Recommender Systems

Marie Al Ghossein

June 10, 2024

Crossing Minds

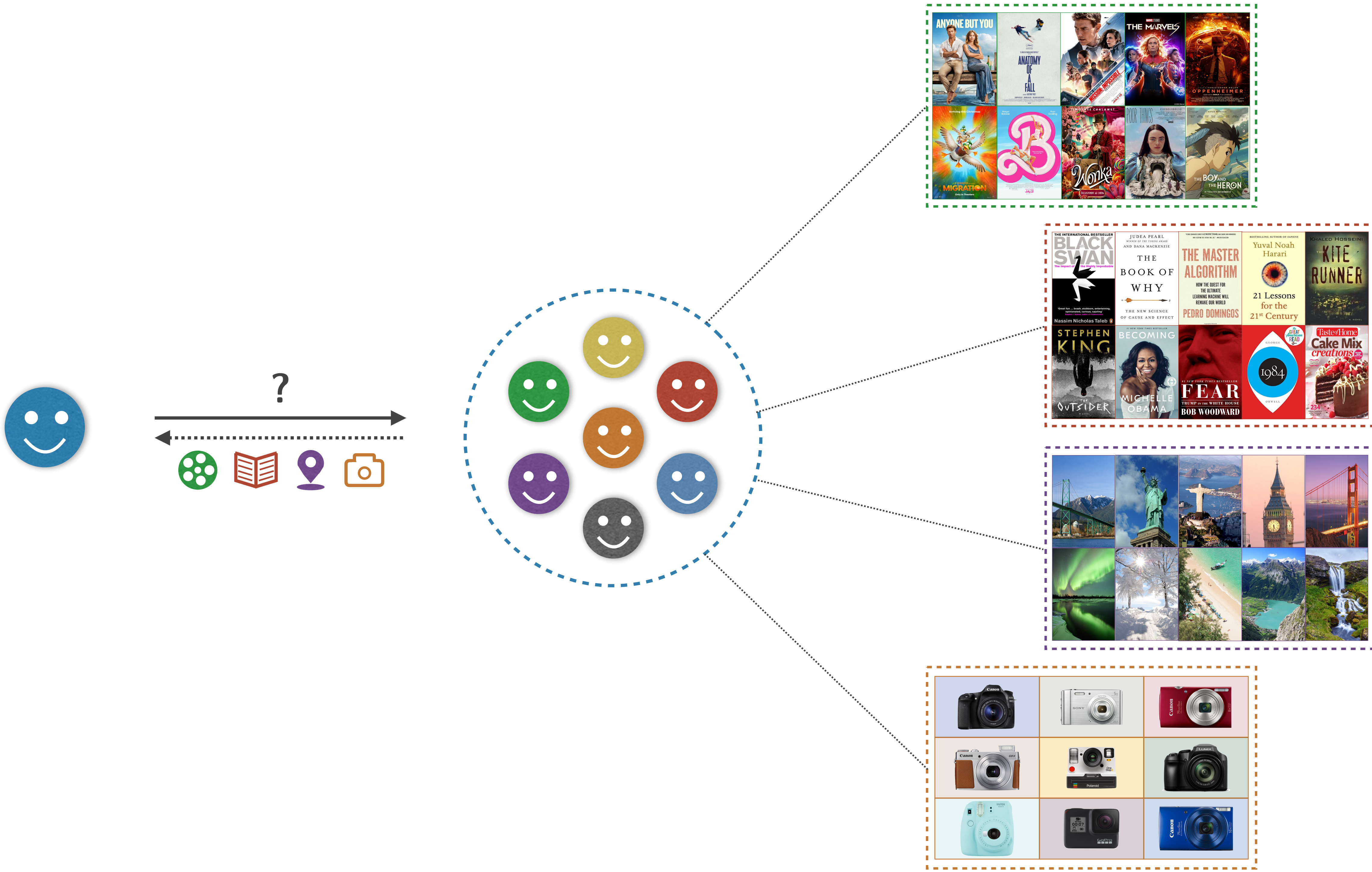
Dauphine | PSL
UNIVERSITÉ PARIS

Outline

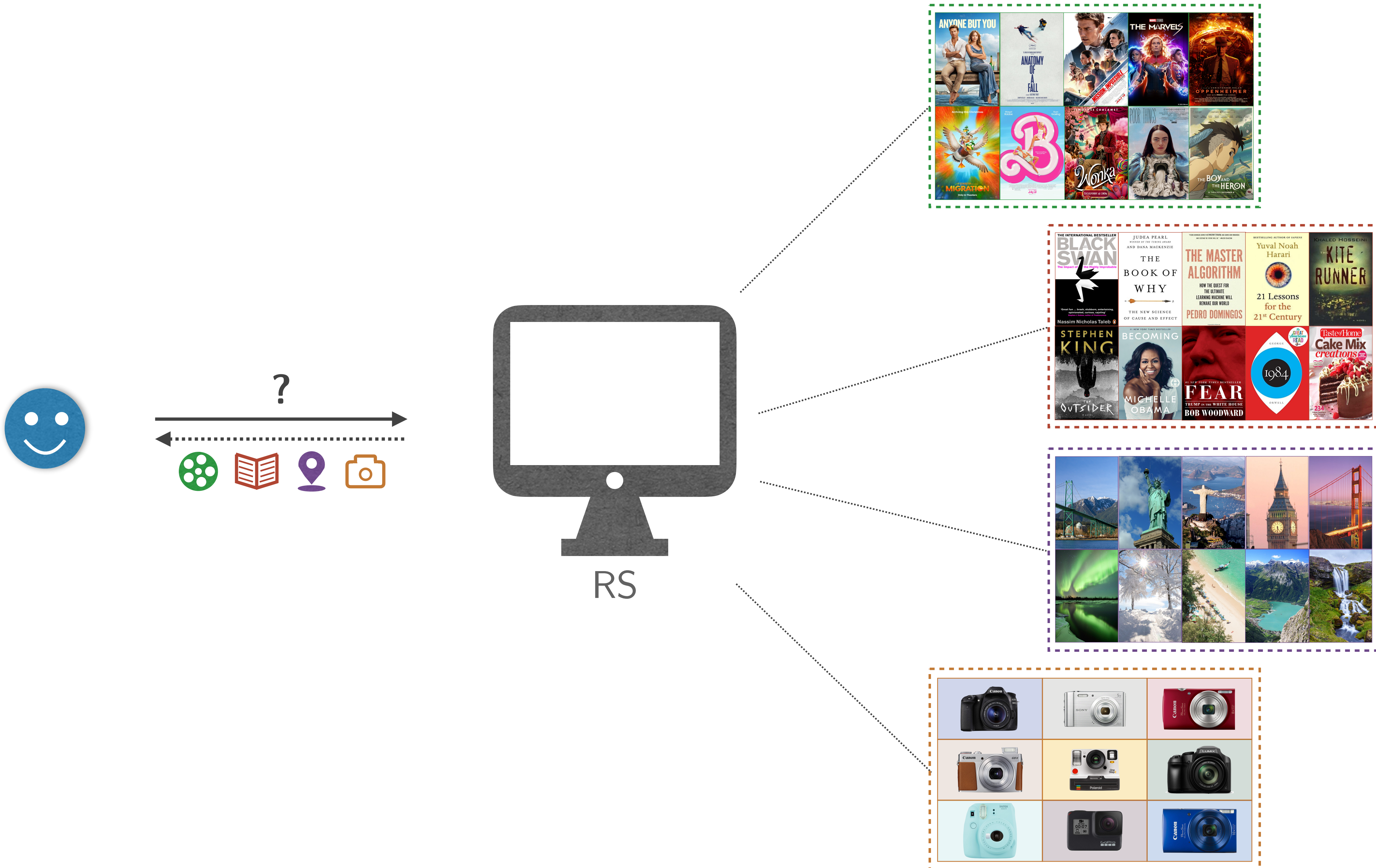
1. Introduction
2. The Recommendation Problem
3. User Feedback
4. Challenges
5. Evaluation of RS
6. Collaborative Filtering
7. Content-Based Filtering
8. Hybrid Approaches
9. Context-Aware Recommendation
10. Reinforcement Learning for Recommender Systems
11. Deep Learning for Recommender Systems
12. Short Intro to Large Language Models for Recommender Systems
13. A Practical Example: Recommendation in Location-Based Social Networks

