Building a Board Game Recommender System

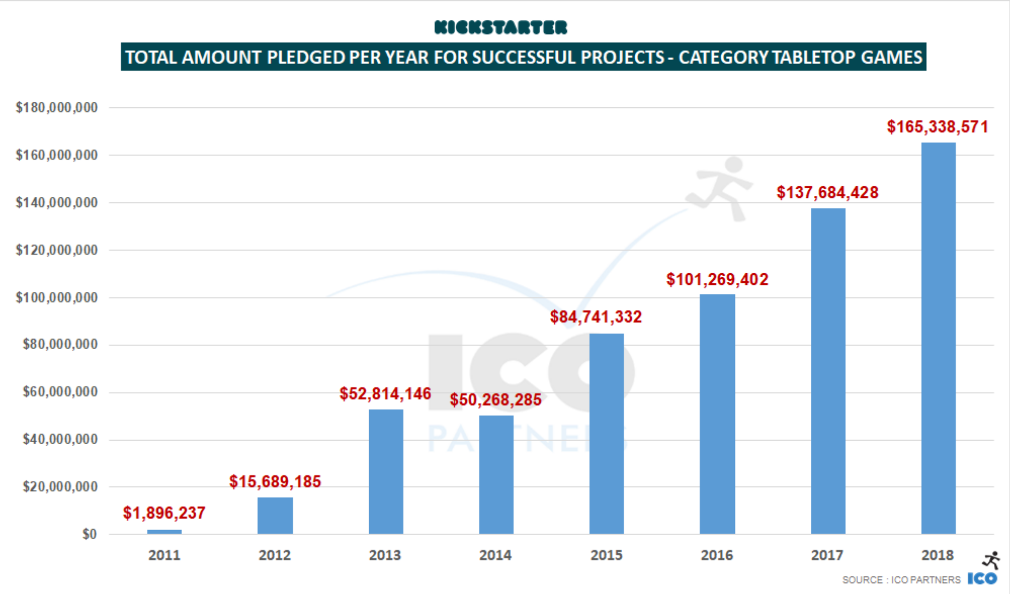
JPMorgan Chase Tech Talk

Aaron Huang

April 17, 2019

Tabletop games a rapidly growing market

- Tabletop games accounted for over ¼ of all funds ($608M) raised on Kickstarter in 2018
- Retail sales of tabletop games exceeded $1.5B in 2017 in North America, $7.2B worldwide
- Value of board game market forecasted to reach $12B worldwide by 2023



Source : icopartners.com/2019/01/games-and-crowdfunding-in-2018/

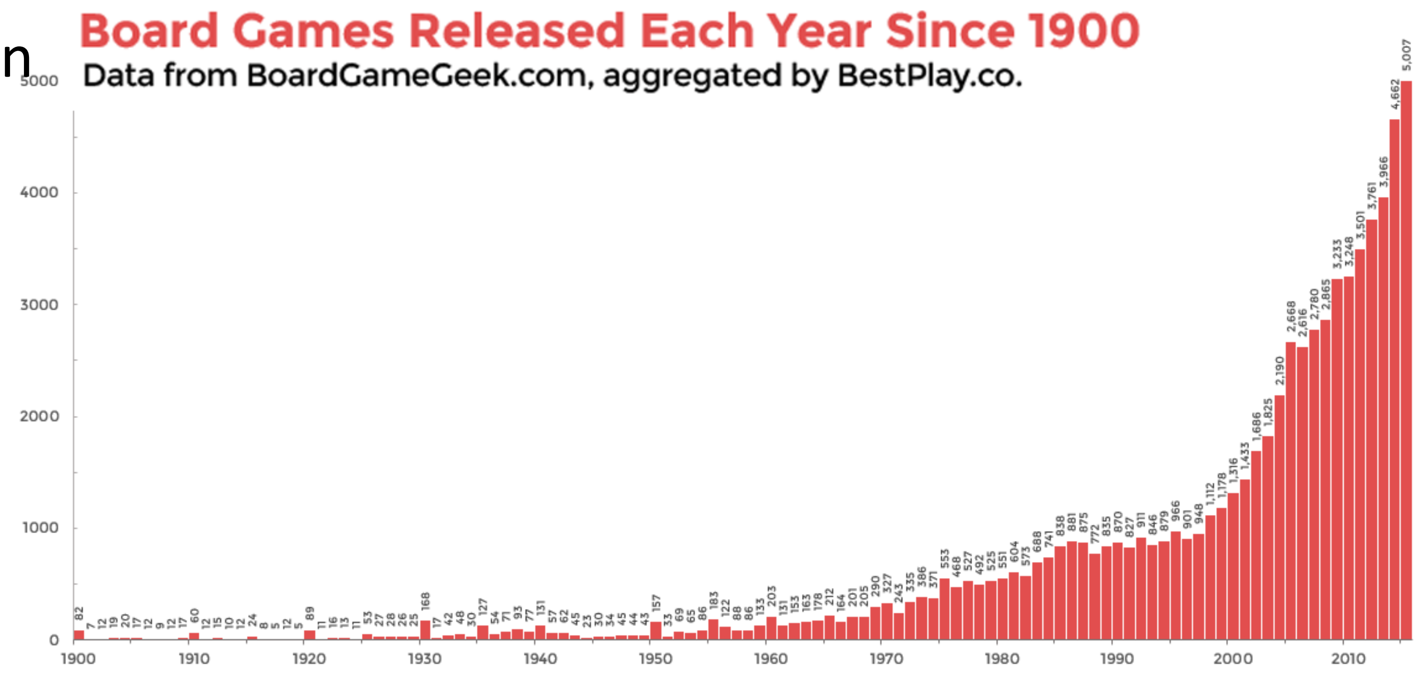
U.S./CANADA GAMES SALES – 2017

HOBBY GAMES CATEGORY	RETAIL SALES (IN MILLIONS)
COLLECTIBLE GAMES	$725
NON-COLLECTIBLE MINIATURE GAMES	$270
HOBBY BOARD GAMES	$345
HOBBY CARD AND DICE GAMES	$150
ROLEPLAYING GAMES	$55
TOTAL HOBBY GAMES	$1,545

Source : https://icv2.com/articles/news/view/41016/hobby-games-top-1-5-billion

Large product selection confusing for consumers

- Over 5,000 games released in 2016 alone
- Over 100,000 results on Amazon
- Specialty shops typically carry 5,000-10,000 products



Source : https://www.bestplay.co/

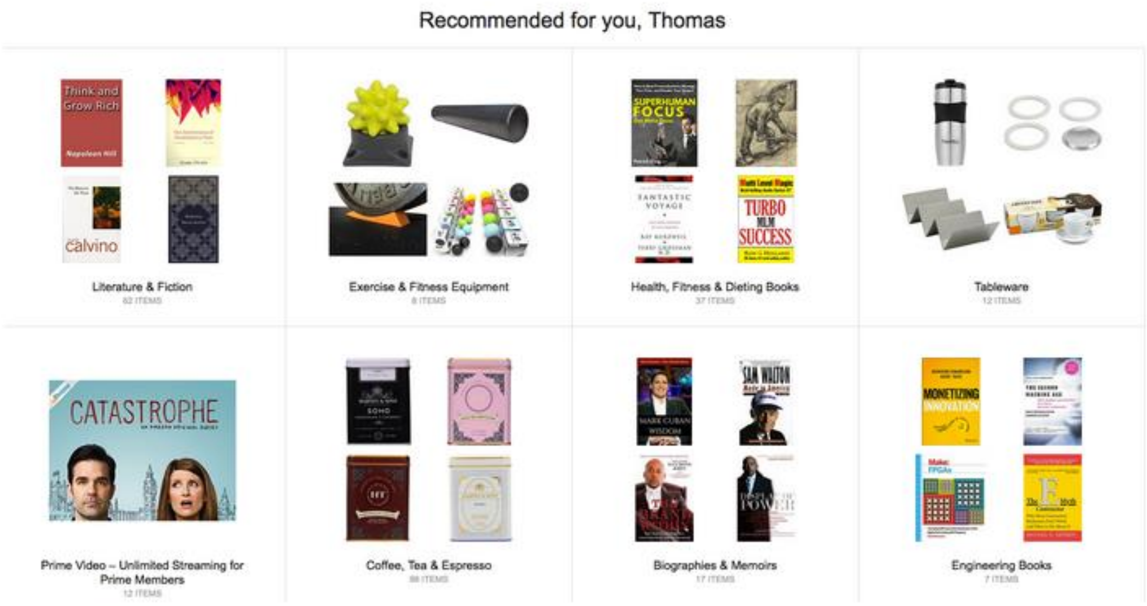

Bombay
COLONY
NINJA
Marrakech
SKYLINE 3000
BED, WED, DEAD
cubo
Death Inc.


