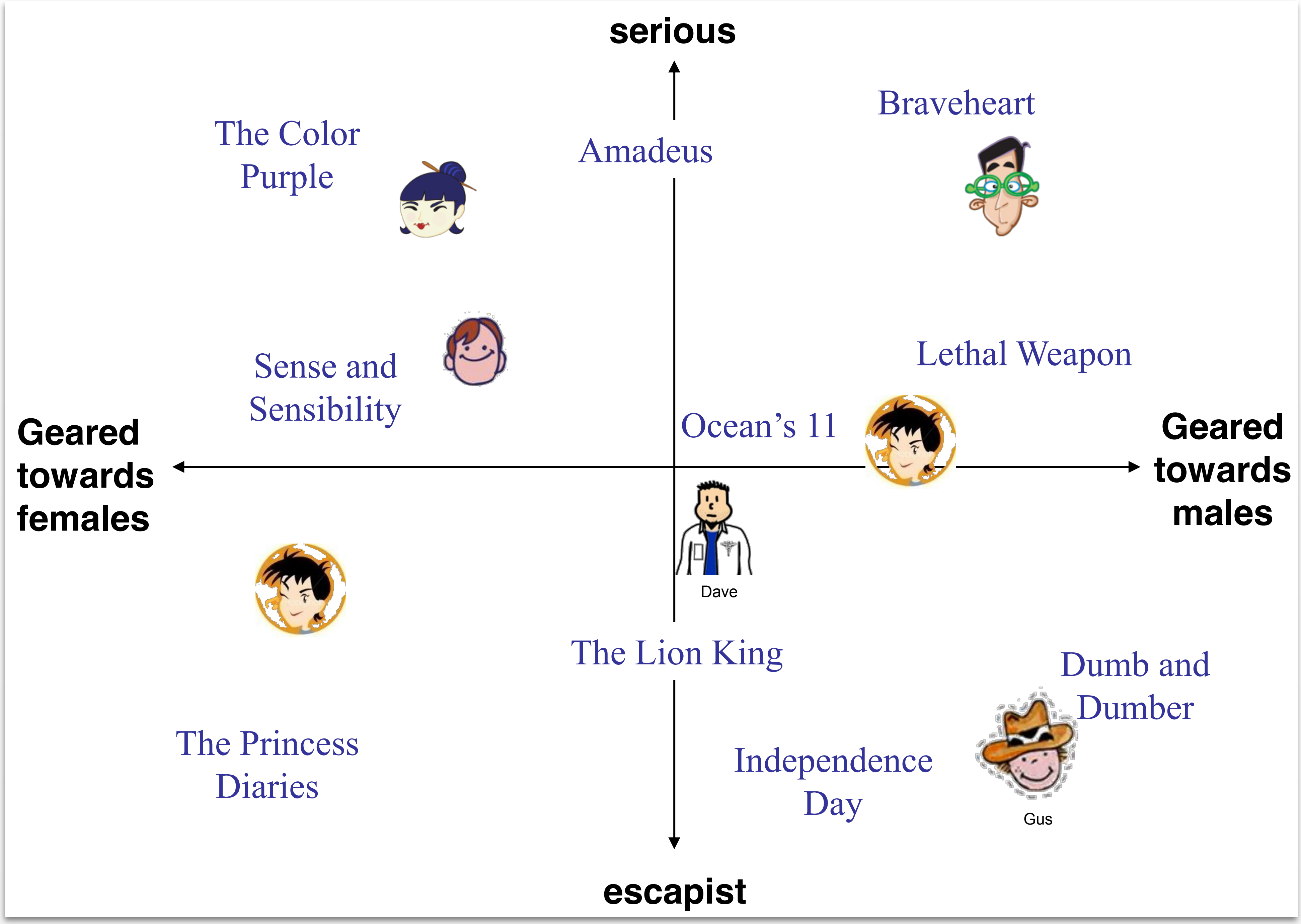
Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