Introduction

The (black) art of recommendation



The (black) art of recommendation



The value of recommender systems

For consumers

- Find good and relevant products
- Improve decision making
- Help them overcome the *information overload*

The value of recommender systems

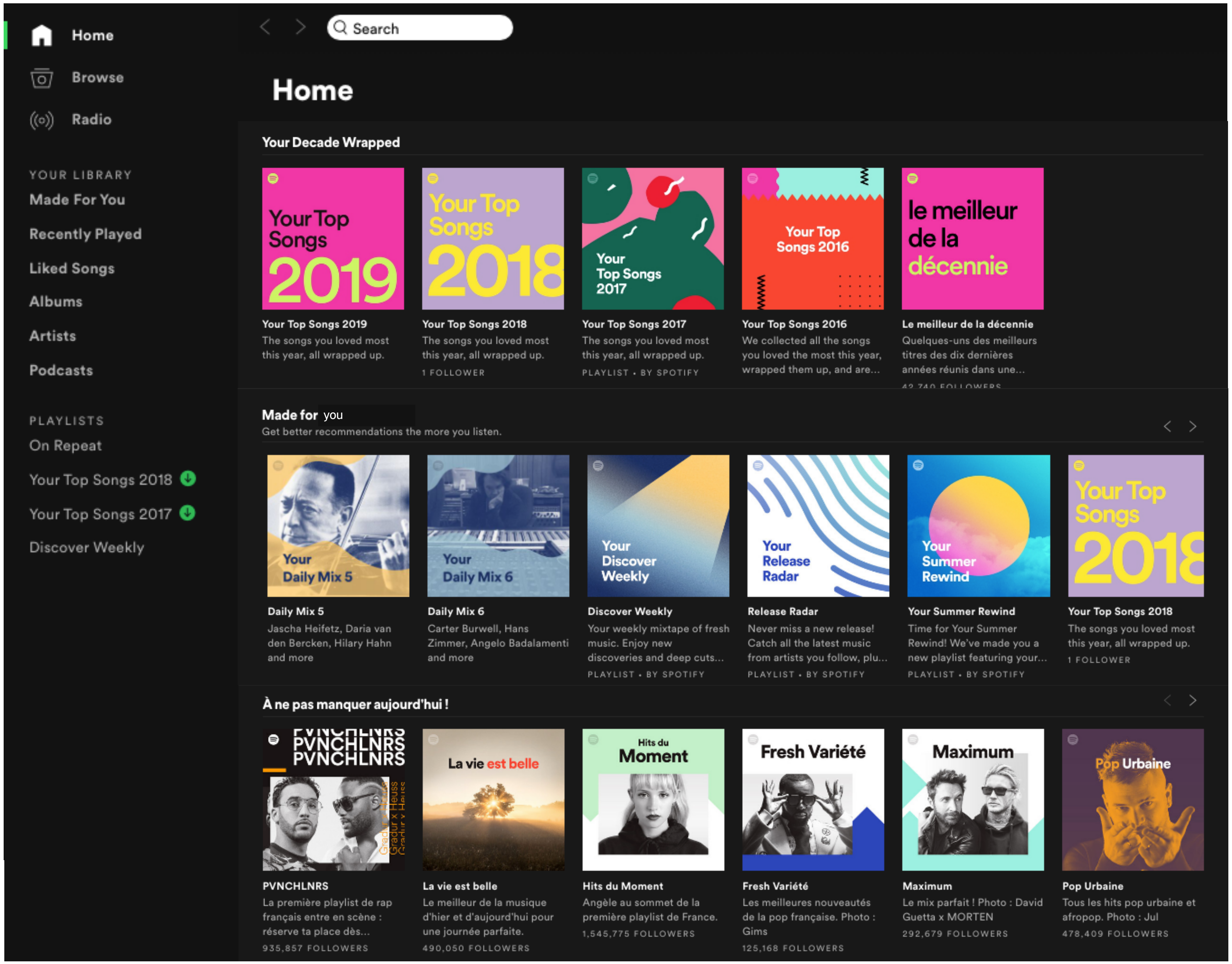
For consumers

- Find good and relevant products
- Improve decision making
- Help them overcome the *information overload*

For service providers

- Increase the number of products sold
- Sell more diverse products
- Increase the user satisfaction and fidelity
- Learn more about consumers

Personalization at Spotify



The value of recommender systems

"Our recommender system influences choice for about 80% of hours streamed at Netflix. [...] We think the combined effect of personalization and recommendations save us more than $1B per year." (2015)

C. A. Gomez-Uribe and N. Hunt. *The Netflix Recommender System: Algorithms, Business Value, and Innovation.* ACM Trans. Management Inf. Syst. 6(4), 2015.

The value of recommender systems

"Our recommender system influences choice for about 80% of hours streamed at Netflix. [...] We think the combined effect of personalization and recommendations save us more than $1B per year." (2015)

C. A. Gomez-Uribe and N. Hunt. *The Netflix Recommender System: Algorithms, Business Value, and Innovation.* ACM Trans. Management Inf. Syst. 6(4), 2015.

"35% of what consumers purchase on Amazon comes from product recommendations." (2013)

https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers

The value of recommender systems

"Our recommender system influences choice for about 80% of hours streamed at Netflix. [...] We think the combined effect of personalization and recommendations save us more than $1B per year." (2015)

C. A. Gomez-Uribe and N. Hunt. *The Netflix Recommender System: Algorithms, Business Value, and Innovation*. ACM Trans. Management Inf. Syst. 6(4), 2015.

"35% of what consumers purchase on Amazon comes from product recommendations." (2013)

https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers

"Recommendations account for about 60% of all video clicks from the homepage." (2010)

J. Davidson et al. *The YouTube Video Recommendation System*. In RecSys'10, 2010.

The value of recommender systems

"Our recommender system influences choice for about 80% of hours streamed at Netflix. [...] We think the combined effect of personalization and recommendations save us more than $1B per year." (2015)

C. A. Gomez-Uribe and N. Hunt. *The Netflix Recommender System: Algorithms, Business Value, and Innovation.* ACM Trans. Management Inf. Syst. 6(4), 2015.