Co-op
BATTLES OF WESTEROS

The need for recommender systems

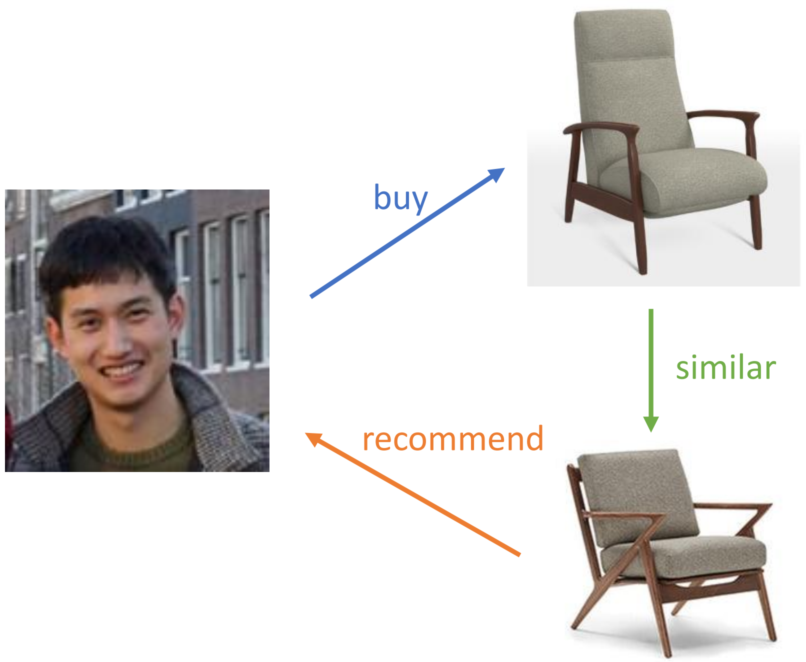
15,000 titles on Netflix

600 million products on Amazon

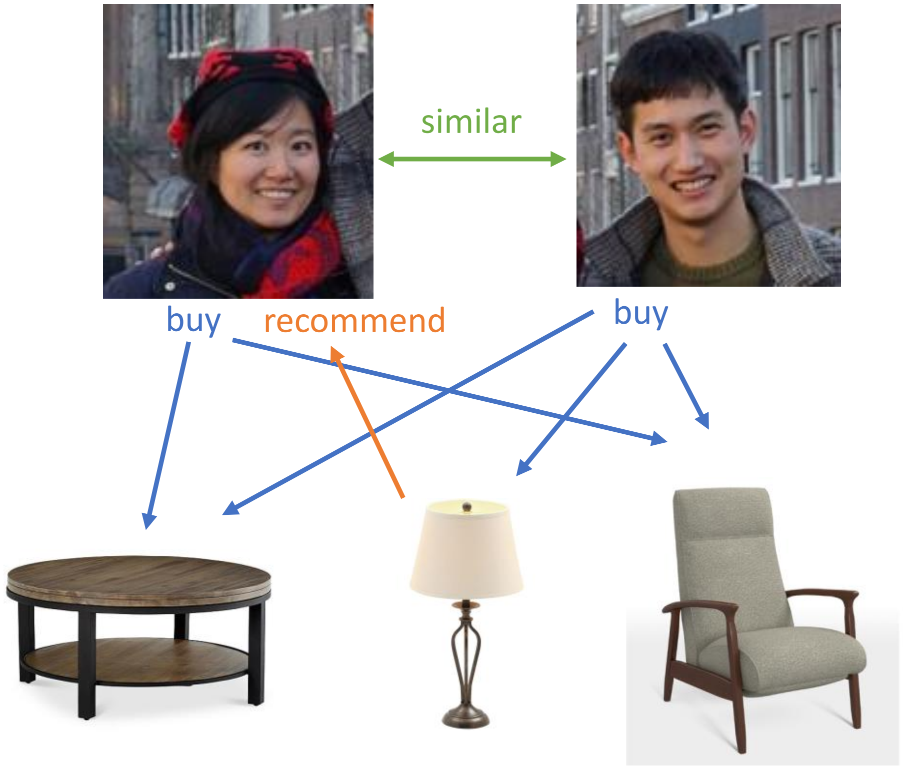


Two common types of recommenders

Content-Based



Collaborative Filtering



Gathering data

- BoardGameGeek.com
  - 2 million registered users, 4 million unique visitors per month
  - 100,000 board games in database
  - Public API

Output

Input
Item “id”


```
−<comments page="1" totalitems="16210">
<comment username="-Johnny-" rating="7" value="It has a good, quick pace, and works wel
<comment username="-xXx-" rating="5" value="Too much luck. A bit too simple to spark in
<comment username="00daniel00" rating="8" value="I don't know why I haven't buy this ga
<comment username="0Felippe0" rating="N/A" value="na stanie"/>
<comment username="1 Family Meeple" rating="6" value="OWN EXPANSION(S): EXPAN
<comment username="1000days" rating="6" value="I was too young and/or too ignorant to g
games I'd learned to enjoy by then."/>
<comment username="1000rpm" rating="7" value="Bizarrely, my copy of this game, which
personal frustration. Okay, its got a large random factor with the dice and a popularity/shouting
<comment username="100pcBlade" rating="8" value="I would originally have rated this as a
perfect gateway game and will always select it when introducing newbies. I always preferred th
<comment username="1024b" rating="8.25" value="Отсутствует одна карточка МОНОПС
<comment username="12thManStanding" rating="6" value="Haven't played this one in years
<comment username="143245" rating="3" value="I don't have any visceral hate for the game
expansion), and there is a permanence in the area/point controls. Catan just takes less time is the
<comment username="144creations" rating="3" value="Perhaps I just over-played this one in
stuck playing). "/>
<comment username="14cross" rating="9" value="I like the game but the only problem I hav
13 year old daughter and son and we all kept rolling 9s all of the time, which gave my daughter
few more times and, while I still don't see a ton of strategy in this game, I am seeing more possi
9 for me. We're having a lot of fun with it!"/>
<comment username="16note" rating="7" value="Rachel's half"/>
<comment username="1amgreg77" rating="8.7" value="One of my favorite 4 player games.
recommend the cities and knights expansion. "/>
<comment username="1Aspielerin" rating="N/A" value="wooden pieces"/>
<comment username="1awesomeguy" rating="5" value="Takes WAY too long for what it is.
<comment username="1hooman" rating="N/A" value="Box says original title "The Settlers c
```


Register | Browse ▾ | Forums ▾ | GeekLists ▾ | Bazaar ▾ | Misc ▾ | Help

Board Game Rank ▲	Title	Geek Rating	Avg Rating	Num Voters
1	Gloomhaven (2017)	8.612	8.91	24206
2	Pandemic Legacy: Season 1 (2015)	8.491	8.65	30730
3	Through the Ages: A New Story of Civilization (2015)	8.258	8.54	15404
4	Terraforming Mars (2016)	8.241	8.40	37736
5	Twilight Struggle (2005)	8.175	8.33	34112
6	Star Wars: Rebellion (2016)	8.162	8.45	16327

.../boardgame/187645/star-wars-rebellion

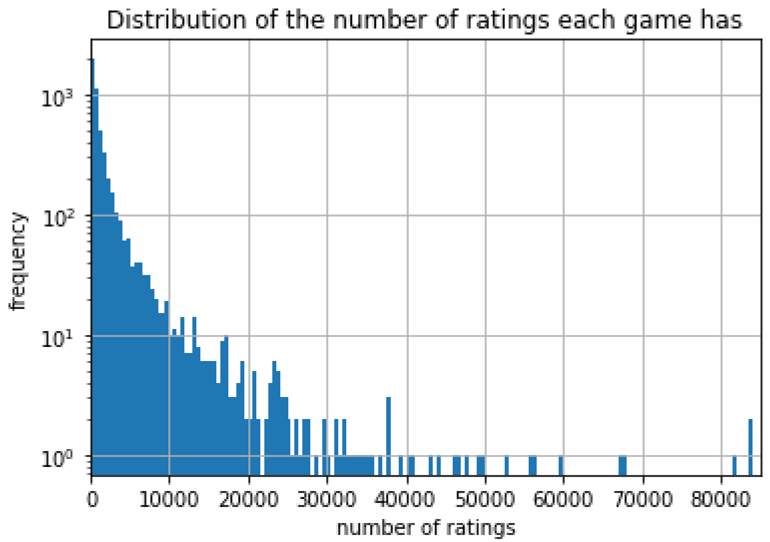
Item “id”

