

Recommender Systems

lecture 1: introduction

Alexey Grishanov

Moscow Institute of Physics and Technology

Spring 2024

«What is recommender system?» evolution:

In a typical recommender system people provide recommendations as inputs, which the system then aggregates and directs to appropriate recipients. In some cases the primary transformation is in the aggregation; in others the system's value lies in its ability to make good matches between the recommenders and those seeking recommendations.

● **Resnick, P. and Varian, H. R.: 1997, 'Recommender Systems'. Communications of the ACM**

More formally, the recommendation problem can be formulated as follows: Let C be the set of all users and let S be the set of all possible items that can be recommended. Let u be a utility function that measures the usefulness of item s to user c , that is, $u: C \times S \Rightarrow \mathbb{R}$, where \mathbb{R} is a totally ordered set (for example, nonnegative integers or real numbers within a certain range). Then, for each user $c \in C$, we want to choose such item $s' \in S$ that maximizes the user's utility.

● **Adomavicius, G., and Tuzhilin, A. 2005. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering**

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	0.8564	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	Qora Solutions and Vandelv United	0.8573	9.89	2009-07-25 20:05:52
5	Vandelv Industries I	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	Qora_	0.8603	9.58	2009-07-24 17:18:43
9	Qora Solutions	0.8611	9.49	2009-07-26 18:02:08
10	BellKor	0.8612	9.48	2009-07-26 17:18:11
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	xianjiang	0.8633	9.26	2009-07-21 02:04:40
14	Gruyter	0.8634	9.25	2009-07-26 15:58:34
15	Cms	0.8642	9.17	2009-07-25 17:42:38
16	Invisible Ideas	0.8644	9.14	2009-07-20 03:28:12
17	Just a guy in a garage	0.8660	8.98	2009-07-22 14:10:42
18	Craig Carmichael	0.8666	9.02	2009-07-25 16:00:54
19	J.Dennis Su	0.8668	9.00	2009-03-11 08:41:54
20	actvhill	0.8669	8.99	2009-04-16 06:29:35
Progress Prize 2007 - RMSE = 0.8712 - Winning Team: KorBell				
Cinematch score on split subset - RMSE = 0.9514				

One of the first and most notable RecSys competitions (02.10.2006 - 26.06.2009)
netflixprize.com/leaderboard

Motivation: money

Amazon

- 35% of sales estimated to be generated from recommendation system
- billions of dollars in revenue generated through personalized product recommendations

Netflix

- 80% of content watched on Netflix is a result of recommendation
- "the combined effect of personalization and recommendations save us more than \$1B per year".^a

^a<https://www.businessinsider.com/netflix-recommendation-engine-worth-1-billion-per-year-2016-6>

Motivation: beyond money

- 1 **Better discovery & diversity of recommendations:** broaden users' perspectives, encourage them to try new things, and create a more inclusive and diverse experience.
- 2 **Increased user engagement:** provide personalized and relevant experience for users to help users find what they're looking for and engage with the service.
- 3 **Better use of data:** provide insights into user behavior which helps businesses make more informed decisions.

My team's mission:

We connect **listeners and **creators** in a unique and enriching way**

Example mission (Spotify), from: <https://www.youtube.com/watch?v=Ysh0CXUbXsU>

Recommender Systems taxonomy

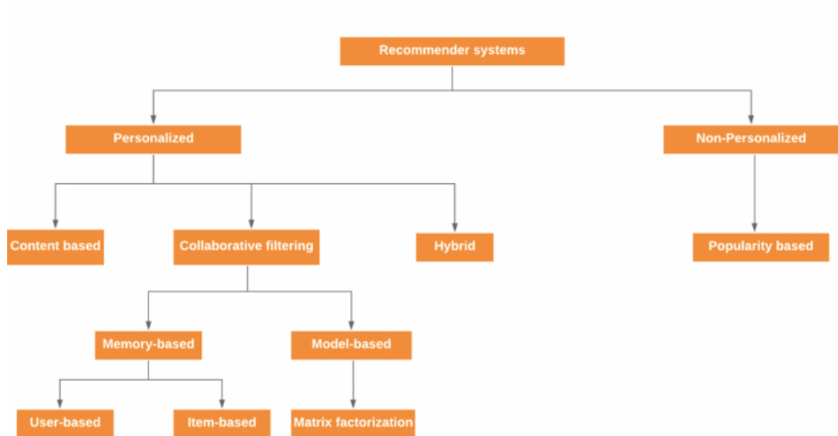


image credit: <https://thingsolver.com/introduction-to-recommender-systems>

Problem statement

Given:

- $U = \{u_j | j \in 1, \dots, n_{users}\}$ — set of users
- $I = \{i_j | j \in 1, \dots, n_{items}\}$ — set of items.
- $R = \|r_{ui}\|$ — relation matrix of shape $n_{users} \times n_{items}$,

$r_{ui} \in \begin{cases} \text{typically } \{0, 1\} \text{— implicit feedback} \\ \text{typically } \{1, 2, 3, 4, 5\} \text{— explicit feedback} \end{cases}$

Possible tasks:

Rating prediction

- Predict unknown r_{ui} — regression or (multi-class) classification

Top-k ranking

- Rank top-k recommendations for items — item2item
- Rank top-k recommendations for users — **item2user (course focus)**

Evaluation metrics (item2user)

1 nDCG

$$DCG@k(u) = \sum_{j=1}^k \frac{2^{r_{uj}} - 1}{\log_2(j+1)}$$

$$IDCG@k(u) = \max(DCG@k(u))$$

$$nDCG@k(u) = \frac{DCG@k(u)}{IDCG@k(u)}$$

$$nDCG@k = \frac{\sum_{u=1}^{n_{users}} nDCG@k(u)}{n_{users}}$$

2 Mean Average Precision

$$p@k(u) = \frac{\sum_{j=1}^k [r_{uj} > 0]}{k}$$

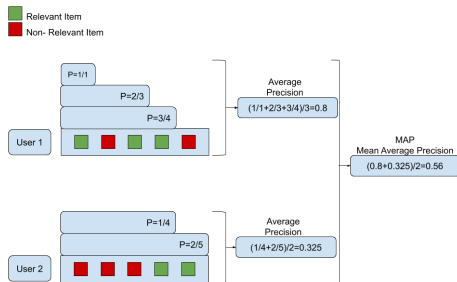
$$ap@k(u) = \frac{\sum_{j=1}^m [r_{uj} > 0] p@j(u)}{m}$$

$$MAP@k = \frac{\sum_{u=1}^{n_{users}} ap@k(u)}{n_{users}}$$

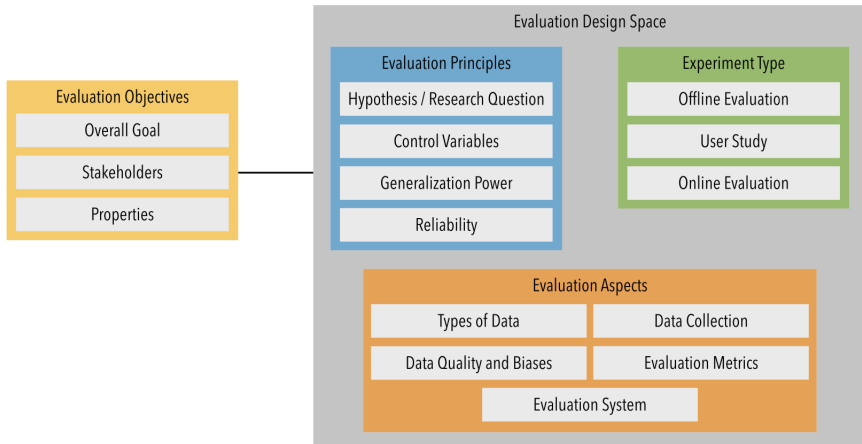
3 etc.

image credit:

<https://medium.com/swlh/rank-aware-recsys-evaluation-metrics-5191bba16832>



Overall evaluation



Evaluation framework example

image credit: <https://dl.acm.org/doi/10.1145/3556536>

- ① *Charu C. Aggarwal*. (2016). Recommender Systems: The Textbook.
- ② *F. Ricci, L. Rokach, B. Shapira*. (2011). Recommender Systems Handbook.
- ③ *Adomavicius, G., Tuzhilin, A.* (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734-749.
- ④ *Eva Zangerle, Christine Bauer* (2022). Evaluating Recommender Systems: Survey and Framework. *ACM Computing Surveys*, Volume 55, Issue 8, Article No.: 170, pp 1–38