"35% of what consumers purchase on Amazon comes from product recommendations." (2013)

https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers

"Recommendations account for about 60% of all video clicks from the homepage." (2010)

J. Davidson et al. *The YouTube Video Recommendation System.* In RecSys'10, 2010.

"Recommendations generate 38% more clicks compared to popularity-based recommendations." (2009)

A. S. Das et al. *Google News Personalization: Scalable Online Collaborative Filtering.* In WWW '07, 2007.

Measuring the business value of recommender systems

- Click-Through Rate (CTR)

Dietmar Jannach and Michael Jugovac. *Measuring the Business Value of Recommender Systems*. arXiv 2019

Measuring the business value of recommender systems

- Click-Through Rate (CTR)
- Conversion rates

Dietmar Jannach and Michael Jugovac. *Measuring the Business Value of Recommender Systems*. arXiv 2019

Measuring the business value of recommender systems

- Click-Through Rate (CTR)
- Conversion rates
- Sales and revenue

Dietmar Jannach and Michael Jugovac. *Measuring the Business Value of Recommender Systems.* arXiv 2019

Measuring the business value of recommender systems

- Click-Through Rate (CTR)
- Conversion rates
- Sales and revenue
- Effects on sales distributions

Dietmar Jannach and Michael Jugovac. *Measuring the Business Value of Recommender Systems*. arXiv 2019

Measuring the business value of recommender systems

- Click-Through Rate (CTR)
- Conversion rates
- Sales and revenue
- Effects on sales distributions
- User engagement and behavior

Dietmar Jannach and Michael Jugovac. *Measuring the Business Value of Recommender Systems*. arXiv 2019

What is a recommender system?

→ A Recommender System (RS) provides personalized suggestions of **items** for **users**, the items being drawn from a large catalog, based on knowledge of the user, the items, and **interactions** between users and items.

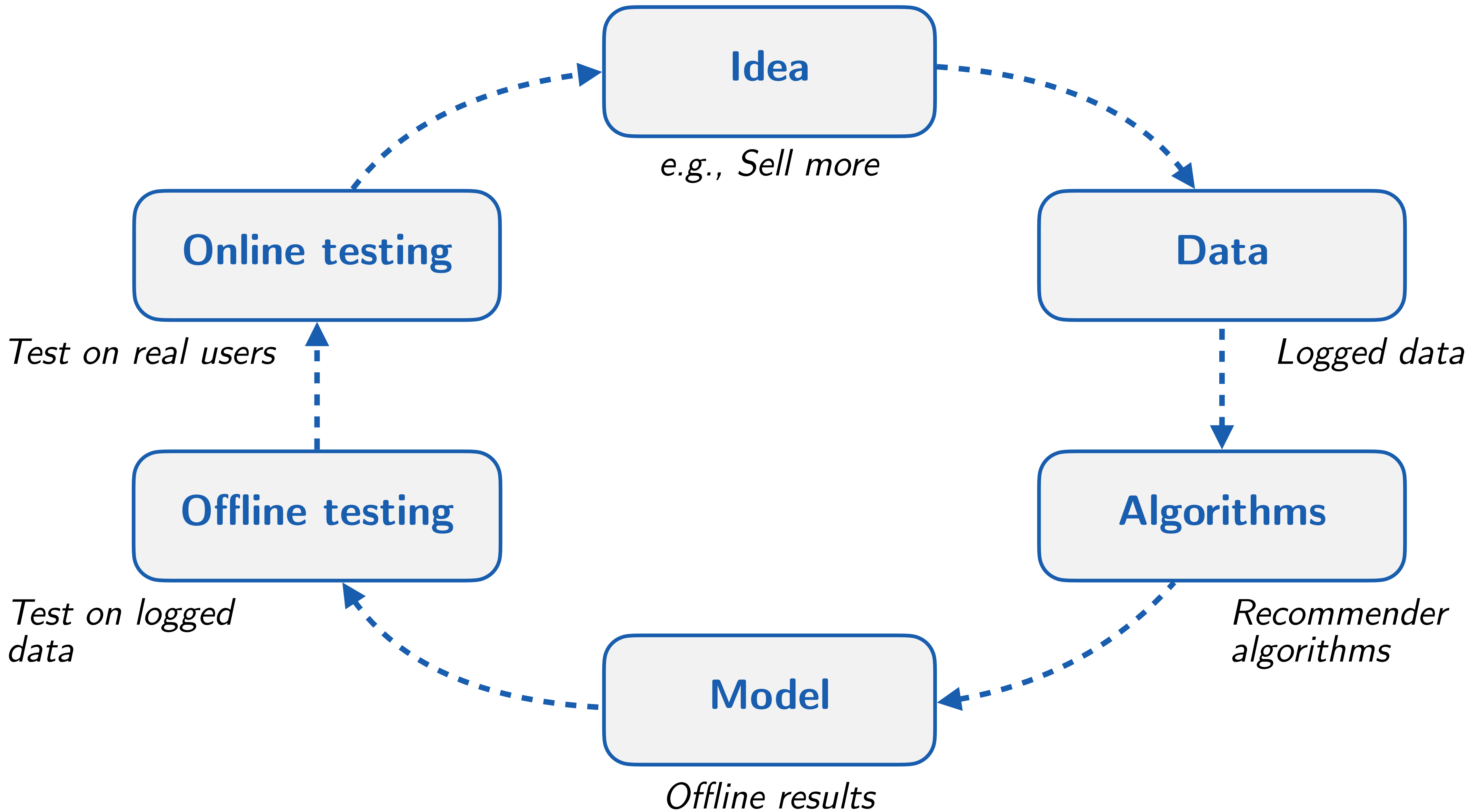
What is a recommender system?

- Knowledge of the user
  - Historical data - interactions with items
    - Purchases, clicks, views, ratings,...
  - Demographics
    - Age, gender, nationality, political views, income,...
  - Context
    - Device, location, time of day/week/year, alone or in a group,...

What is a recommender system?

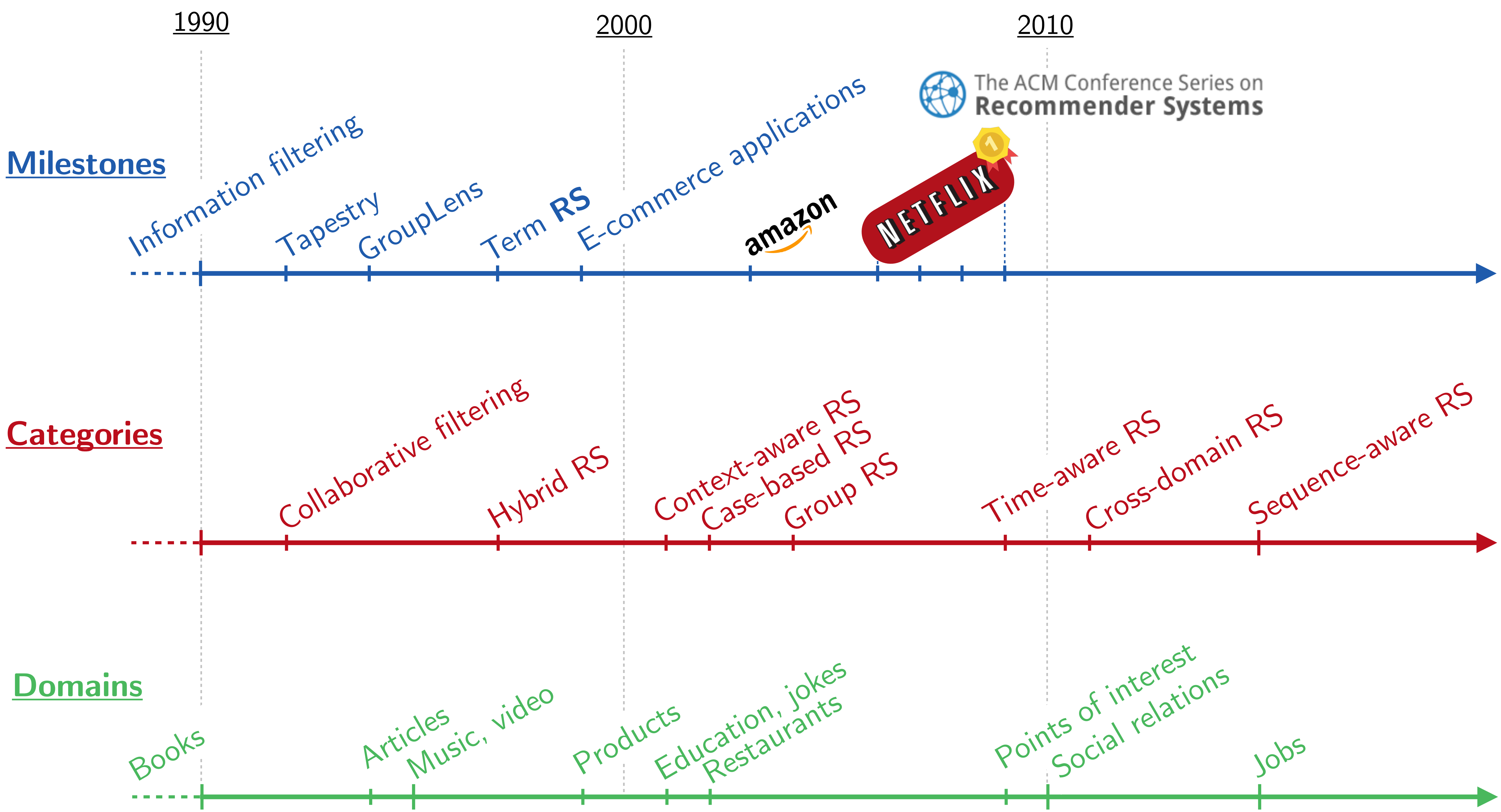
- Knowledge of the user
  - Historical data - interactions with items
    - Purchases, clicks, views, ratings,...
  - Demographics
    - Age, gender, nationality, political views, income,...
  - Context
    - Device, location, time of day/week/year, alone or in a group,...
- Knowledge of the item
  - Content information
    - Features, tags, text, images, multimedia signals,...