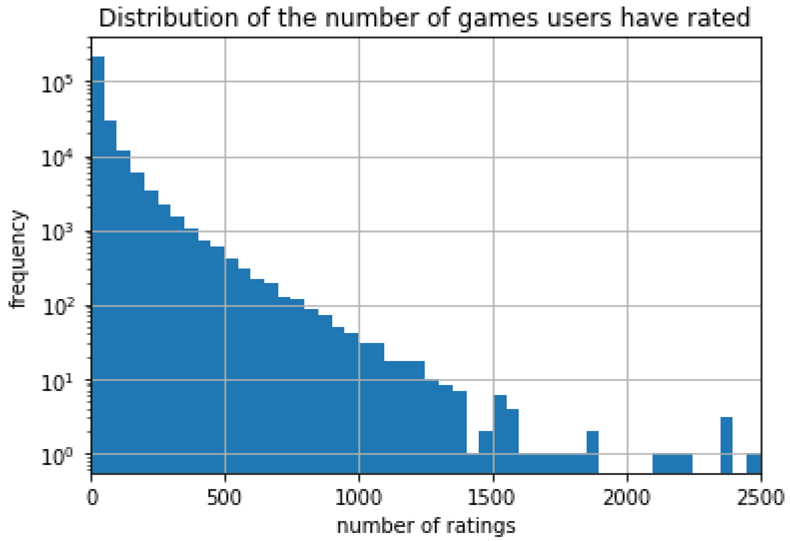
Gathering data

- Full data set (bgg-113)
  - Minimum user ratings = 1
  - Minimum item ratings = 100
  - 11.3 million ratings (1-10 scale)
  - 272,553 users
  - 5,077 items

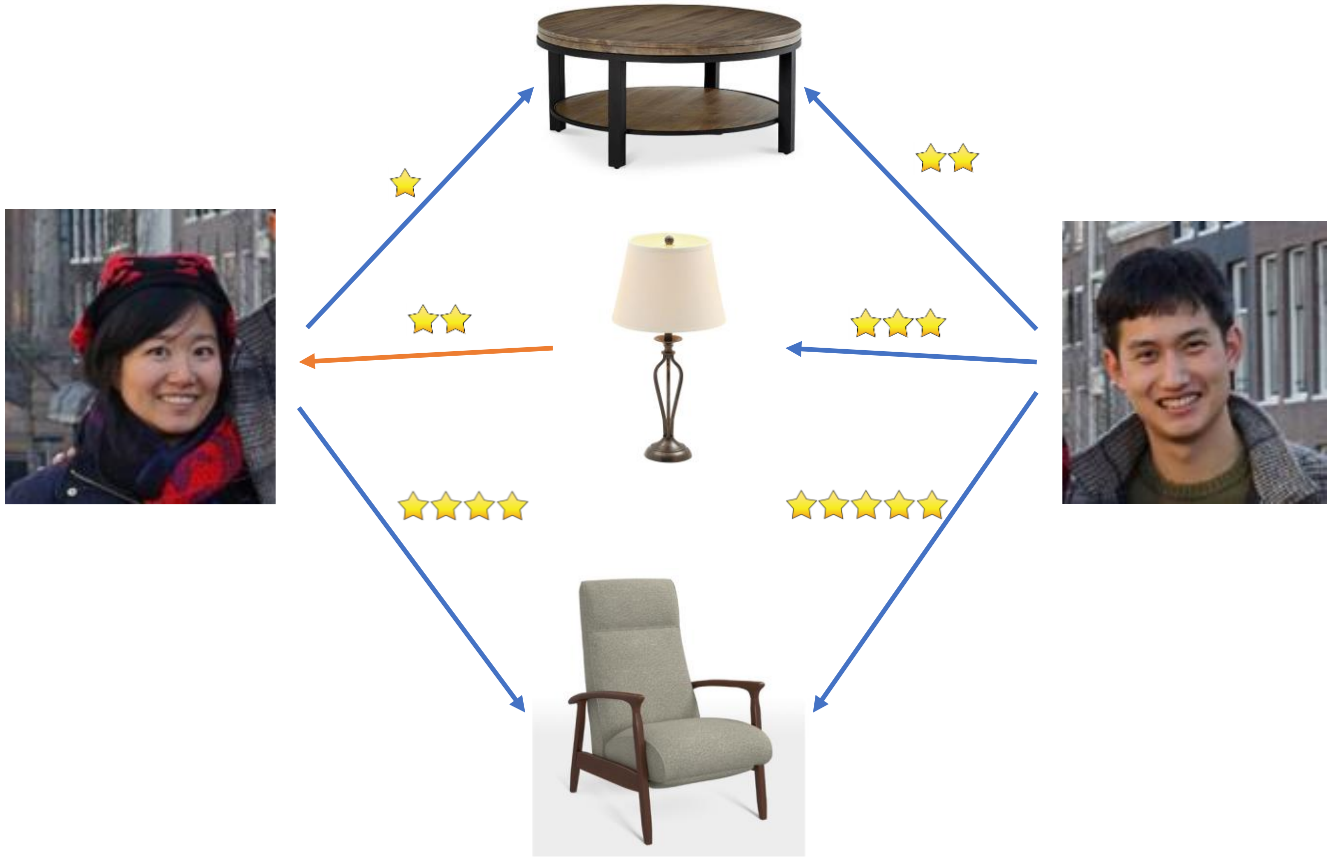
- “Medium” data set (bgg-86)
  - Minimum user ratings = 50
  - Minimum item ratings = 100
  - 8.6 million ratings
  - 58,919 users
  - 4,905 items

- “Small” data set (bgg-26)
  - Minimum user ratings = 300
  - Minimum item ratings = 100
  - 2.6 million ratings
  - 5,387 users
  - 3,827 items