- *Task*: Predict user preferences for unseen items

Content-based Filtering

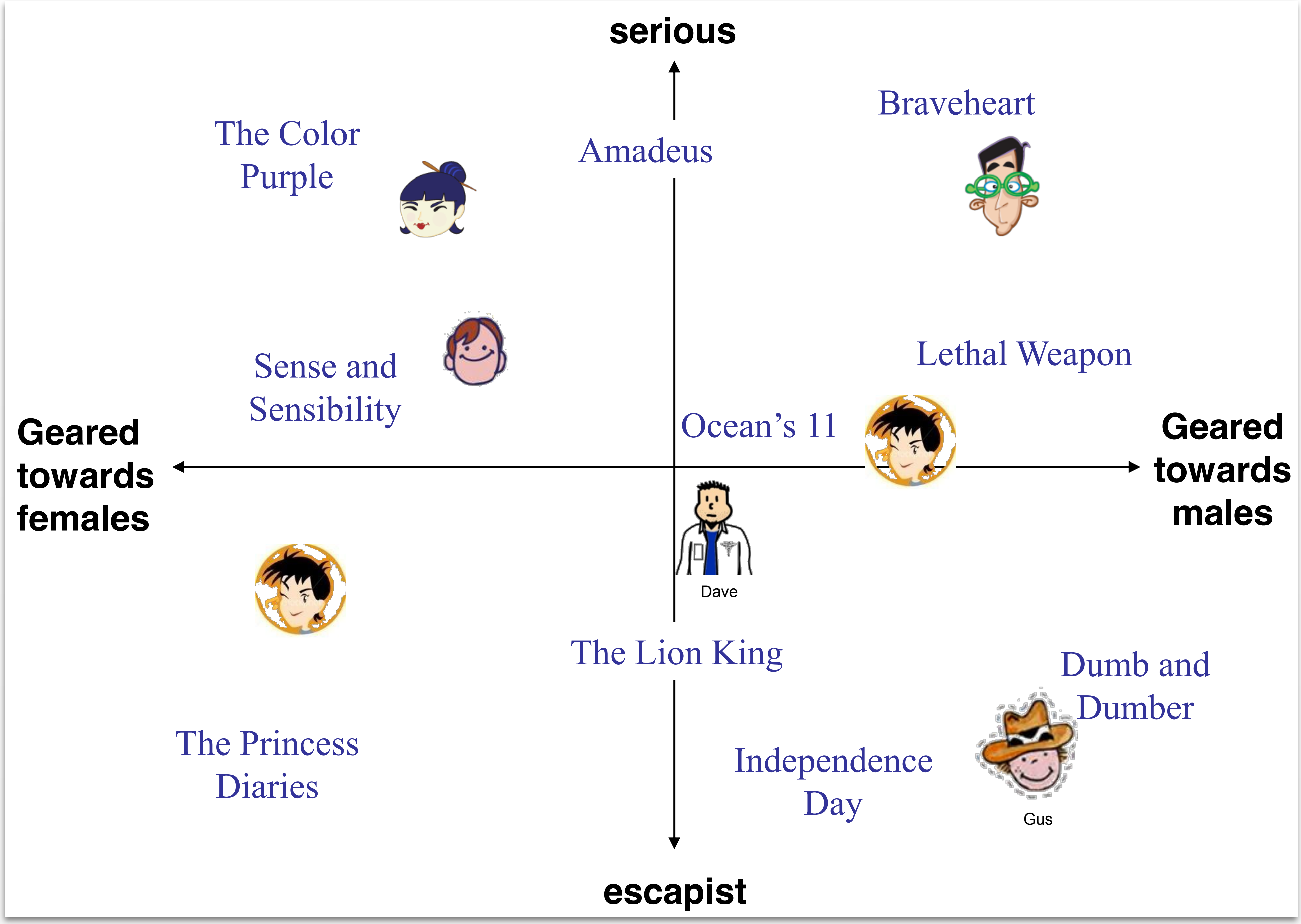


Content-based Filtering



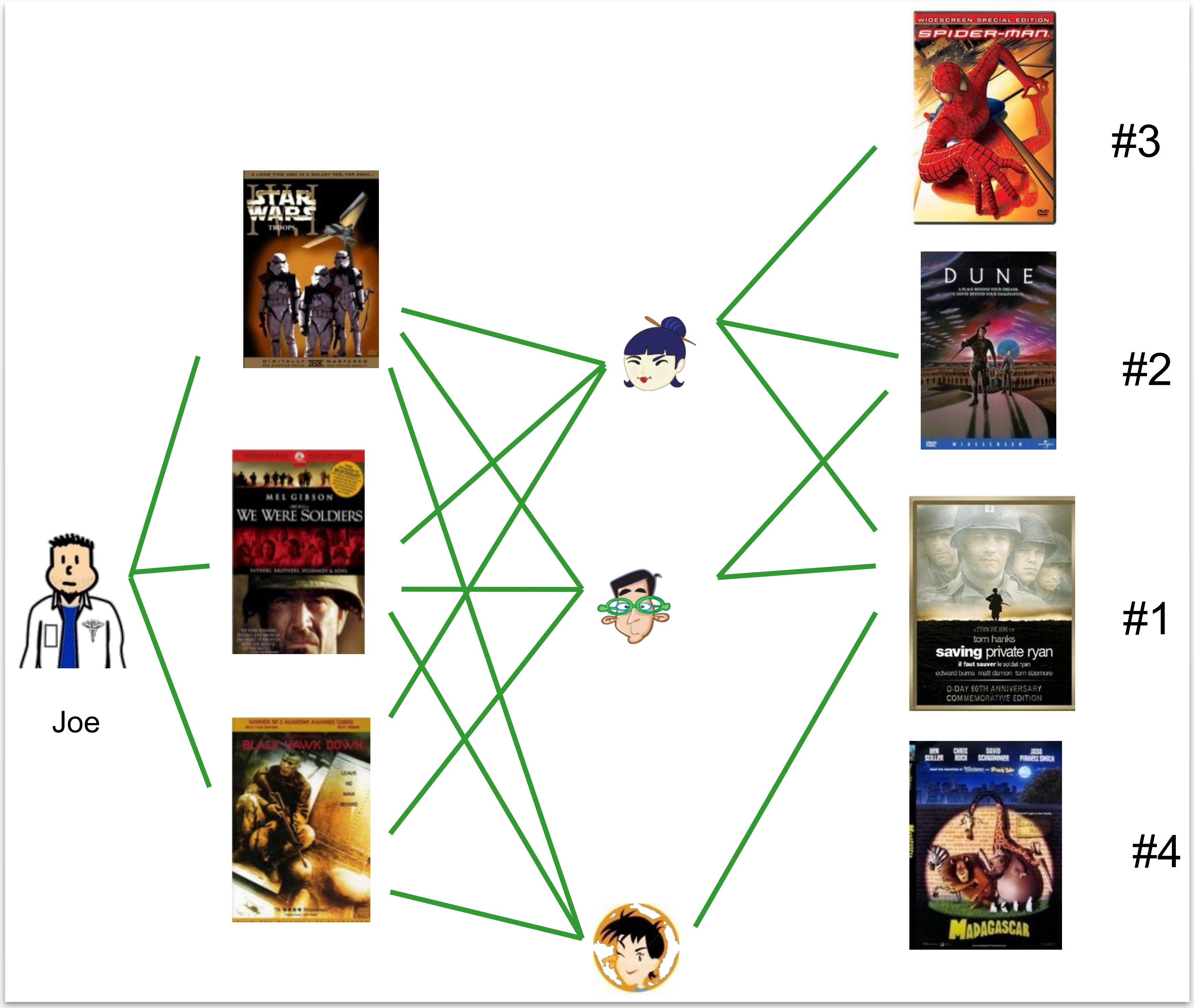
Idea: Predict rating using ***item*** features on a ***per-user*** basis

Content-based Filtering



Idea: Predict rating using ***user*** features on a ***per-item*** basis

Collaborative Filtering



Idea: Predict rating based on similarity to other users

Problem Setting

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

- *Task*: Predict user preferences for unseen items
- *Content-based filtering*: Model user/item features
- *Collaborative filtering*: Implicit similarity of users items

Recommender Systems

- Movie recommendation (Netflix)
- Related product recommendation (Amazon)
- Web page ranking (Google)
- Social recommendation (Facebook)
- News content recommendation (Yahoo)
- Priority inbox & spam filtering (Google)
- Online dating (OK Cupid)
- Computational Advertising (Everyone)

Challenges

- *Scalability*
  - Millions of objects
  - 100s of millions of users
- *Cold start*
  - Changing user base
  - Changing inventory
- *Imbalanced dataset*
  - User activity / item reviews power law distributed
  - Ratings are not missing at random

Running Example: Netflix Data

Training data

user	movie	date	score
1	21	5/7/02	1
1	213	8/2/04	5
2	345	3/6/01	4
2	123	5/1/05	4
2	768	7/15/02	3
3	76	1/22/01	5
4	45	8/3/00	4
5	568	9/10/05	1
5	342	3/5/03	2
5	234	12/28/00	2
6	76	8/11/02	5
6	56	6/15/03	4

Test data

user	movie	date	score
1	62	1/6/05	?
1	96	9/13/04	?
2	7	8/18/05	?
2	3	11/22/05	?
3	47	6/13/02	?
3	15	8/12/01	?
4	41	9/1/00	?
4	28	8/27/05	?
5	93	4/4/05	?
5	74	7/16/03	?
6	69	2/14/04	?
6	83	10/3/03	?

- Released as part of $1M competition by Netflix in 2006
- Prize awarded to BellKor in 2009

Running Yardstick: RMSE

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

$$\mathrm{rmse}(S) = \sqrt{|S|^{-1} \sum_{(i,u)\in S} (\hat{r}_{ui} - r_{ui})^2}$$

Running Yardstick: RMSE

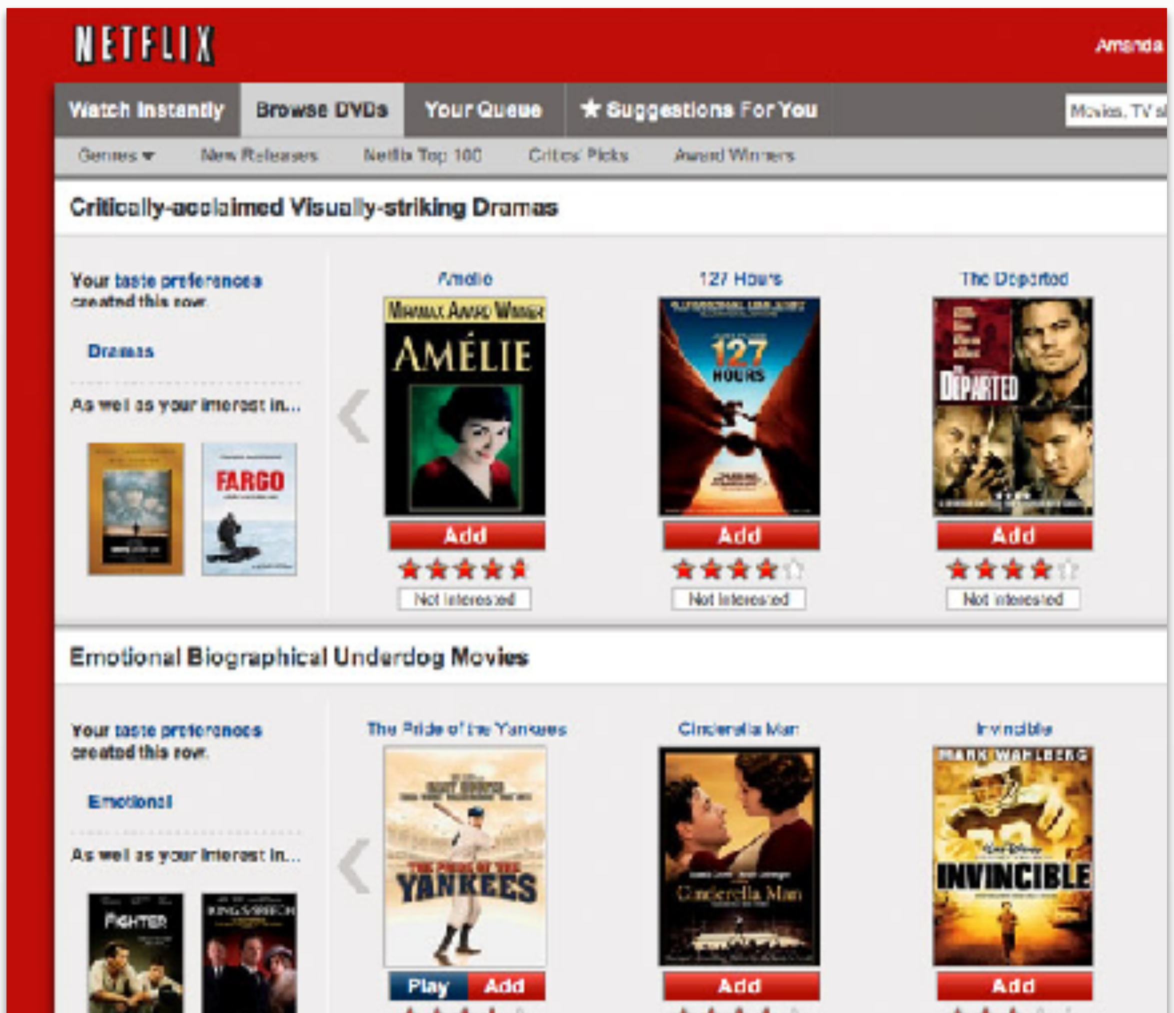
Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

$$\mathrm{rmse}(S) = \sqrt{|S|^{-1} \sum_{(i,u)\in S} (\hat{r}_{ui} - r_{ui})^2}$$