Building a recommender system

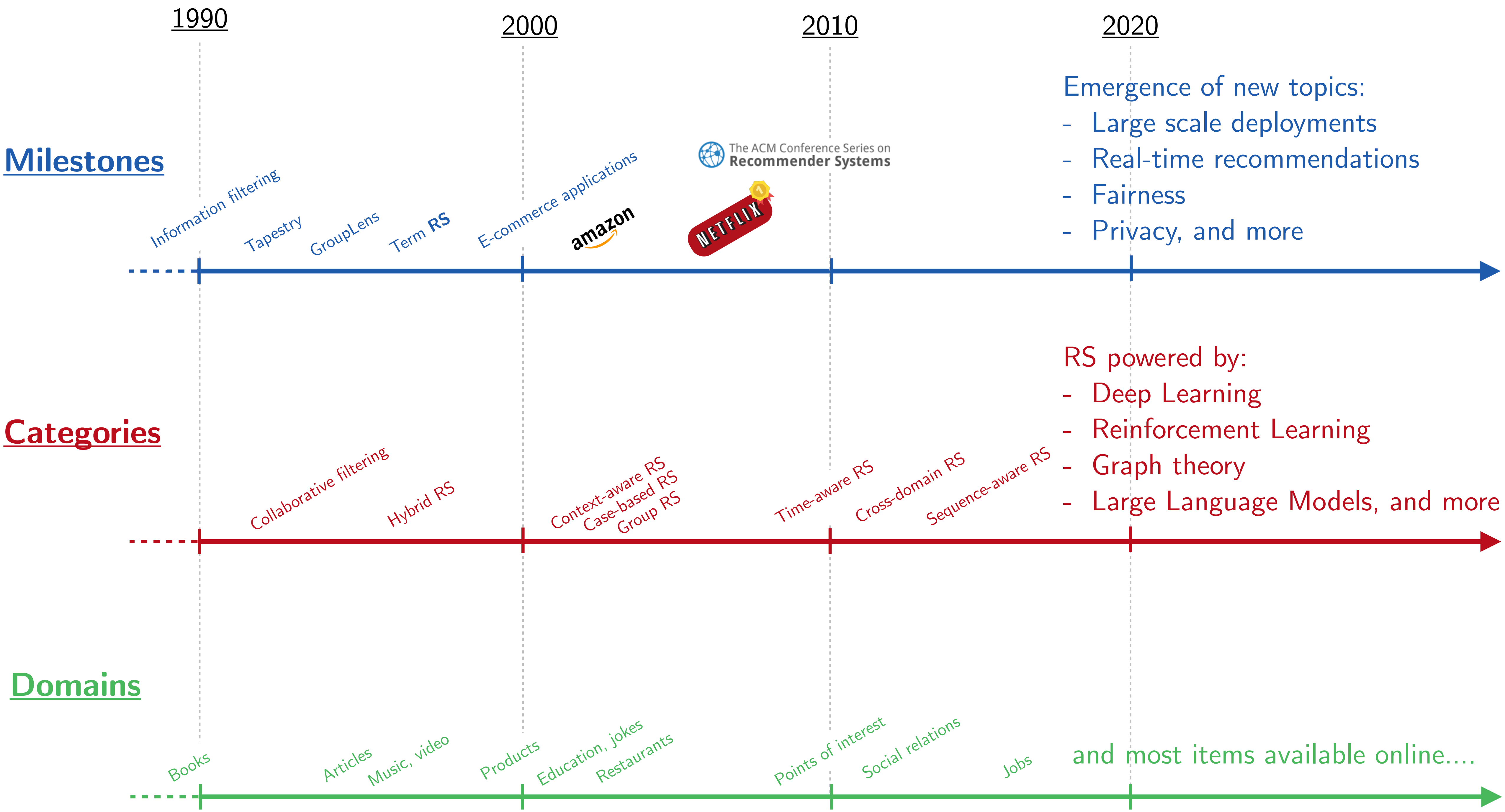


Kim Falk. 2019. *Practical Recommender Systems.* Manning, 2019.

Historical perspective on recommender systems



Historical perspective on recommender systems



The recommendation problem

The Netflix Prize (2006)

- **End-goal:** Recommend movies to Netflix users
- Competition problem:
  - Predict movie ratings on a 5-star scale
  - Dataset of 100M ratings
    from 480k users
    on 18k movies
  - $1M prize for 10% improvement over accuracy of existing solution

Ratings

James Bennett and Stan Lanning. *The Netflix Prize*. In Proc. KDD Cup and Workshop, 2007.

The Netflix Prize (2006)

- **End-goal:** Recommend movies to Netflix users

- Competition problem:
  - Predict movie ratings on a 5-star scale
  - Dataset of 100M ratings
    from 480k users
    on 18k movies
  - \$1M prize for 10% improvement over accuracy of existing solution

Ratings

→ Recommendation problem formulated as **rating prediction**

James Bennett and Stan Lanning. *The Netflix Prize*. In Proc. KDD Cup and Workshop, 2007.

The recommendation problem

Rating prediction

- The task of the RS is to predict the rating that a user will give to an item he did not rate.
  - Objective: Minimize the error between predicted and observed ratings
  - Recommendation: Items having the highest ratings for a user

The recommendation problem

Rating prediction

- The task of the RS is to predict the rating that a user will give to an item he did not rate.
  - Objective: Minimize the error between predicted and observed ratings
  - Recommendation: Items having the highest ratings for a user

Limitation: In real-world RS, it is more valuable to provide relevant top recommendations than accurately predicting ratings for all items, including those that are at the bottom list.

The recommendation problem

Rating prediction

- The task of the RS is to predict the rating that a user will give to an item he did not rate.
  - Objective: Minimize the error between predicted and observed ratings
  - Recommendation: Items having the highest ratings for a user

Limitation: In real-world RS, it is more valuable to provide relevant top recommendations than accurately predicting ratings for all items, including those that are at the bottom list.

