Recommender Systems
lecture 1: introduction

Alexey Grishanov

Moscow Institute of Physics and Technology

Fall 2023

Recommender Systems taxonomy

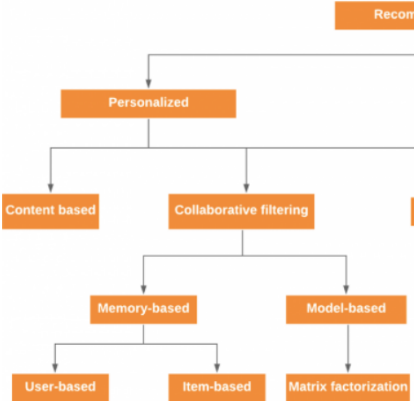


image credit: https://thingsolver.com/introduction-to-recommender-systems

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

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 $n_{users} \times n_{items}$,

$$r_{ui} \in \left[\begin{array}{c} \text{typically } \{0,1\}\text{— implicit feedback} \\ \text{typically } \{1,2,3,4,5\}\text{— explicit feedback} \end{array} \right.$$

Possible tasks:

Rating prediction

- Predict unknown r_{ui} (regression)

Top-k recommendation

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

1. **Normalized Discounted Cumulative Gain**

$$nDCG@k = \frac{DCG@k}{IDCG@k},$$

IDCG — ideal DCG@k.

$$DCG@k = \sum_{j=1}^{k} \frac{2^{y_j} - 1}{\log_2 j + 1}.$$

2. **Mean Average Precision**

$$MAP@k = \frac{1}{n_{users}} ap@k$$

$$ap@k = \frac{1}{k} \sum_{j=1}^{k} [y_j > 0] \, p@k$$

$$p@k = \frac{1}{k} \sum_{j=1}^{k} y_j$$

3. etc.

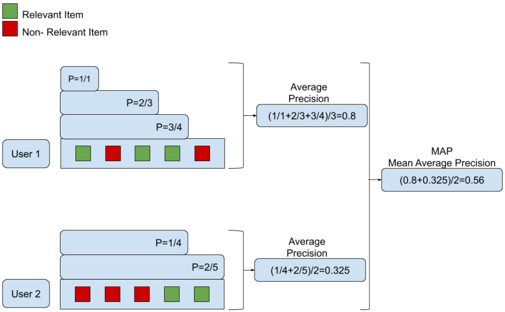
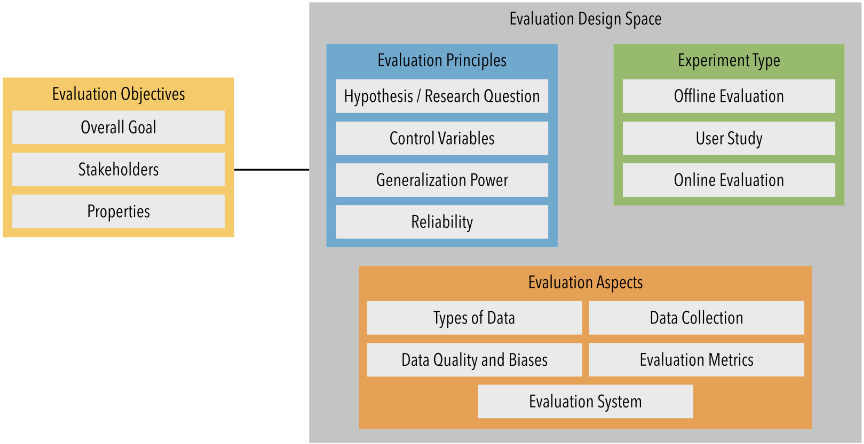


image credit:
https://medium.com/swlh/rank-aware-recsys-evaluation-metrics-5191bba16832

Overall evaluation



Evaluation framework example

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

Literature

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