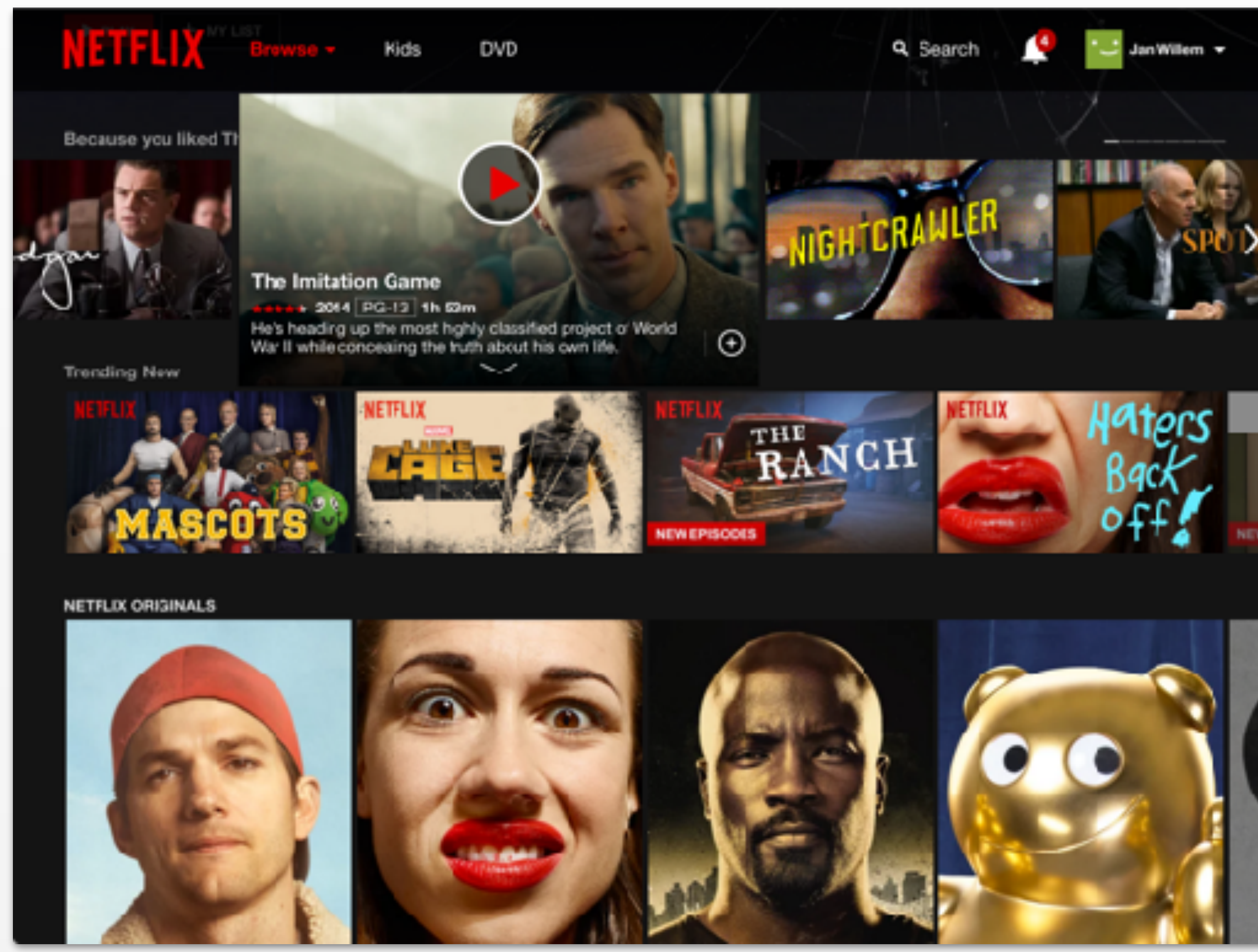
(doesn't tell you how to actually do recommendation)

Ratings aren't everything

Netflix then



Netflix now



Content-based Filtering

Item-based Features

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)
Love at last	5	5	0	0
Romance forever	5	?	?	0
Cute puppies of love	?	4	0	?
Nonstop car chases	0	0	5	4
Swords vs. karate	0	0	5	?

Item-based Features

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9

Item-based Features

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9

Per-user Regression

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9

Learn a set of regression coefficients for each user

$$\boldsymbol{w}_u = \underset{\boldsymbol{w}}{\operatorname{argmin}} |\boldsymbol{r}_u - \boldsymbol{X}\boldsymbol{w}|^2$$

Bias

Bias

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9

Bias

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9
Moonrise Kingdom	4	5	4	4	0.3	0.2

Bias

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9
Moonrise Kingdom	4	5	4	4	0.3	0.2

Problem: Some movies are universally loved / hated

Bias

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	3	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	3	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9
Moonrise Kingdom	4	3	4	4	0.3	0.2

Problem: Some movies are universally loved / hated
some users are more picky than others

Bias

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9
Moonrise Kingdom	4	5	4	4	0.3	0.2

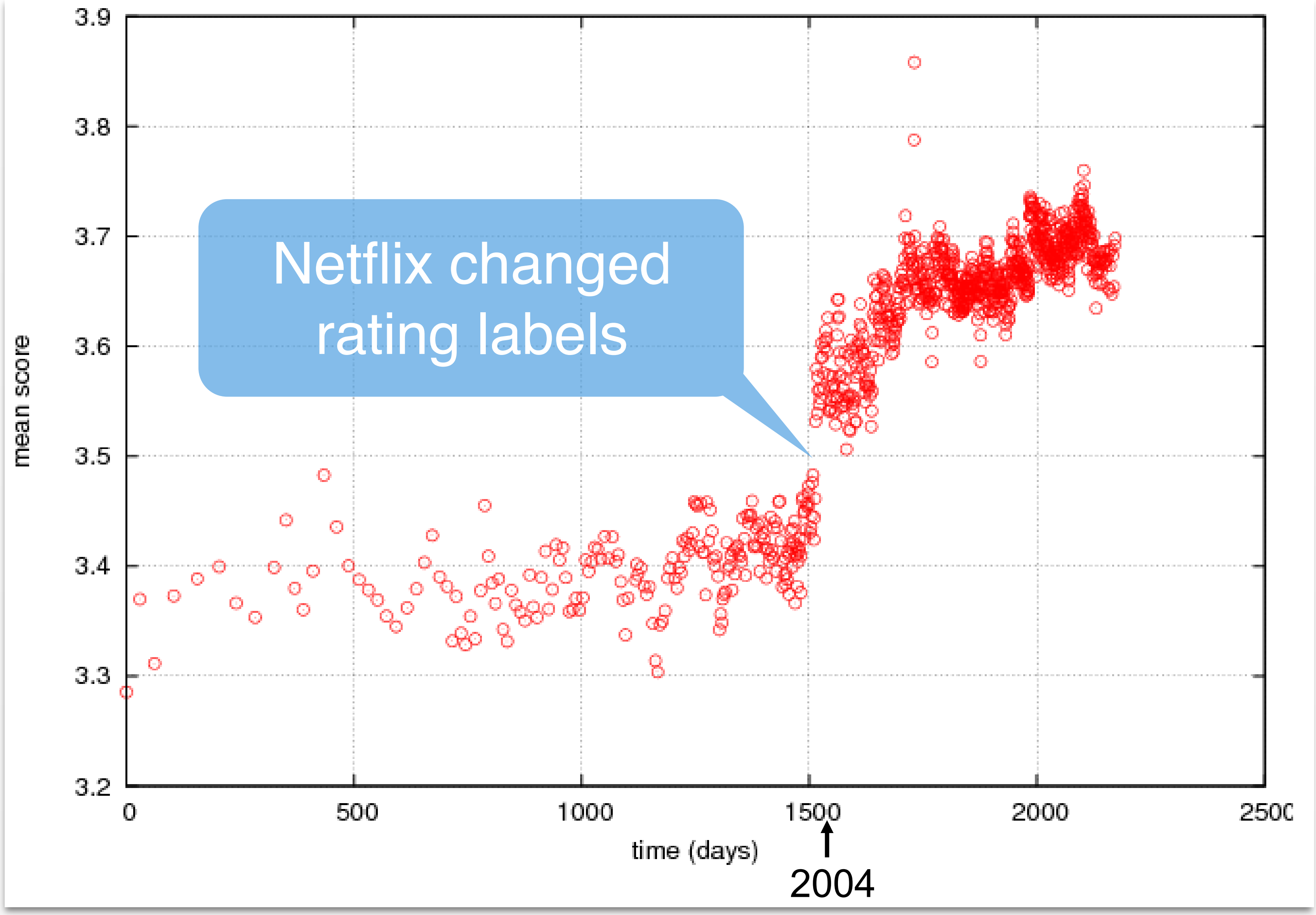
Problem: Some movies are universally loved / hated some users are more picky than others