Top-N recommendation

- The task of the RS is to predict the relevance of an item for a user, *i.e.*, predict whether the user will choose an item or not.

General formulation

- Let $\mathcal{U}$ be the set of users and $\mathcal{I}$ the set of items.
- Let f be a utility function that measures the usefulness, e.g., rating or relevance, of an item for a user: $f : \mathcal{U} \times \mathcal{I} \rightarrow \mathcal{R}$, where $\mathcal{R}$ is a totally ordered set.
- For each user $u \in \mathcal{U}$, we recommend the items maximizing f :

$$\forall u \in \mathcal{U}, i^* = \arg\max_{i \in \mathcal{I}} f(u, i)$$

f can represent the rating or the relevance of an item

A recommendation algorithm

Algorithm 1 A recommendation algorithm

Input: user u, number of items to recommend N

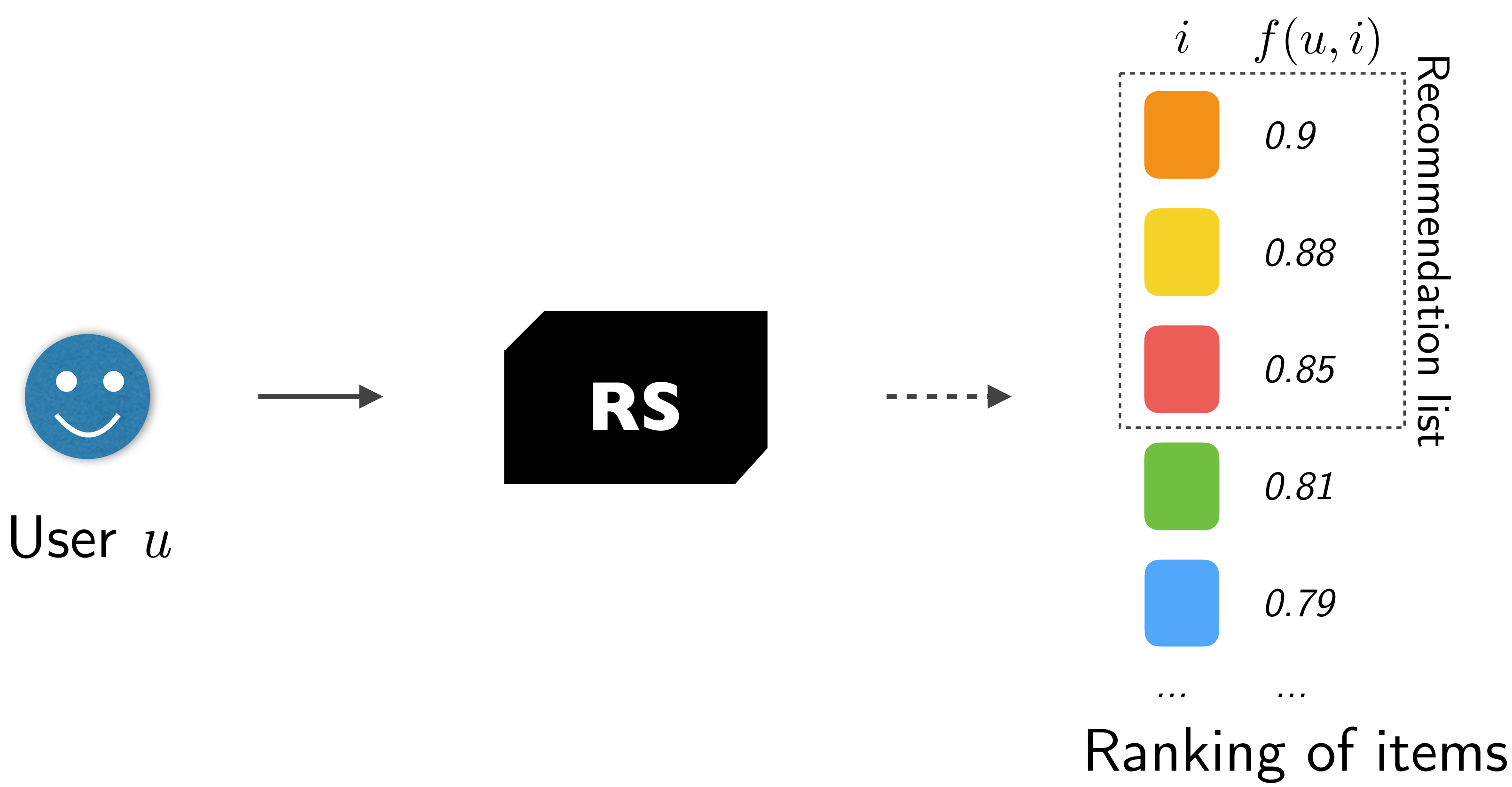
for i in $\mathcal{I} \setminus \mathcal{I}_u$ **do**

  Predict $\hat{r}_{ui}$

end for

Create list of $\mathcal{I} \setminus \mathcal{I}_u$ items ordered by decreasing order of $\hat{r}_{ui}$

Return N first items of the list

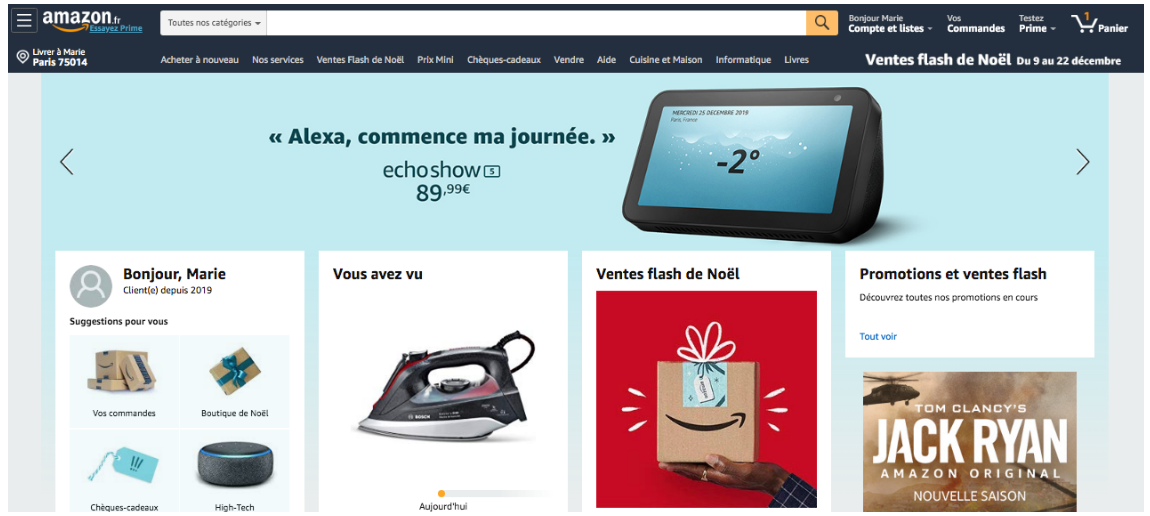


User feedback

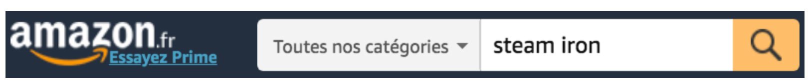
What is a recommender system?

- Knowledge of the user
  - Historical data - interactions with items
    - Purchases, clicks, views, ratings,...
  - Demographics
    - Age, gender, nationality, political views, income,...
  - Context
    - Device, location, time of day/week/year, alone or in a group,...
- Knowledge of the item
  - Content information
    - Features, tags, text, images, multimedia signals,...