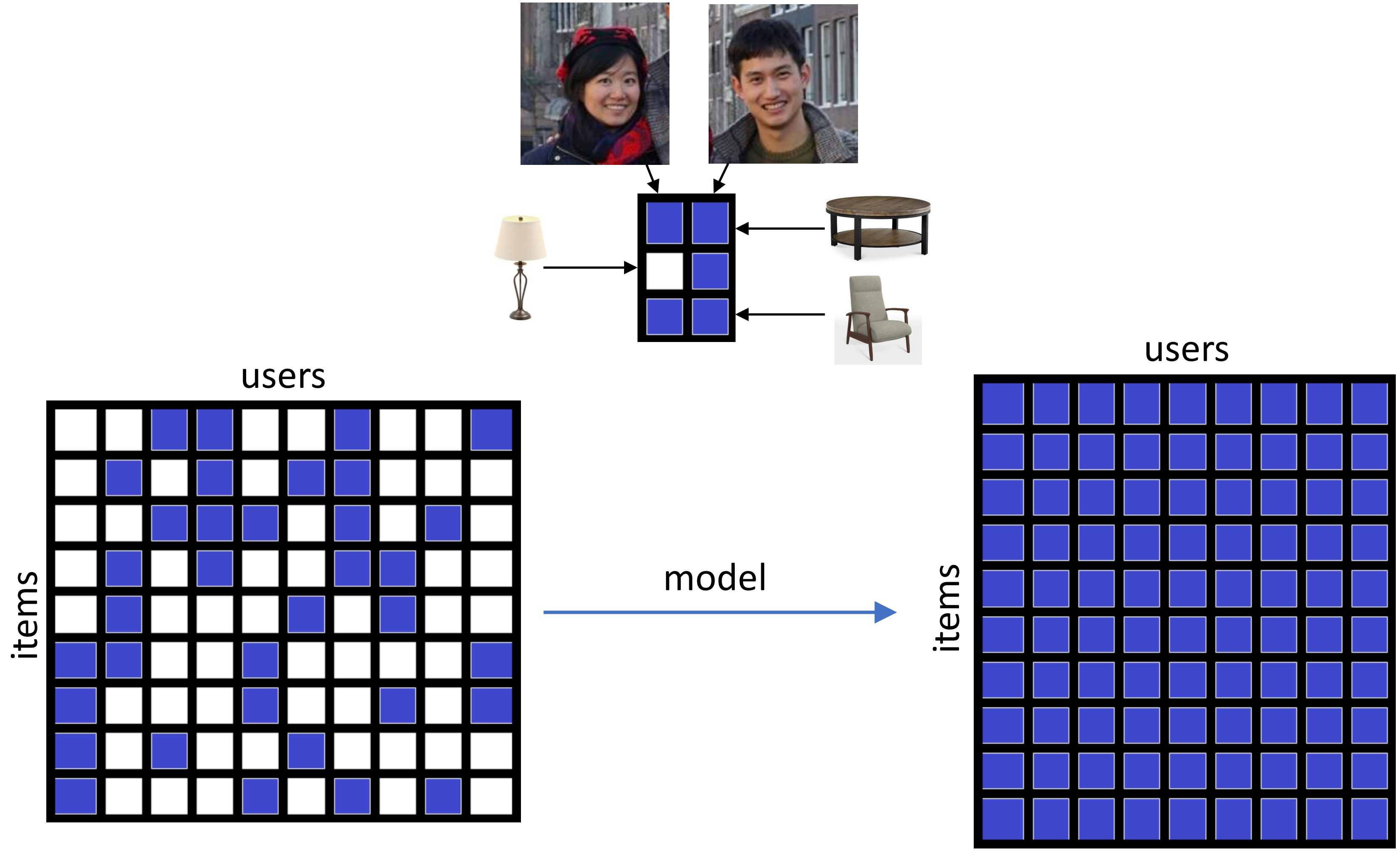




Collaborative filtering



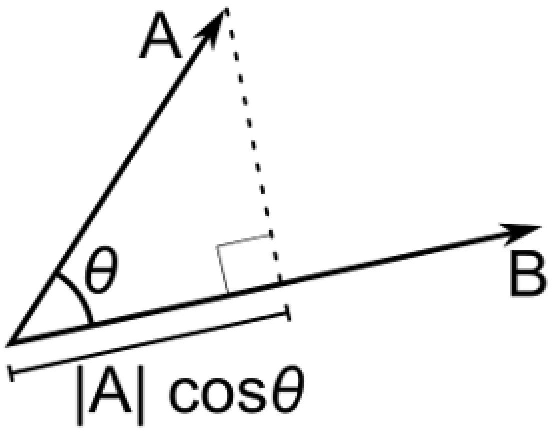
Matrix completion problem



Neighborhood approach

$$s_{u,u'} = \sum_i \frac{r_{ui} r_{u'i}}{(\sum_i r_{ui}^2)^{1/2} (\sum_i r_{u'i}^2)^{1/2}}$$

User-user likeness calculated as cosine similarity



$$\hat{r}_{ui} = \frac{\sum_{u'} s_{u,u'} r_{u',i}}{\sum_{u'} s_{u,u'}}$$

Ratings calculated as the average rating given by all users weighted by similarity score

- Validation and testing
  - 70% training, 15% validation, 15% testing
- Error
  - "Masked" RMSE

$$mRMSE = \sqrt{mean(r_{ui} - \hat{r}_{ui})^2_{r_{ui} \neq 0}}$$

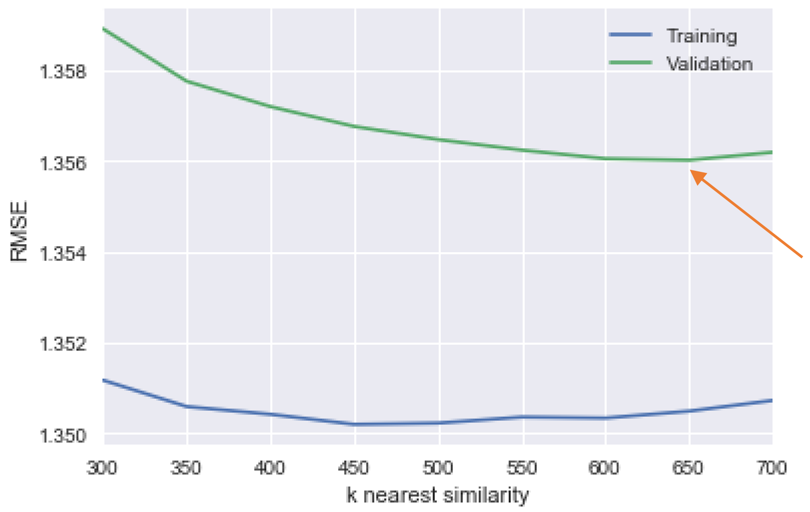
- Baseline method
  - Fill all missing values with item mean

Method	**Test RMSE (bgg-26)**
Baseline	1.368
Neighborhood	1.364

Considering only the *k* most similar users improves predictions

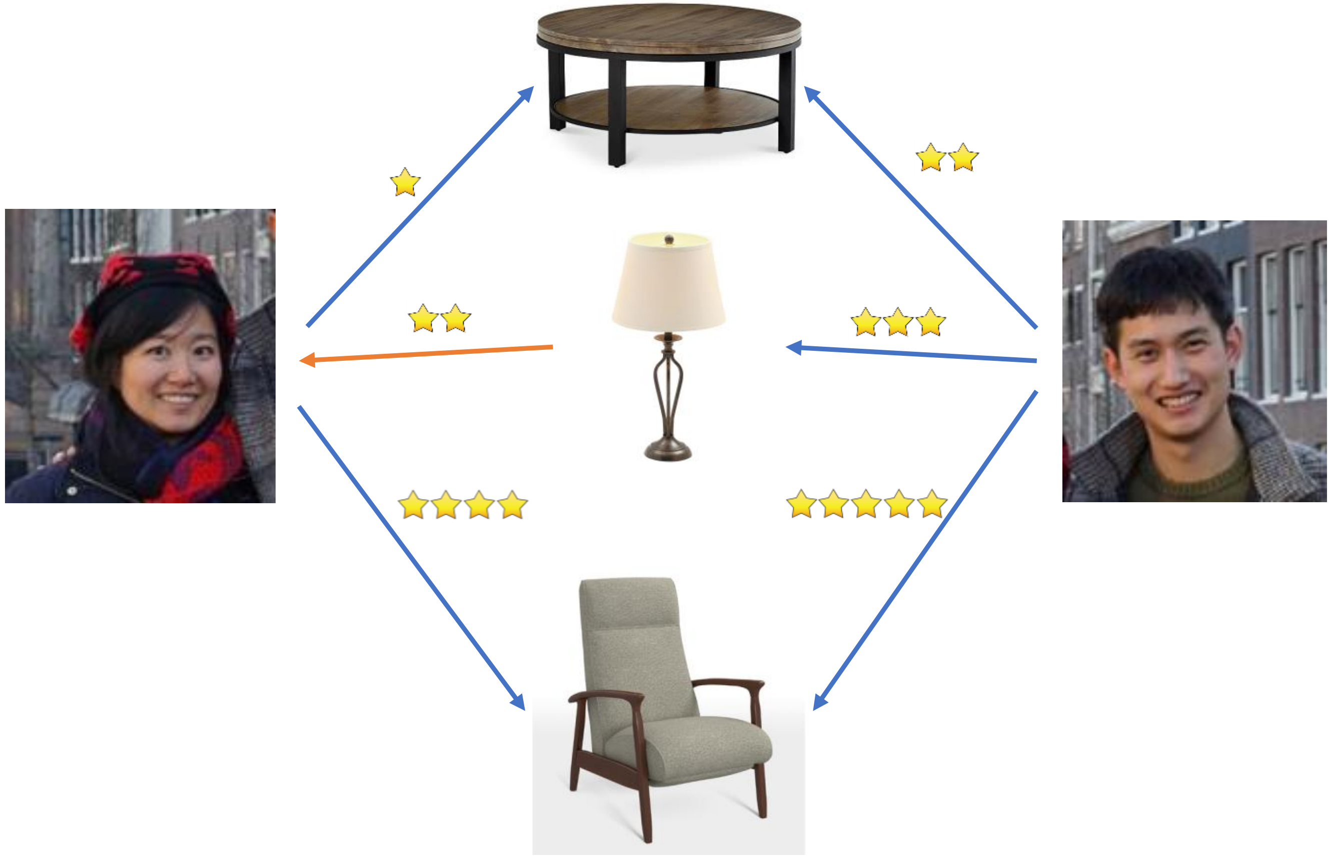
$$s_{u,u'} = \sum_i \frac{r_u r_{u'i}}{(\sum_i r_{ui}^2)^{1/2} (\sum_i r_{u'i}^2)^{1/2}} \qquad \hat{r}_{ui} = \frac{\sum_{u'} s_{u,u'} r_{u',i}}{\sum_{u'} s_{u,u'}}$$