Solution: Introduce a per-movie and per-user bias

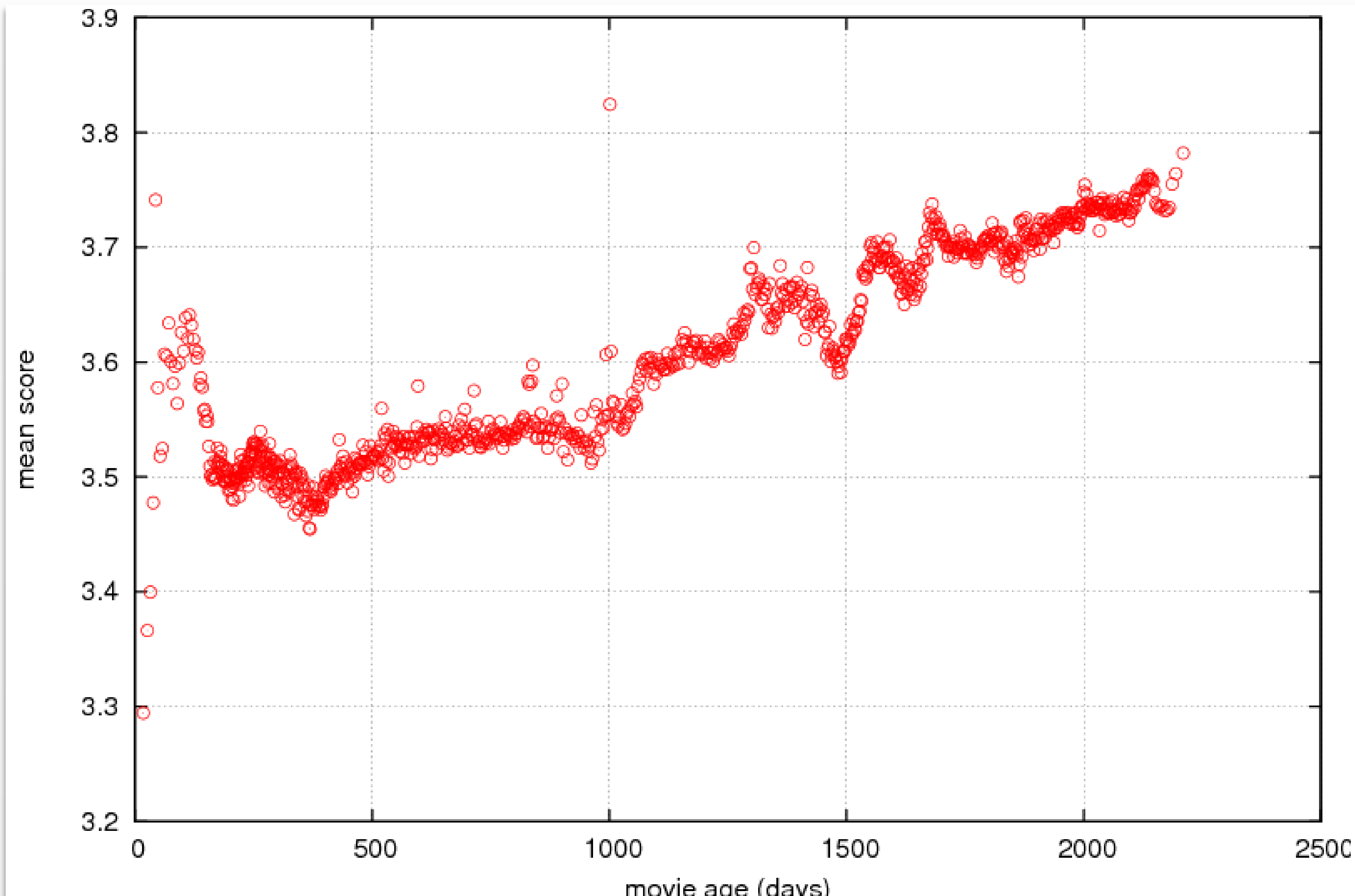
$$\hat{r}_{ui} = \mu + b_u + b_i + \boldsymbol{x}_i^\top \boldsymbol{w}_u$$

Temporal Effects

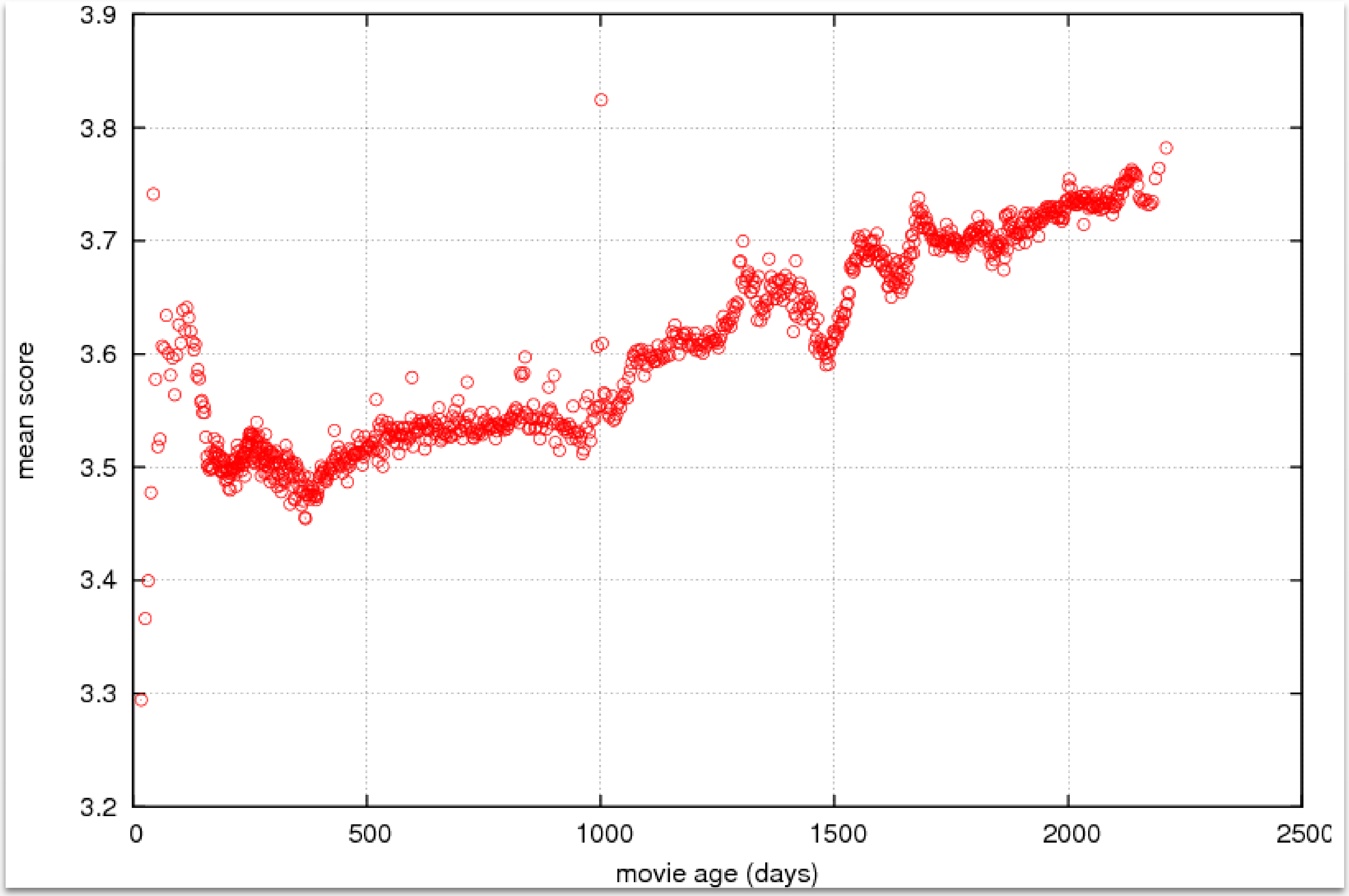
Changes in user behavior



Movies get better with time?