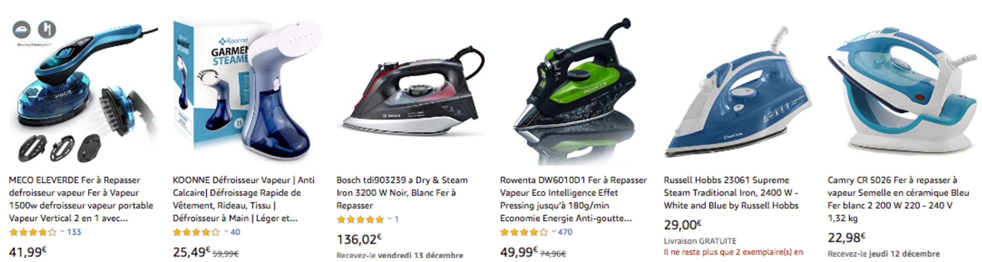
Browsing on Amazon...



Click on card #2



Search with keywords "w_1 w_2"



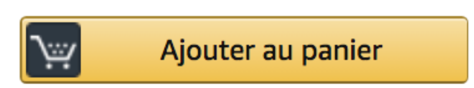
Impression with items #1, 2, 3, 4, 5, 6

Click on item #3

Informations sur le produit

Marque	Bosch
Numéro de modèle	TDI903239A
Couleur	Noir, Blanc
Poids de l'article	2,12 Kg
Dimensions du colis	33,8 x 20,2 x 12,4 cm
Capacité	0.4 litres
Puissance	3200 Watts
Matériau	Aluminium

Read description of item #3



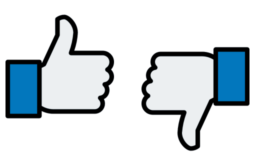
Add item #3 to cart

Rate and review item #3

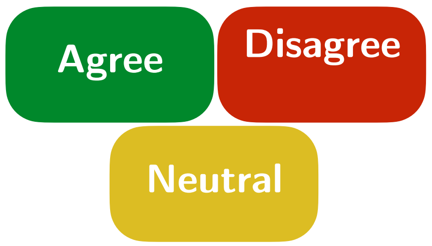
Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.

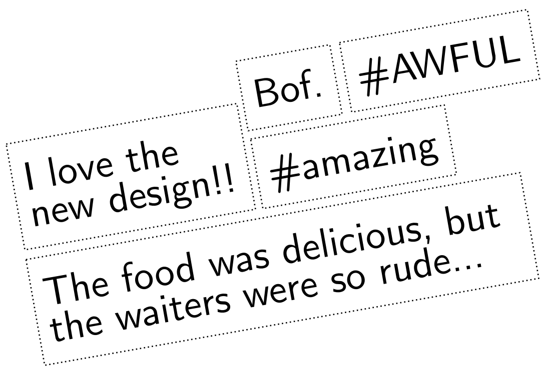
Ratings, 5-star scale



Like/Dislike



Ordinal scale



Textual reviews

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.
- Limitations:
  - Users do not always provide explicit feedback.

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.
- Limitations:
  - Users do not always provide explicit feedback.
  - The system may be unable to collect explicit feedback, e.g., cannot collect rating of people for Linkedin's "*People You May Know*".

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.
- Limitations:
  - Users do not always provide explicit feedback.
  - The system may be unable to collect explicit feedback, e.g., cannot collect rating of people for Linkedin's "*People You May Know*".
  - Is the rating revealing the true opinion of the user?

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.
- Limitations:
  - Users do not always provide explicit feedback.
  - The system may be unable to collect explicit feedback, e.g., cannot collect rating of people for Linkedin's "*People You May Know*".
  - Is the rating revealing the true opinion of the user?
  - Is the user being influenced by his social circle when rating an item?

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.
- Limitations:
  - Users do not always provide explicit feedback.
  - The system may be unable to collect explicit feedback, e.g., cannot collect rating of people for Linkedin's "*People You May Know*".
  - Is the rating revealing the true opinion of the user?
  - Is the user being influenced by his social circle when rating an item?
  - Is the user rating the complete package or is there an item's detail affecting the overall rating?

Explicit feedback

- Explicit feedback is a form of feedback directly reported by the user to the system.
- Limitations:
  - Users do not always provide explicit feedback.
  - The system may be unable to collect explicit feedback, e.g., cannot collect rating of people for Linkedin's "*People You May Know*".
  - Is the rating revealing the true opinion of the user?
  - Is the user being influenced by his social circle when rating an item?
  - Is the user rating the complete package or is there an item's detail affecting the overall rating?
  - Explicit feedback exhibit some biases, e.g., some users tend to give higher ratings than others for a similar level of appreciation.

Implicit feedback

- Implicit feedback is collected by the system without the intervention of the user.

Clicks


Informations sur le produit

Marque	Bosch
Numéro de modèle	TDI903239A
Couleur	Noir, Blanc
Poids de l'article	2,12 Kg
Dimensions du colis	33,8 x 20,2 x 12,4 cm
Capacité	0.4 litres
Puissance	3200 Watts
Matériau	Aluminium

Views of item's description

Purchases

Add to Wishlist

Implicit feedback

- Implicit feedback is collected by the system without the intervention of the user.

Clicks

Informations sur le produit

Marque	Bosch
Numéro de modèle	TDI903239A
Couleur	Noir, Blanc
Poids de l'article	2,12 Kg
Dimensions du colis	33,8 x 20,2 x 12,4 cm
Capacité	0.4 litres
Puissance	3200 Watts
Matériau	Aluminium

Views of item's description

Acheter cet article

Purchases

Ajouter à votre liste

Add to Wishlist

$\longrightarrow$ More abundant than explicit feedback

Implicit feedback

- Implicit feedback is collected by the system without the intervention of the user.