Similar to neighborhood approach except only *k* most similar users are used for prediction



Method	Test RMSE (bgg-26)
Baseline	1.368
Neighborhood	1.364
Top-k	1.356

Correcting for rating scale bias per user

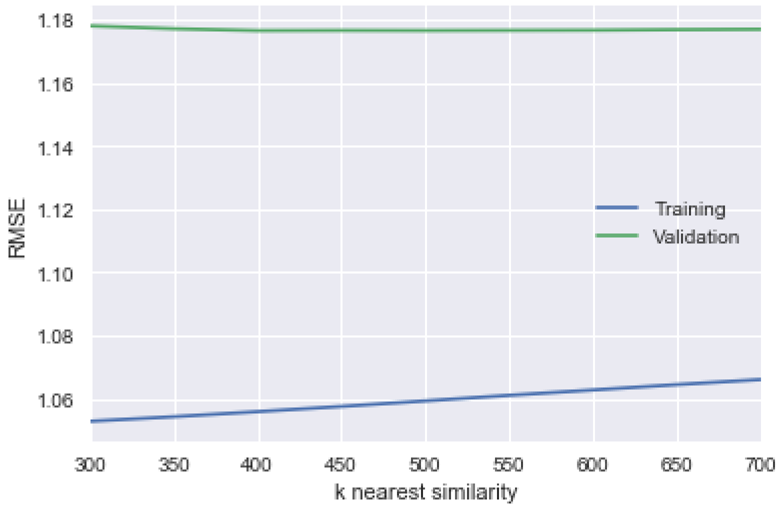


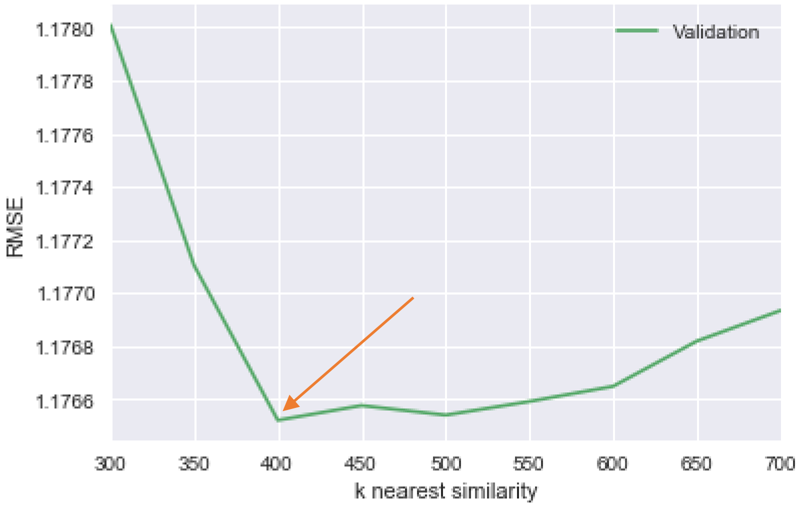
Different users may have similar tastes but have a different interpretation of rating scale values

Centering ratings for each user significantly improves predictions

$$r'_{ui} = r_{ui} - \bar{r}_u$$

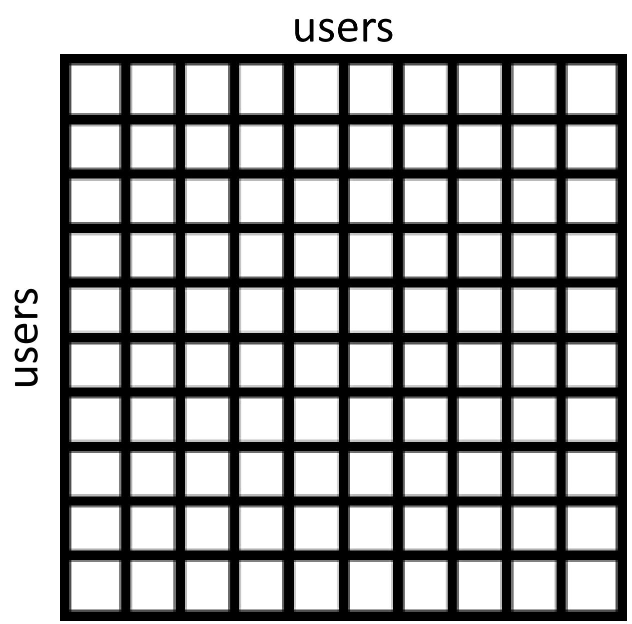
Method	Test RMSE (bgg-26)
Baseline	1.368
Neighborhood	1.364
Top-k	1.356
Neighborhood, bias subtracted	1.184
Top-k, bias subtracted	1.178



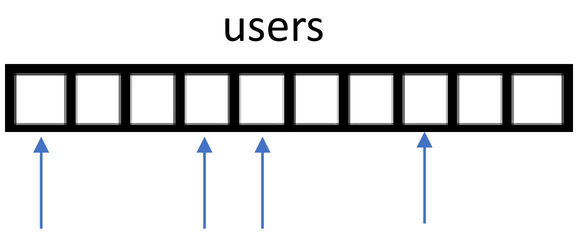


Scalability considerations

$$s_{u,u'} = \sum_i \frac{r_{ui} r_{u'i}}{(\sum_i r_{ui}^2)^{1/2} (\sum_i r_{u'i}^2)^{1/2}}$$

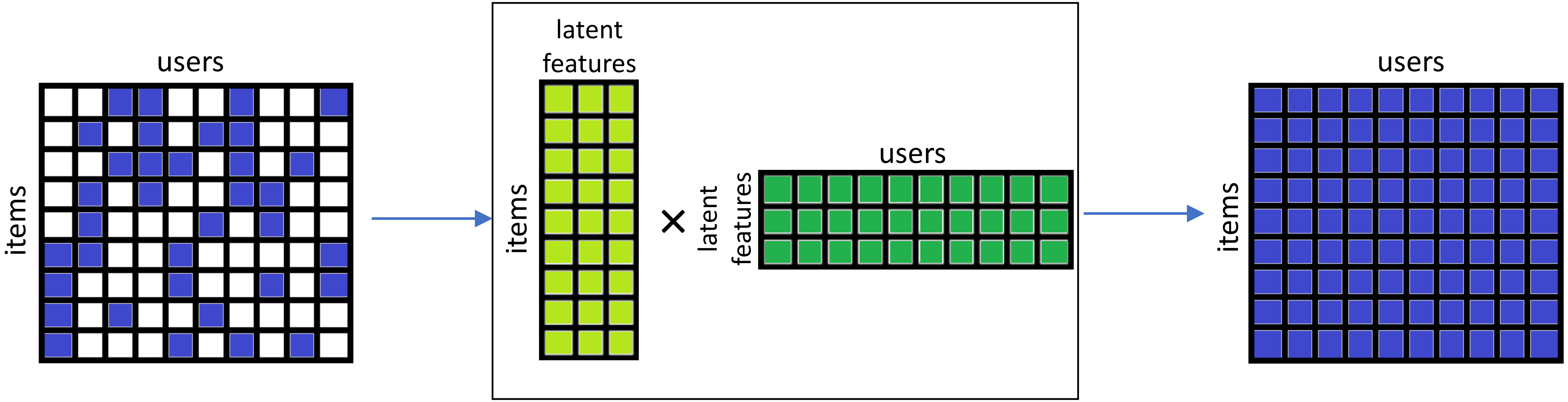


- Typically far more users than items
- Scaling to bgg-86 would result in 58,919 × 58,919 similarity matrix



- In practice, only care about *k* most similar
- *k* most similar may not be best predictors
- May not need to compare every item
- How can we reduce dimensions to best predictors?