Temporal Effects

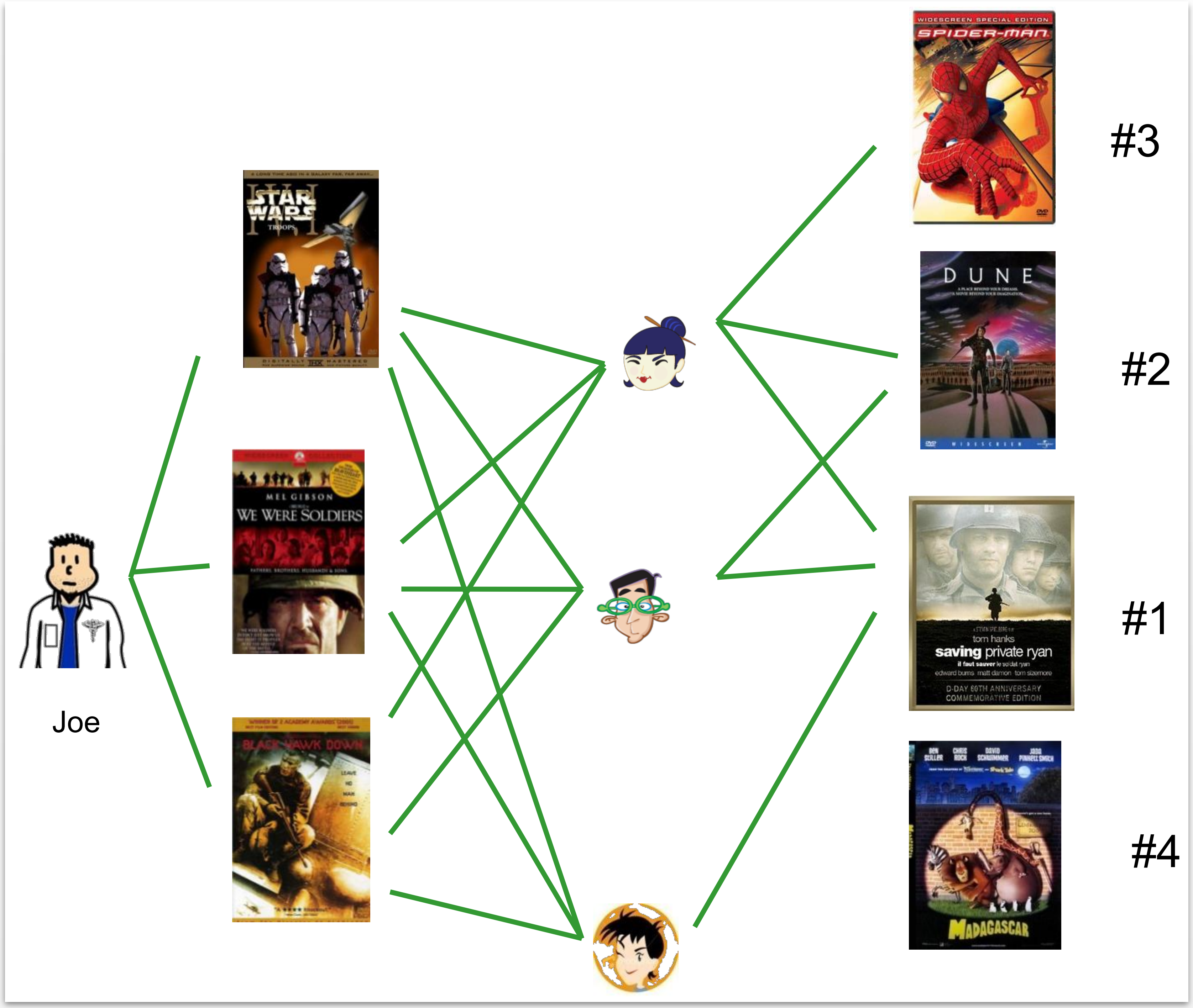


Solution: Model temporal effects in bias not weights

$$\hat{r}_{ui} = \mu(t) + b_u(t) + b_i(t) + \boldsymbol{x}_i^\top \boldsymbol{w}_u$$

Neighborhood Methods

Neighborhood Based Methods



Users and items form a bipartite graph (edges are ratings)

Neighborhood Based Methods

(user, user) similarity

- predict rating based on average from k-nearest users
- good if item base is smaller than user base
- good if item base changes rapidly

(item,item) similarity

- predict rating based on average from k-nearest items
- good if the user base is small
- good if user base changes rapidly

Parzen-Window Style CF

$$\hat{r}_{ui} = b_{ui} + \frac{\sum_{j \in s_k(i,u)} s_{ij}(r_{uj} - b_{uj})}{\sum_{j \in s_k(i,u)} s_{ij}} \qquad b_{ui} = \mu + b_u + b_i$$

- Define a similarity s_{ij} between items
- Find set $s_k(i,u)$ of k-nearest neighbors to i that were rated by user u
- Predict rating using weighted average over set
- How should we define s_{ij}?

Pearson Correlation Coefficient

User ratings for item i:

1	?	?	5	5	3	?	?	?	4	2	?	?	?	?	4	?	5	4	1	?

User ratings for item j:

?	?	4	2	5	?	?	1	2	5	?	?	2	?	?	3	?	?	?	5	4

$$s_{ij} = \frac{\mathrm{Cov}[r_{ui}, r_{uj}]}{\mathrm{Std}[r_{ui}]\mathrm{Std}[r_{uj}]}$$

(item,item) similarity

Empirical estimate of Pearson correlation coefficient

$$\hat{\rho}_{ij} = \frac{\sum_{u \in U(i,j)} (r_{ui} - b_{ui})(r_{uj} - b_{uj})}{\sqrt{\sum_{u \in U(i,j)} (r_{ui} - b_{ui})^2 \sum_{u \in U(i,j)} (r_{uj} - b_{uj})^2}}$$

Regularize towards 0 for small support

$$s_{ij} = \frac{|U(i,j)| - 1}{|U(i,j)| - 1 + \lambda} \hat{\rho}_{ij}$$

Regularize towards baseline for small neighborhood

$$\hat{r}_{ui} = b_{ui} + \frac{\sum_{j \in s_k(i,u)} s_{ij} (r_{uj} - b_{uj})}{\lambda + \sum_{j \in s_k(i,u)} s_{ij}}$$

Similarity for binary labels

Pearson correlation not meaningful for binary labels
(e.g. Views, Purchases, Clicks)

Jaccard similarity

$$s_{ij} = \frac{m_{ij}}{\alpha + m_i + m_j - m_{ij}}$$

Observed / Expected ratio

$$s_{ij} = \frac{\text{observed}}{\text{expected}} \approx \frac{m_{ij}}{\alpha + m_i m_j / m}$$

m_i users acting on i

m_{ij} users acting on both i and j

m total number of users

Matrix Factorization Methods

Matrix Factorization

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9
Moonrise Kingdom	4	5	4	4	0.3	0.2

$$\hat{r}_{ui} = \mu + b_u + b_i + \boldsymbol{x}_i^\top \boldsymbol{w}_u$$

Matrix Factorization

Movie	Alice (1)	Bob (2)	Carol (3)	Dave (4)	x_1 (romance)	x_2 (action)
Love at last	5	5	0	0	0.9	0
Romance forever	5	?	?	0	1.0	0.01
Cute puppies of love	?	4	0	?	0.99	0
Nonstop car chases	0	0	5	4	0.1	1.0
Swords vs. karate	0	0	5	?	0	0.9
Moonrise Kingdom	4	5	4	4	0.3	0.2

$$\hat{r}_{ui} = \mu + b_u + b_i + \boldsymbol{x}_i^\top \boldsymbol{w}_u$$

Idea: pose as (biased) matrix factorization problem

$$\hat{\boldsymbol{R}} = \boldsymbol{B} + \boldsymbol{X}\boldsymbol{W}^\top$$

Matrix Factorization

users

items

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

~

items

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

●

users

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

A rank-3 SVD approximation

Prediction

users

items

1		3			5			5		4	
		5		?		4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

~

items

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

•

users

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

A rank-3 SVD approximation

Prediction

users

items

1		3			5			5		4	
		5		2.4		4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