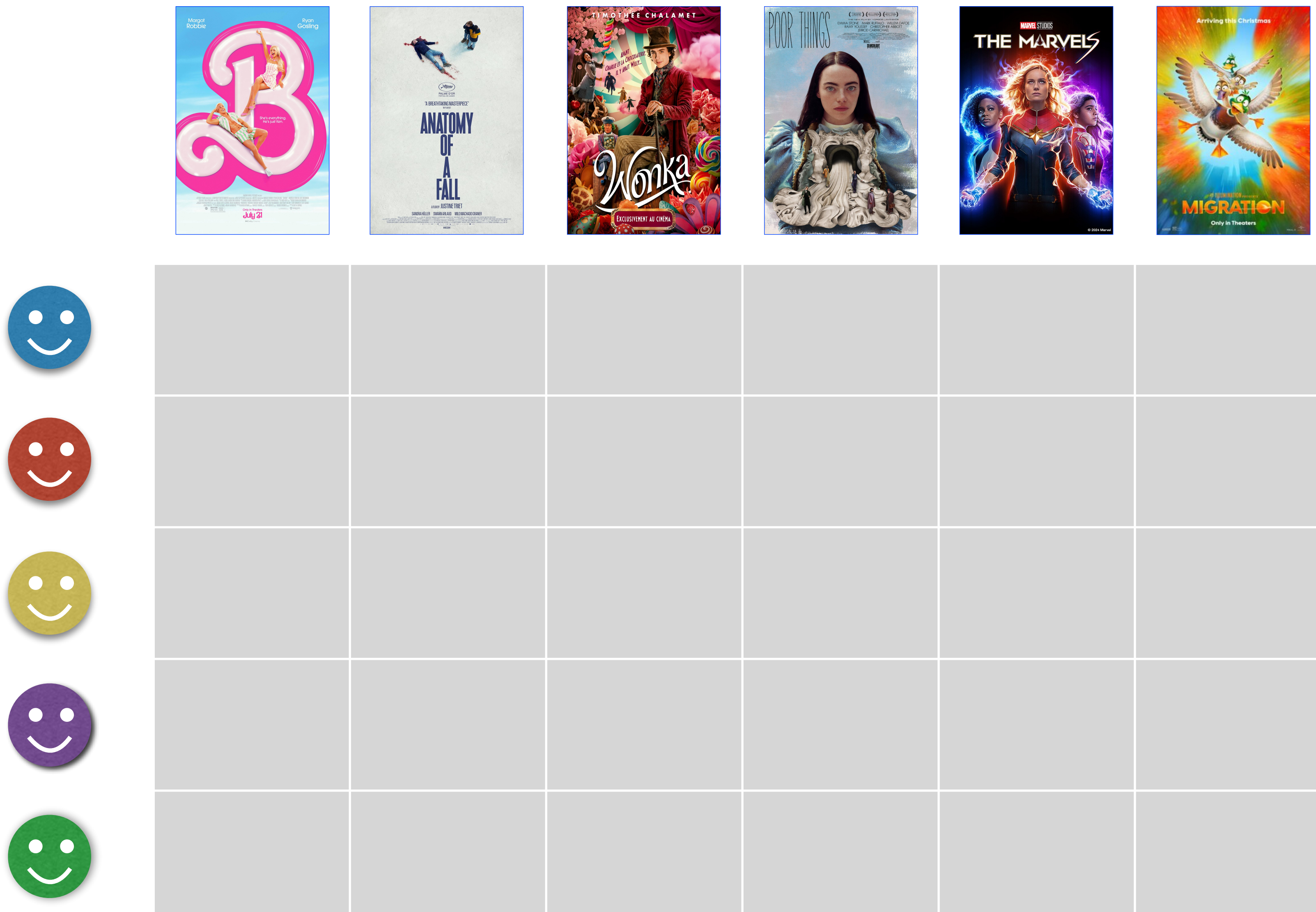
Implicit feedback

- Implicit feedback is collected by the system without the intervention of the user.
- Challenges:
  - Implicit feedback is inherently noisy.
    - Every *positive item*, i.e., item the user has interacted with, is not necessarily an item liked by the user.

Implicit feedback

- Implicit feedback is collected by the system without the intervention of the user.
- Challenges:
  - Implicit feedback is inherently noisy.
    - Every *positive item*, i.e., item the user has interacted with, is not necessarily an item liked by the user.
  - There is no negative feedback.
    - Negative items are a mixture of unliked items and unknown items.

Feedback matrix



Feedback matrix - Explicit feedback

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

Feedback matrix - Implicit feedback (Binary)

1		1	1		
			1	1	1
	1	1	1		
1	1			1	1
	1		1		

Feedback matrix - Implicit feedback (Weights)

2		2.8	0.8		
			3.2	2.4	1
	1.2	1	2		
0.4	0.9			2.1	1.3
	3		0		

Challenges

Rating data sparsity

- Users only interact with a small number items selected from a very large catalog of available items.

Rating data sparsity

- Users only interact with a small number items selected from a very large catalog of available items.

Examples of datasets:

NETFLIX

480k users
18k movies
100M ratings
Period of ~7 years

Density: **1.1%**

James Bennett and Stan Lanning. *The Netflix Prize.* In Proc. KDD Cup and Workshop, 2007.

ACCOR HOTELS Feel Welcome

7.8M users
4.5k hotels
34.7M bookings
Period of ~4 years

Density: **0.1%**

Marie Al-Ghossein et al. *Exploiting Contextual and External Data for Hotel Recommendation.* In Proc. UMAP, 2018.

amazon

20.98M users
9.35M items
82.83M ratings
Period of ~18 years

Density: **0.00004%**

https://cseweb.ucsd.edu/~jmcauley/datasets.html#amazon_reviews

Cold-start

- The *cold-start* problem designates the problems of:
  - Recommending items to new users where the historical data is missing,
    - Recommendations cannot be computed before the user has rated a certain number of items, but the user expects to receive suggestions before that.
  - Recommending new items to existing users.

Cold-start

- The *cold-start* problem designates the problems of:
  - Recommending items to new users where the historical data is missing,
    - Recommendations cannot be computed before the user has rated a certain number of items, but the user expects to receive suggestions before that.
  - Recommending new items to existing users.
- A related problem is the *gray sheep* problem:
  - Recommending items to users who have very unique tastes, such that no other user (or very few users) has interacted with the same items.

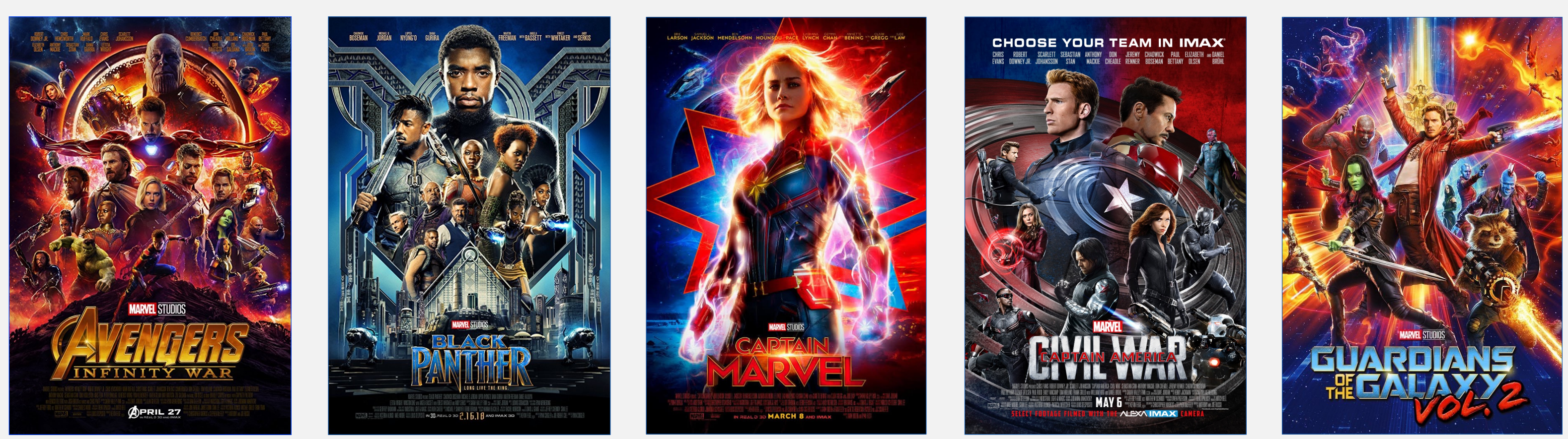
Overspecialization

- Recommendation algorithms tend to recommend items that are too similar to what the user has already experienced, resulting in overspecialized recommendations.

Overspecialization

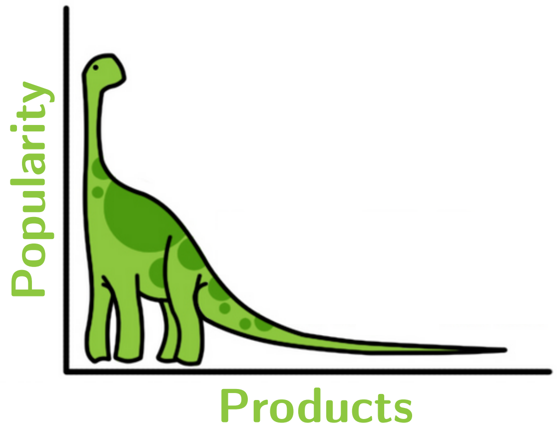
- Recommendation algorithms tend to recommend items that are too similar to what the user has already experienced, resulting in overspecialized recommendations.