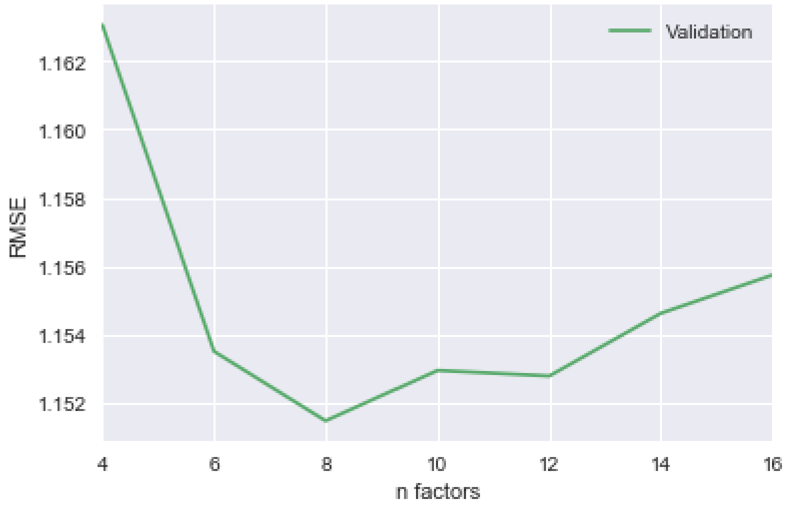
Dimension reduction with singular-value decomposition

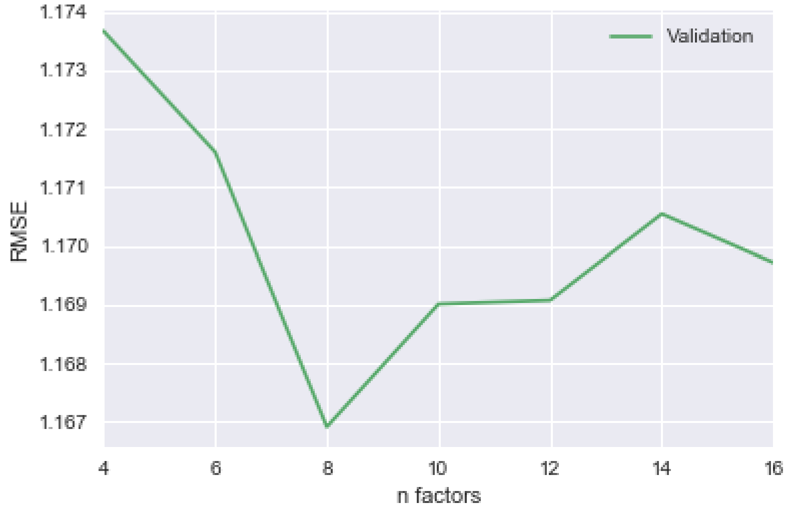


$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T p_u$$

$$\sum_{r_{ui} \in R_{train}} \left(r_{ui} - \hat{r}_{ui}\right)^2 + \lambda \left(b_i^2 + b_u^2 + ||q_i||^2 + ||p_u||^2\right)$$

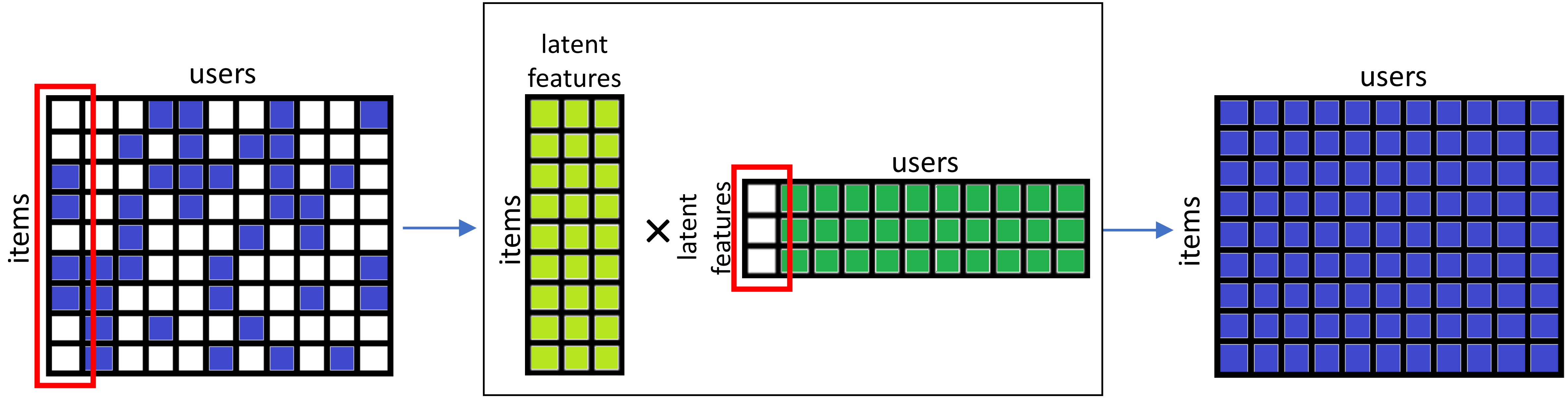
SVD improves prediction and allows scaling to larger datasets





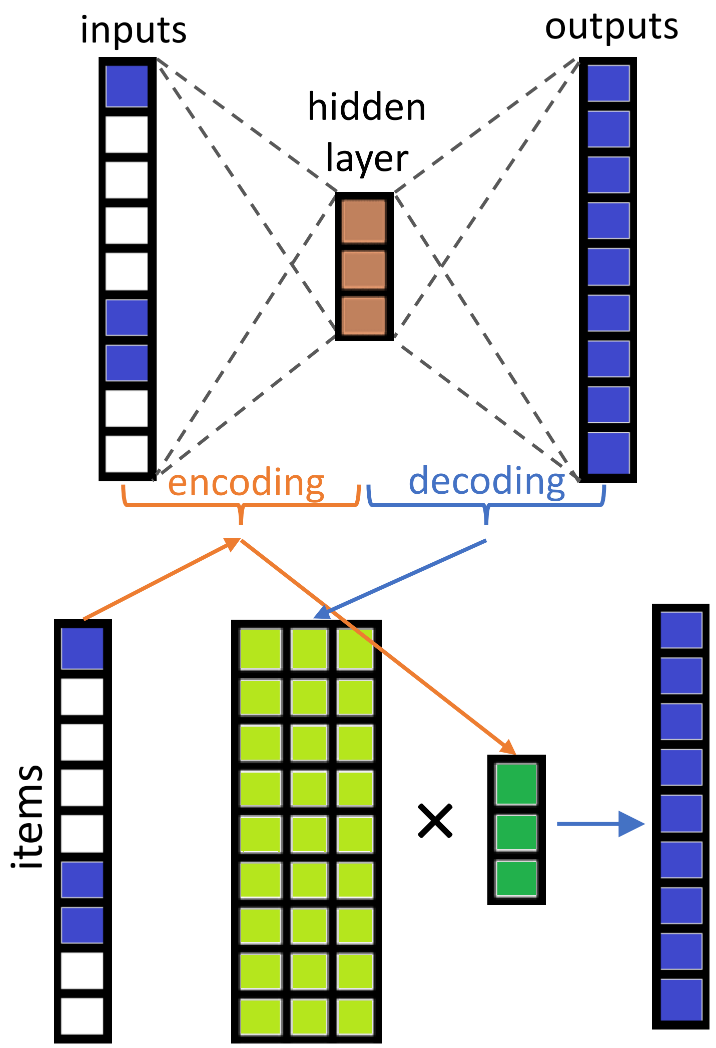
Method	Test RMSE (bgg-26)	Test RMSE (bgg-86)
Baseline	1.368	1.357
Top-k, bias subtracted	1.178	--
SVD	1.154	1.167