~

items

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

●

users

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

A rank-3 SVD approximation

SVD with missing values

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

Pose as regression problem

$$\underset{X,W}{\operatorname{argmin}} \sum_{(u,i)\in S} (r_{ui} - \boldsymbol{w}_u^\top \boldsymbol{x}_i)^2 + \lambda \left(||\boldsymbol{X}||_{\mathrm{F}}^2 + ||\boldsymbol{W}||_{\mathrm{F}}^2 \right)$$

Regularize using Frobenius norm

$$||\boldsymbol{A}||_{\mathrm{F}}^2 = \sum_{ij} |A_{ij}|^2$$

Alternating Least Squares

$\boldsymbol{R}$

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

$\boldsymbol{X}$

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

$\boldsymbol{W}^\top$

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

$$\boldsymbol{w}_u \leftarrow \left[\lambda \boldsymbol{I} + \sum_{i:(u,i)\in S} \boldsymbol{x}_i \boldsymbol{x}_i^\top\right]^{-1} \sum_{i:(u,i)\in S} \boldsymbol{x}_i r_{ui}$$

(regress w_u given X)

Alternating Least Squares

$\boldsymbol{R}$ ~ $\boldsymbol{X}$ $\boldsymbol{W}^\top$

$\boldsymbol{R}$

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

$\boldsymbol{X}$

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

$\boldsymbol{W}^\top$

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

$$\boldsymbol{w}_u \leftarrow \left[\lambda \boldsymbol{I} + \sum_{i:(u,i)\in S} \boldsymbol{x}_i \boldsymbol{x}_i^\top\right]^{-1} \sum_{i:(u,i)\in S} \boldsymbol{x}_i r_{ui}$$

(regress w_u given X)

$L2$: closed form solution

$$\mathbf{w} = (\mathbf{X}^\mathsf{T}\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^\mathsf{T}\mathbf{y}$$

Remember ridge regression?

Alternating Least Squares

$R \sim X \; W^\top$

R

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

X

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

$W^\top$

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

$$\boldsymbol{w}_u \leftarrow \left[\lambda \boldsymbol{I} + \sum_{i:(u,i)\in S} \boldsymbol{x}_i \boldsymbol{x}_i^\top \right]^{-1} \sum_{i:(u,i)\in S} \boldsymbol{x}_i r_{ui}$$

(regress w_u given X)

$$\boldsymbol{x}_i \leftarrow \left[\lambda \boldsymbol{I} + \sum_{u:(u,i)\in S} \boldsymbol{w}_u \boldsymbol{w}_u^\top \right]^{-1} \sum_{u:(u,i)\in S} \boldsymbol{w}_u r_{ui}$$

(regress x_i given W)

Stochastic Gradient Descent

$\boldsymbol{R}$

1		3			5			5		4	
		5	4			4			2	1	3
2	4		1	2		3		4	3	5	
	2	4		5			4			2	
		4	3	4	2					2	5
1		3		3			2			4	

~

$\boldsymbol{X}$

.1	-.4	.2
-.5	.6	.5
-.2	.3	.5
1.1	2.1	.3
-.7	2.1	-2
-1	.7	.3

$\boldsymbol{W}^\top$

1.1	-.2	.3	.5	-2	-.5	.8	-.4	.3	1.4	2.4	-.9
-.8	.7	.5	1.4	.3	-1	1.4	2.9	-.7	1.2	-.1	1.3
2.1	-.4	.6	1.7	2.4	.9	-.3	.4	.8	.7	-.6	.1

$$\boldsymbol{w}_u \leftarrow (1 - \lambda\eta_t)\boldsymbol{w}_u + \eta_t \boldsymbol{x}_i (r_{ui} - \boldsymbol{w}_u^\top \boldsymbol{x}_i)$$

$$\boldsymbol{x}_i \leftarrow (1 - \lambda\eta_t)\boldsymbol{x}_i + \eta_t \boldsymbol{w}_u (r_{ui} - \boldsymbol{w}_u^\top \boldsymbol{x}_i)$$

- No need for locking
- Multicore updates asynchronously (Recht, Re, Wright, 2012 - Hogwild)

Netflix Prize

Netflix Prize

Training data

- 100 million ratings, 480,000 users, 17,770 movies
- 6 years of data: 2000-2005

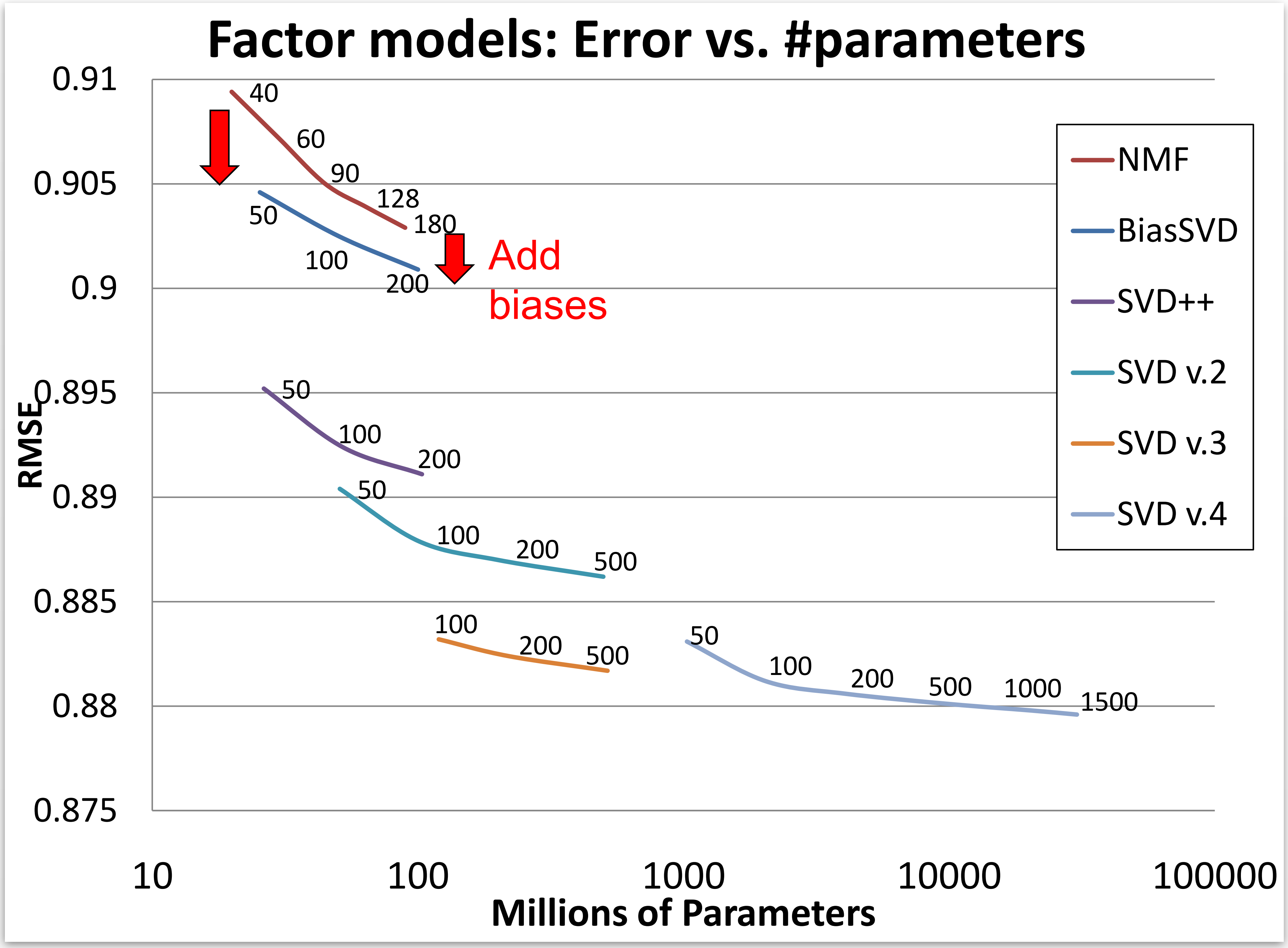
Test data

- Last few ratings of each user (2.8 million)
- Evaluation criterion: Root Mean Square Error (RMSE)

Competition

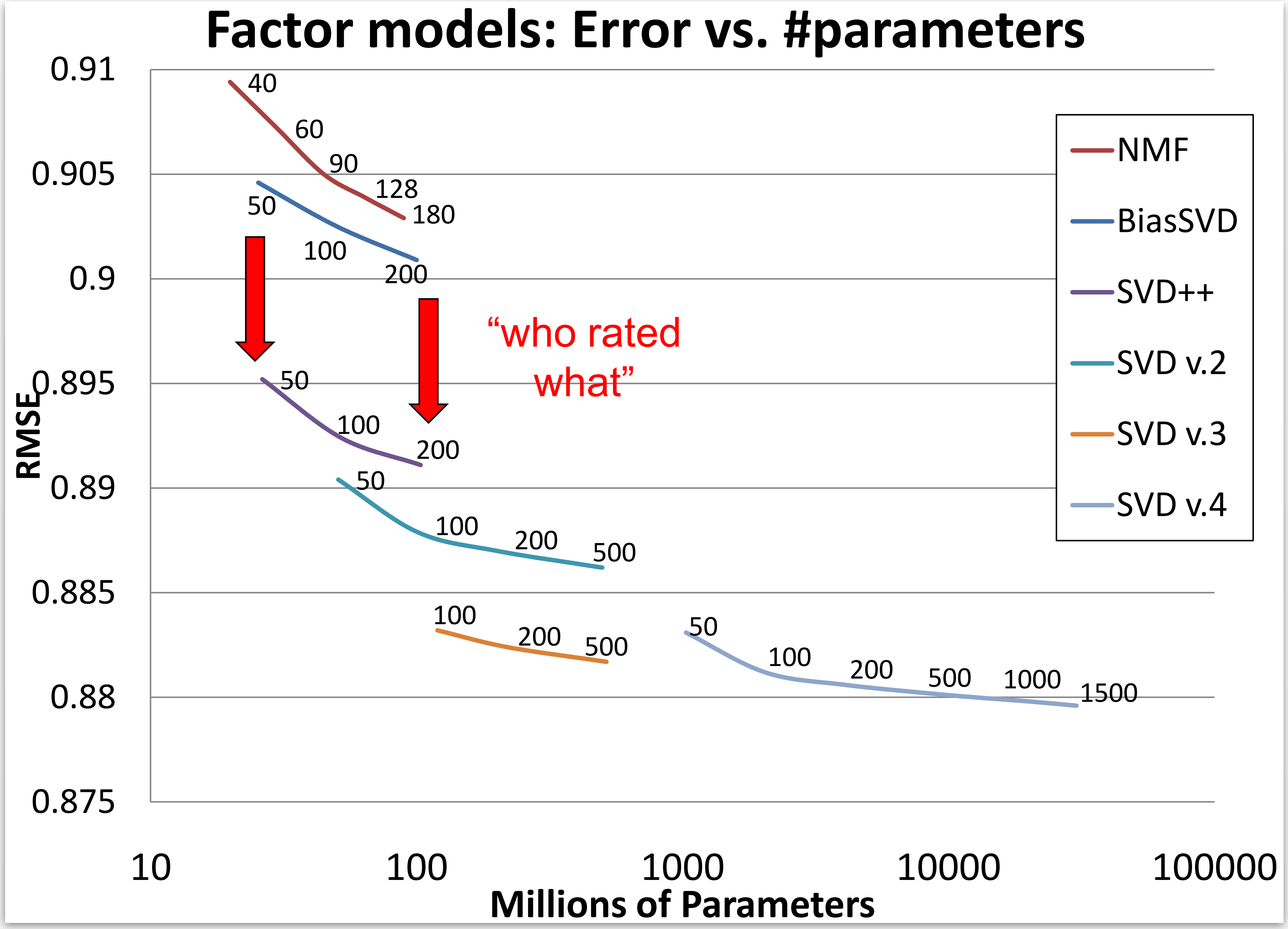
- 2,700+ teams
- Netflix's system RMSE: 0.9514
- $1 million prize for 10% improvement on Netflix

Improvements



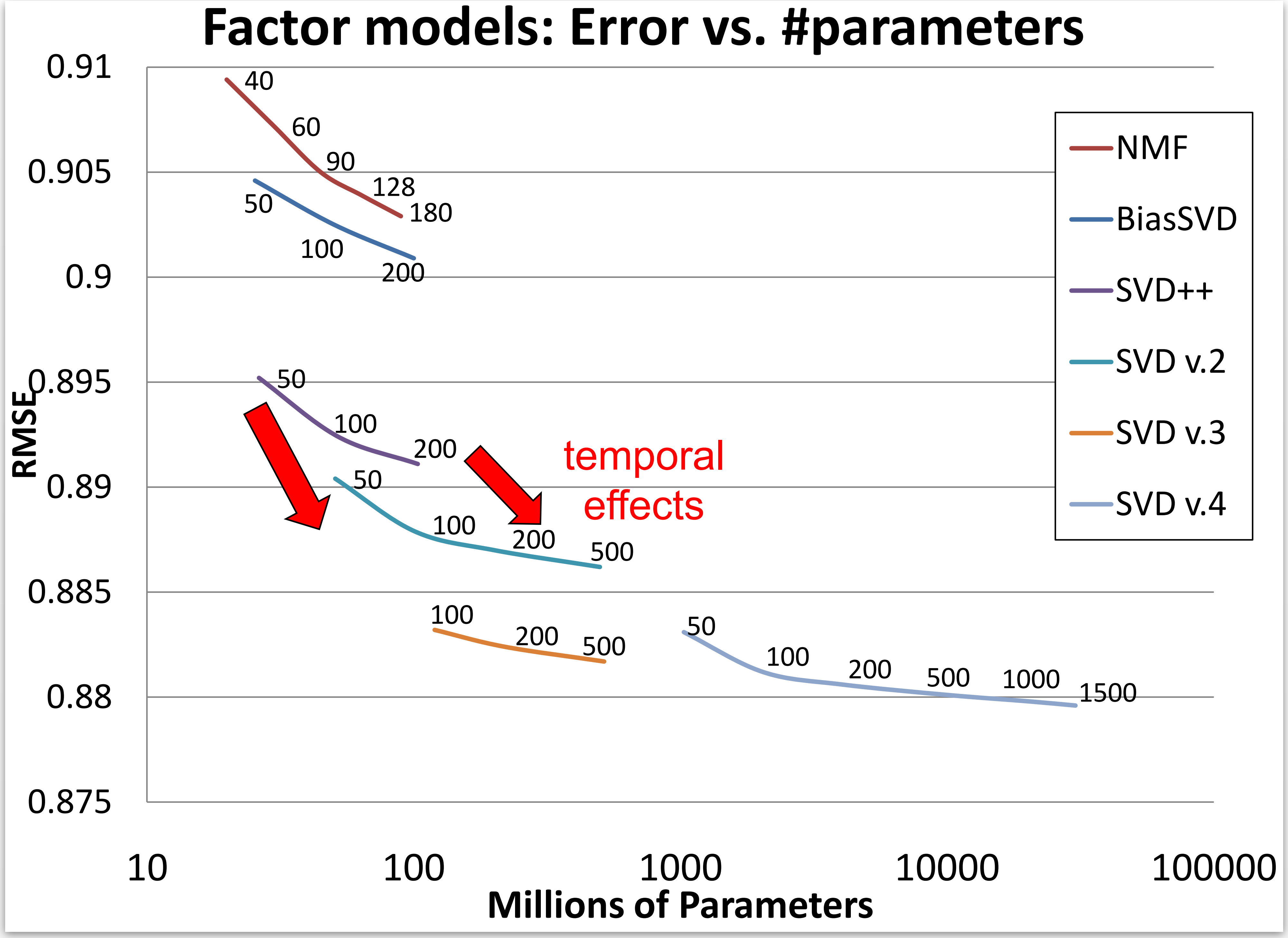
Do SGD, but also learn biases μ, b_u and b_i