SVD does not generalize to new users without additional training



$$\hat{r}_{ui} = \mu + \boxed{b_u} + b_i + q_i^T \boxed{p_u}$$

Autoencoders allow for dimension reduction

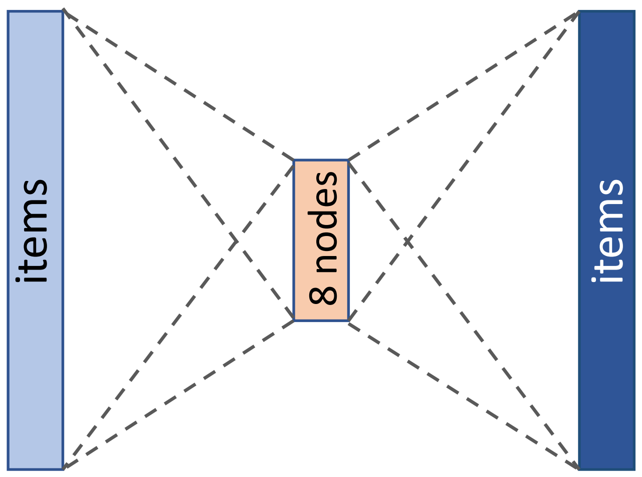


Preprocessing: Bias correction

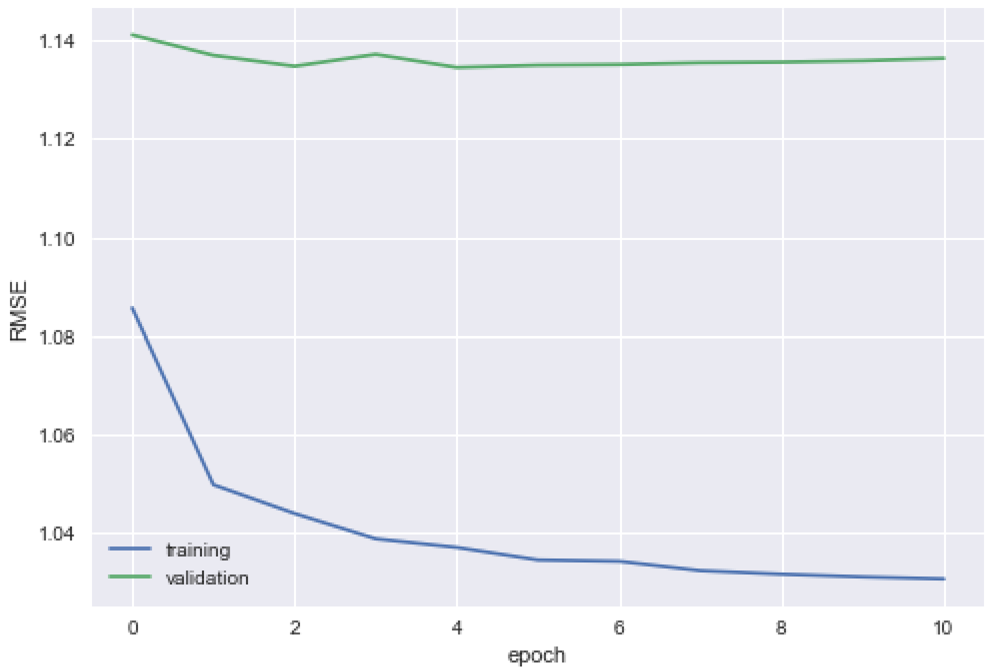
$$r'_{ui} = r_{ui} - \mu - b_i - b_u$$
$$b_i = \bar{r}_i - \mu$$
$$b_u = \bar{r}_u - \mu$$

Assumes enough ratings per user and per item estimate biases

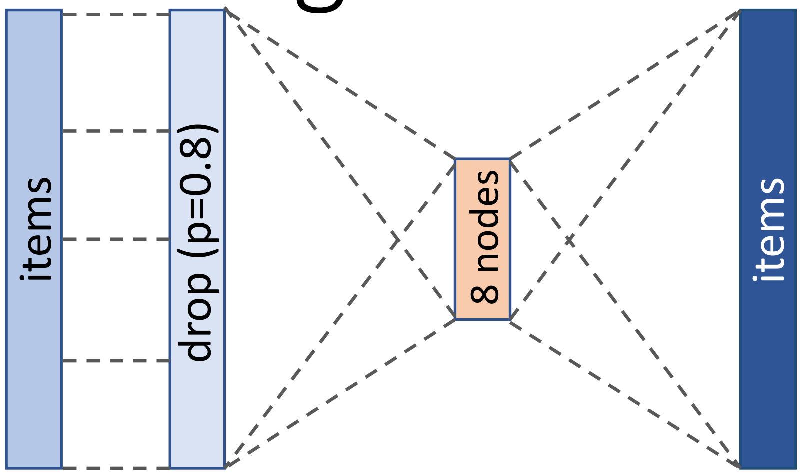
Shallow autoencoder improves on SVD



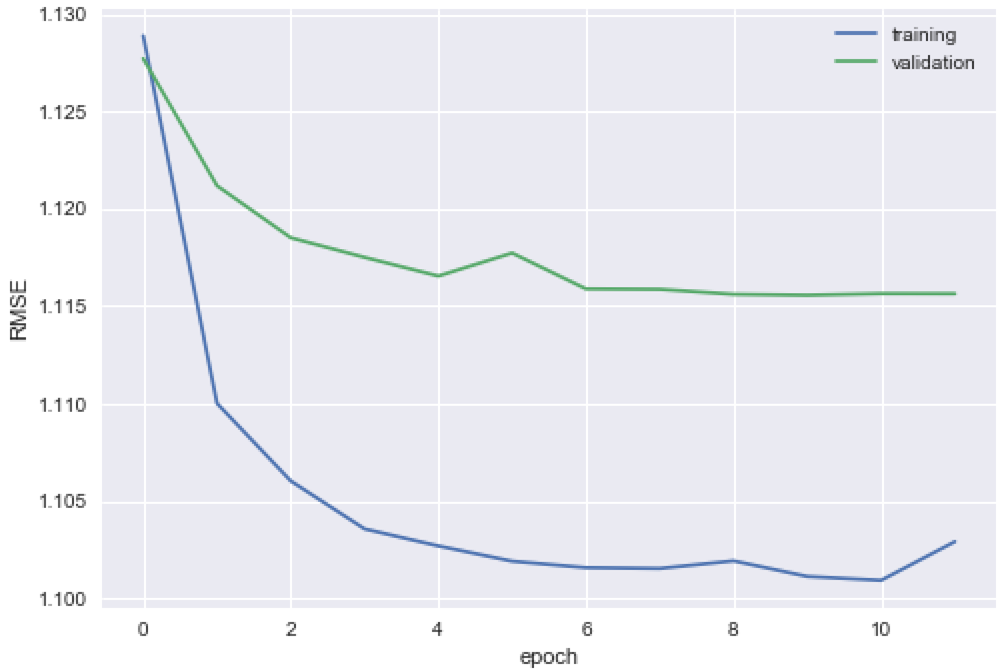
Method	Test RMSE (bgg-86)
Baseline	1.357
SVD	1.167
Shallow AE	1.134



Introducing a dropout layer prevents overfitting

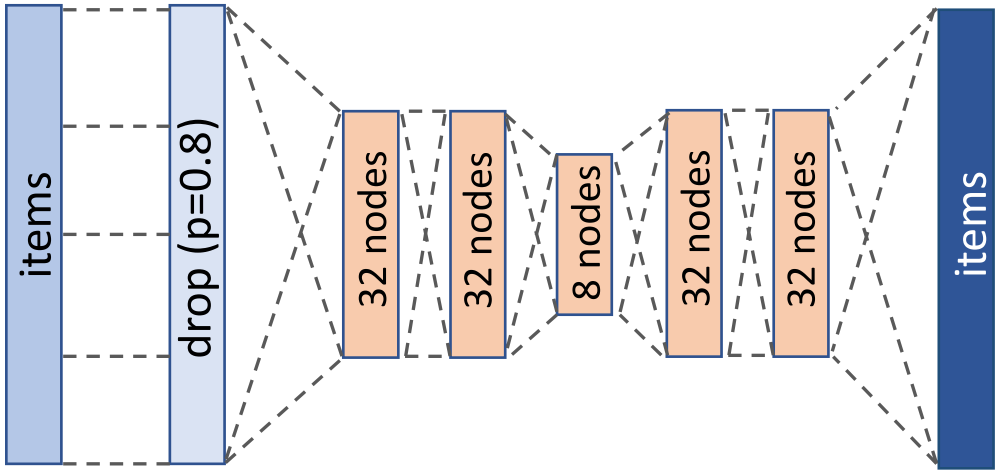


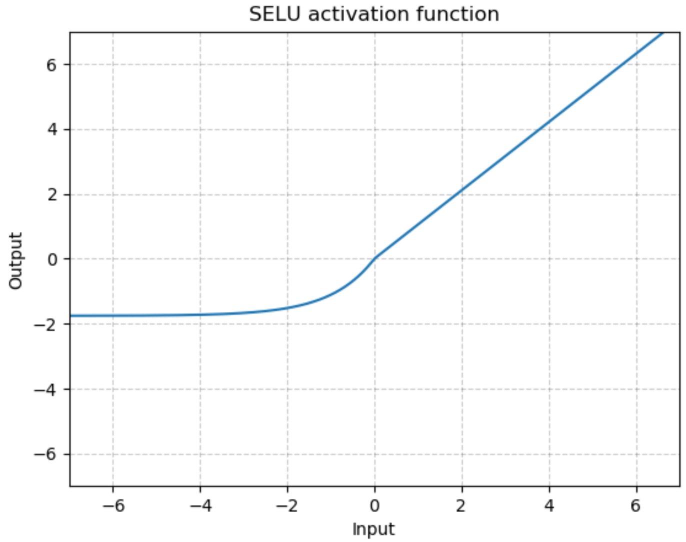
Method	Test RMSE (bgg-86)
Baseline	1.357
SVD	1.167
Shallow AE	1.134
Denoised Shallow AE	1.116