Improvements



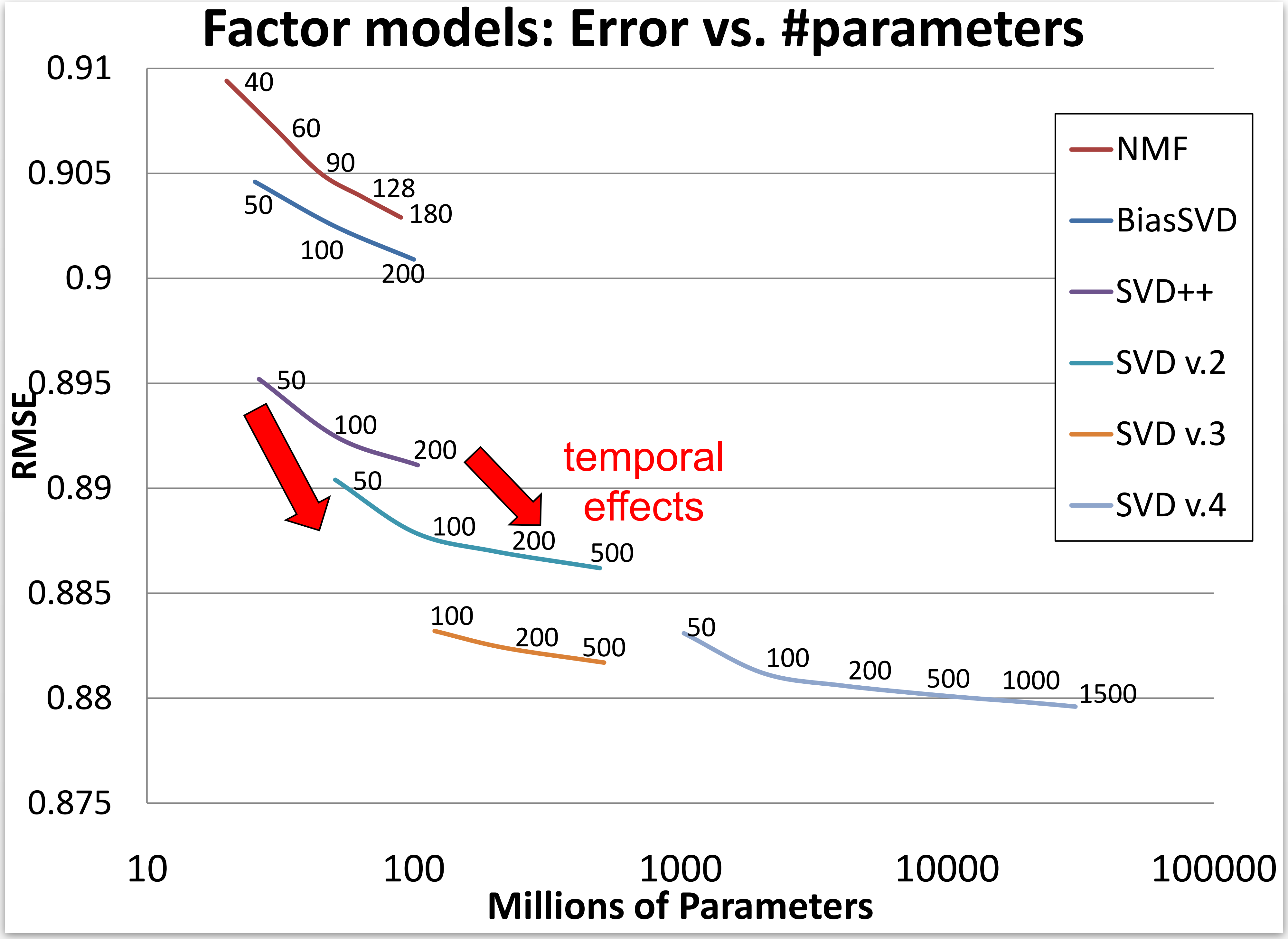
Account for fact that ratings are not missing at random.

Improvements



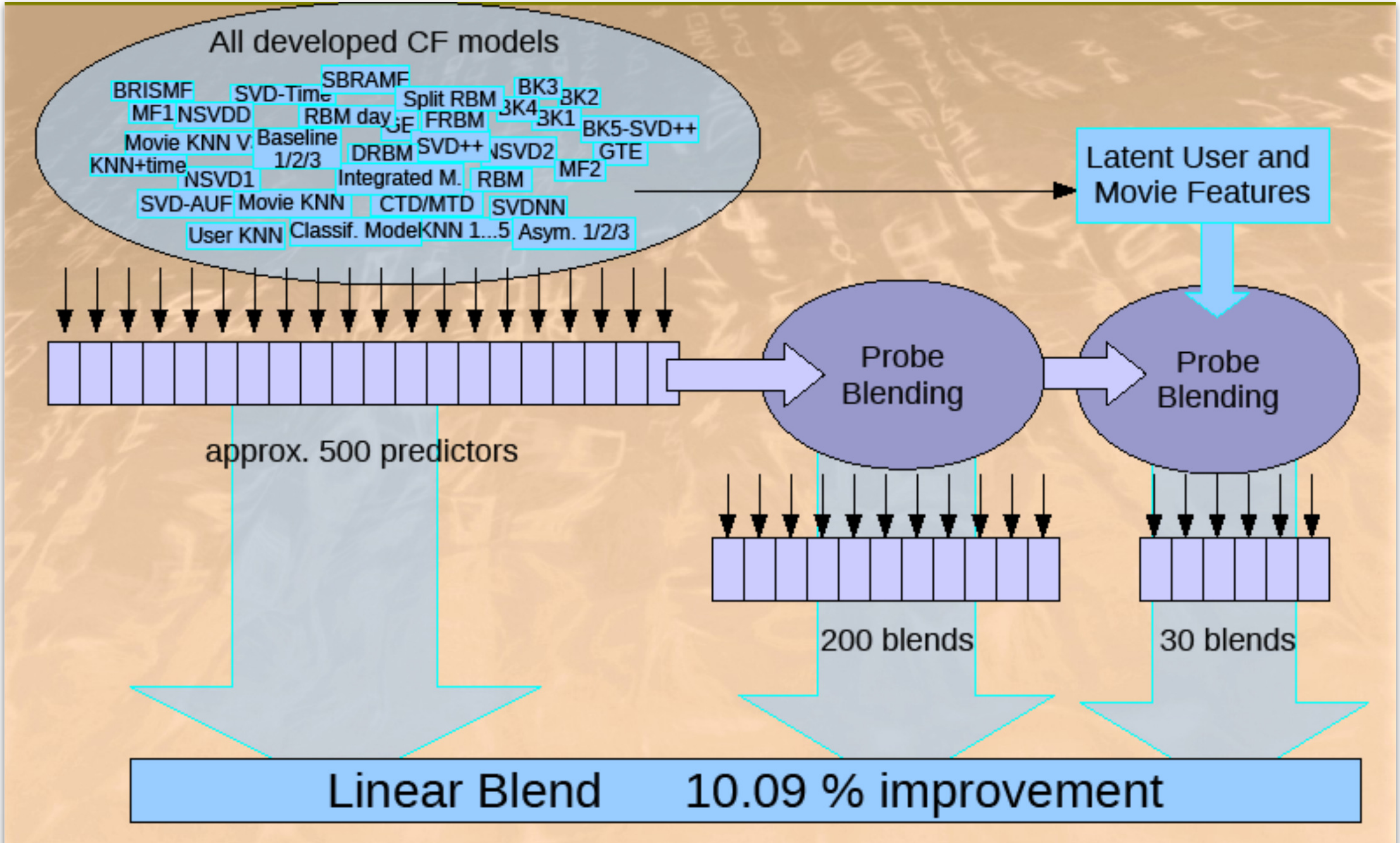
Account for drift in user and item biases

Improvements



Still pretty far from 0.8563 grand prize

Winning Solution from BellKor



Last 30 days

NETFLIX

Netflix Prize

Home | Rules | Leaderboard | Register | Update | Submit | Download

Leaderboard

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
Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos				
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	xlvector	0.8639	9.20	2009-06-26 13:49:04
13	xiangliang	0.8639	9.20	2009-06-26 07:47:34

June 26th submission triggers 30-day “last call”

Last 30 days

NETFLIX

Netflix Prize

Home | Rules | Leaderboard | Register | Update | Submit | Download

Leaderboard

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
Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor in BigChaos				
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	xlvector	0.8639	9.20	2009-06-26 13:49:04
13	xiangliang	0.8639	9.20	2009-06-26 07:47:34

June 26th submission triggers 30-day "last call"

BellKor fends off competitors by a hair

Netflix Prize

COMPLETED

Home | Rules | Leaderboard | Update | Download

Leaderboard

Showing Test Score. Click here to show quiz score

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	[illegible]	[illegible]	[illegible]
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

BellKor fends off competitors by a hair