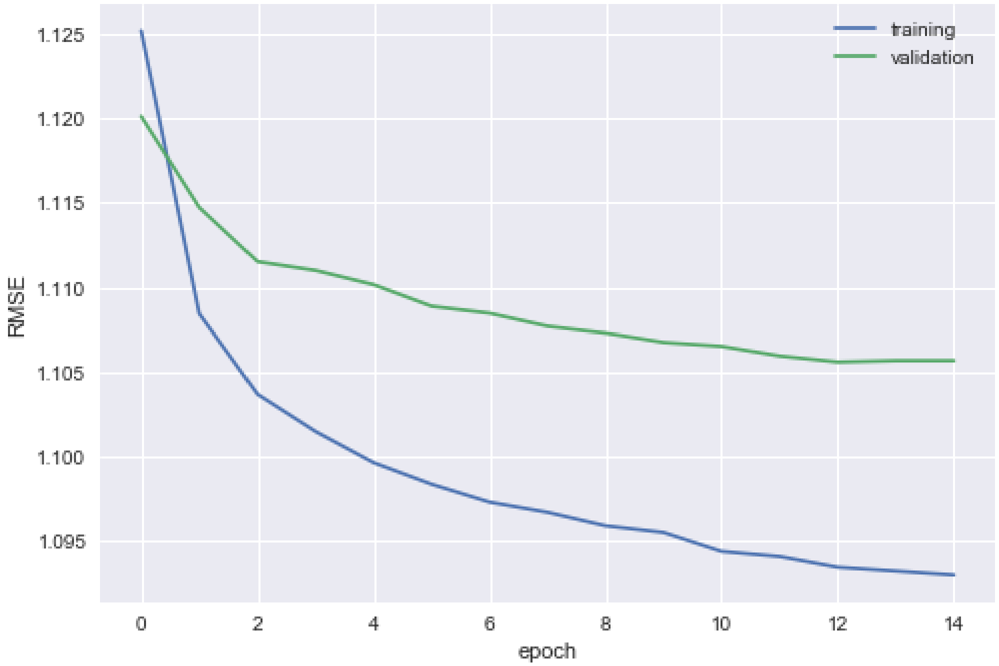
$$\mathrm{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leqslant 0 \end{cases}.$$

Deepening the network

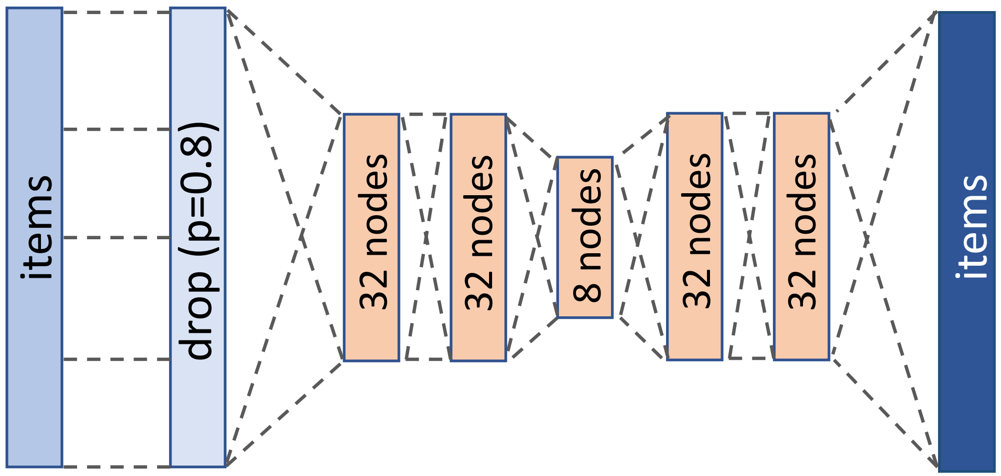




Method	Test RMSE (bgg-86)
Baseline	1.357
SVD	1.167
Shallow AE	1.134
Denoised Shallow AE	1.116
Denoised Deep AE	1.106



Further considerations



- Optimizing number of layers
- Optimizing number of nodes in each layer
- Incorporating user and item biases directly into model
- Adding content-based information
- Postprocessing

Examining predictions

User with 104 ratings

name	my_pred	average_rating	num_voters	popularity
Too Many Bones: Undertow	9.314951	8.96092	631	2718.5
Here I Stand (500th Anniversary Reprint Edition)	9.118732	8.69876	225	4417.0
1844: Switzerland	9.106171	8.33923	180	4686.0
Axis Empires: Totaler Krieg!	9.041317	8.37052	229	4392.5
1822: The Railways of Great Britain	9.002617	8.50138	182	4669.0
Brass: Birmingham	8.982116	8.63405	4815	516.0
Enemy Action: Ardennes	8.963823	8.67158	297	3973.5
1817	8.911486	8.73544	268	4154.5
Gaia Project	8.909788	8.56949	8593	273.0
Riichi Mahjong	8.837430	8.43058	188	4632.5

	movie_id	prediction	title	genres
0	953	4.868923	It's a Wonderful Life (1946)	Drama
1	668	4.866858	Pather Panchali (1955)	Drama
2	1423	4.859523	Hearts and Minds (1996)	Drama
3	3307	4.834415	City Lights (1931)	Comedy|Drama|Romance
4	649	4.802675	Cold Fever (Á köldum klaka) (1994)	Comedy|Drama
5	669	4.797451	Aparajito (1956)	Drama
6	326	4.784828	To Live (Huozhe) (1994)	Drama
7	3092	4.761148	Chushingura (1962)	Drama
8	3022	4.753003	General, The (1927)	Comedy
9	2351	4.720692	Nights of Cabiria (Le Notti di Cabiria) (1957)	Drama
10	926	4.719633	All About Eve (1950)	Drama

Postprocessing

Original

gameid	name	my_pred	num_voters	popularity
235802	Too Many Bones: Undertow	9.314951	631	2718.5
242722	Here I Stand (500th Anniversary Reprint Edition)	9.118732	225	4417.0
7935	1844: Switzerland	9.106171	180	4686.0
32989	Axis Empires: Totaler Krieg!	9.041317	229	4392.5
193867	1822: The Railways of Great Britain	9.002617	182	4669.0
224517	Brass: Birmingham	8.982116	4815	516.0
68820	Enemy Action: Ardennes	8.963823	297	3973.5
63170	1817	8.911486	268	4154.5
220308	Gaia Project	8.909788	8593	273.0
108018	Riichi Mahjong	8.837430	188	4632.5

Filtered by minimum popularity rank

$$(popularity\ rank) \leq 10 * (number\ of\ ratings)$$

name	my_pred	average_rating	num_voters	popularity
Brass: Birmingham	8.982116	8.63405	4815	516.0
Gaia Project	8.909788	8.56949	8593	273.0
Gloomhaven	8.790935	8.91228	23504	60.0
Vinhos Deluxe Edition	8.729107	8.24893	2519	956.5
Food Chain Magnate	8.683999	8.18605	9311	256.0
Spirit Island	8.651888	8.38891	9757	233.0
Lisboa	8.616892	8.19692	3845	640.0
Great Western Trail	8.603271	8.28808	18302	95.0
Brass: Lancashire	8.526196	8.11884	14647	140.0
Indonesia	8.498173	7.83536	3486	714.0

- Other postprocessing methods:
  - Impression discounting
  - Dithering
  - “Common sense”

Final thoughts

- Collaborative filtering effective for recommender systems
- Dimension reduction useful for more accurate modeling and improved scalability
- Deep neural networks capture higher order interactions and allow for additional processing
- Postprocessing necessary to improve recommendation relevance

Final thoughts

- Collaborative filtering effective for recommender systems
- Dimension reduction useful for more accurate modeling and improved scalability
- Deep neural networks capture higher order interactions and allow for additional processing
- Postprocessing necessary to improve recommendation relevance