Is this recommendation list useful?



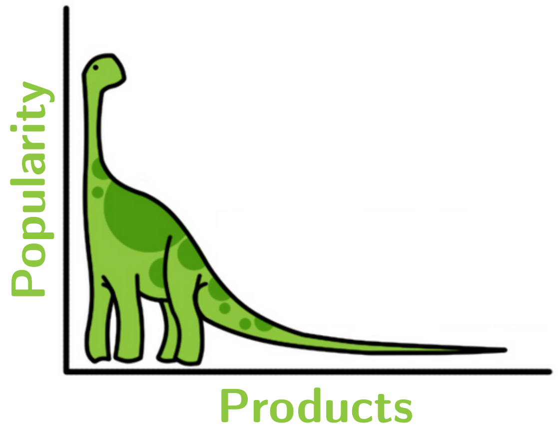
Popularity bias

- Most item catalogs exhibit the *long tail* effect,



Popularity bias

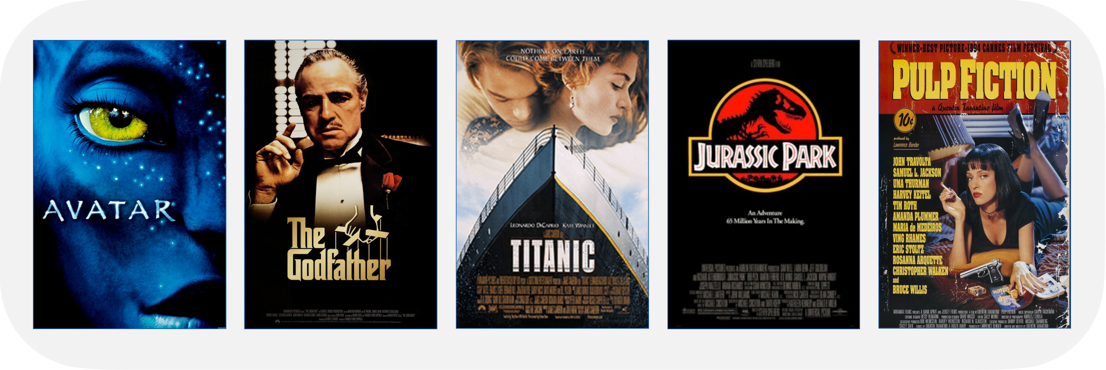
- Most item catalogs exhibit the *long tail* effect,



which raises two issues:

- Risk of recommending popular items to everyone,

Is this recommendation list useful?



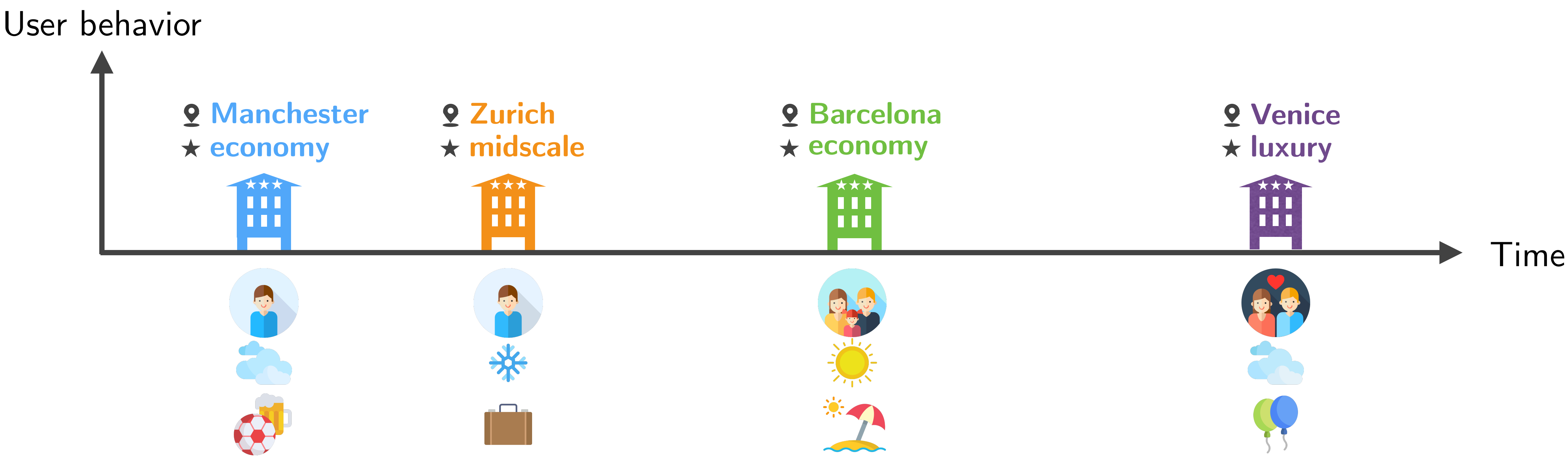
- Difficulty to promote long tail items with little feedback.

Temporal dynamics

- Users preferences and item perceptions are changing over time
  - Reasons: emergence of new items, seasonality factors, etc.

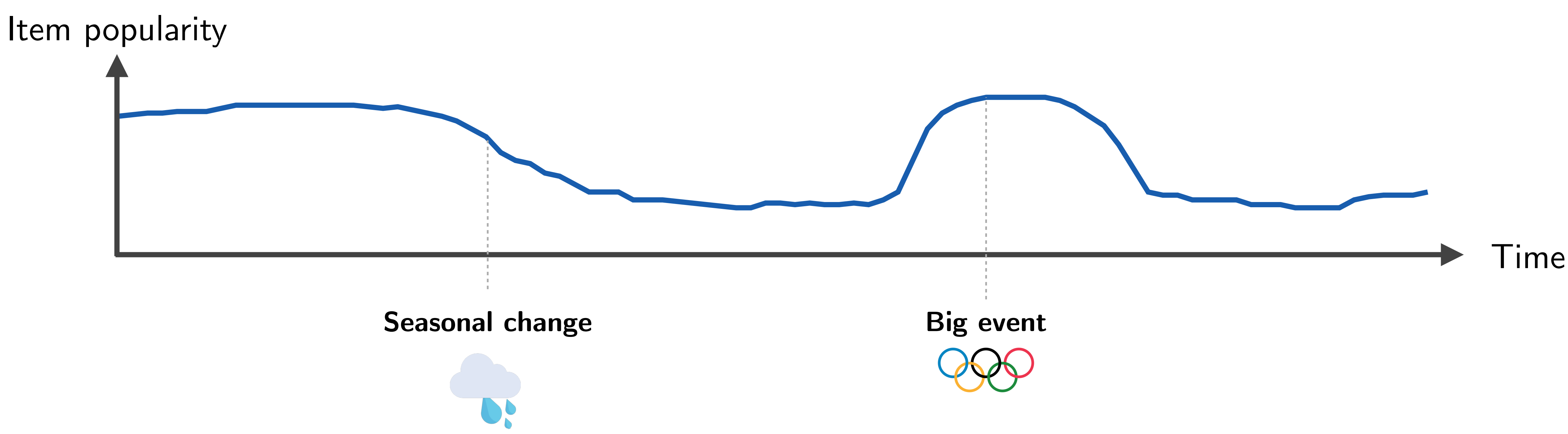
Temporal dynamics

- Users preferences and item perceptions are changing over time
  - Reasons: emergence of new items, seasonality factors, etc.



Temporal dynamics

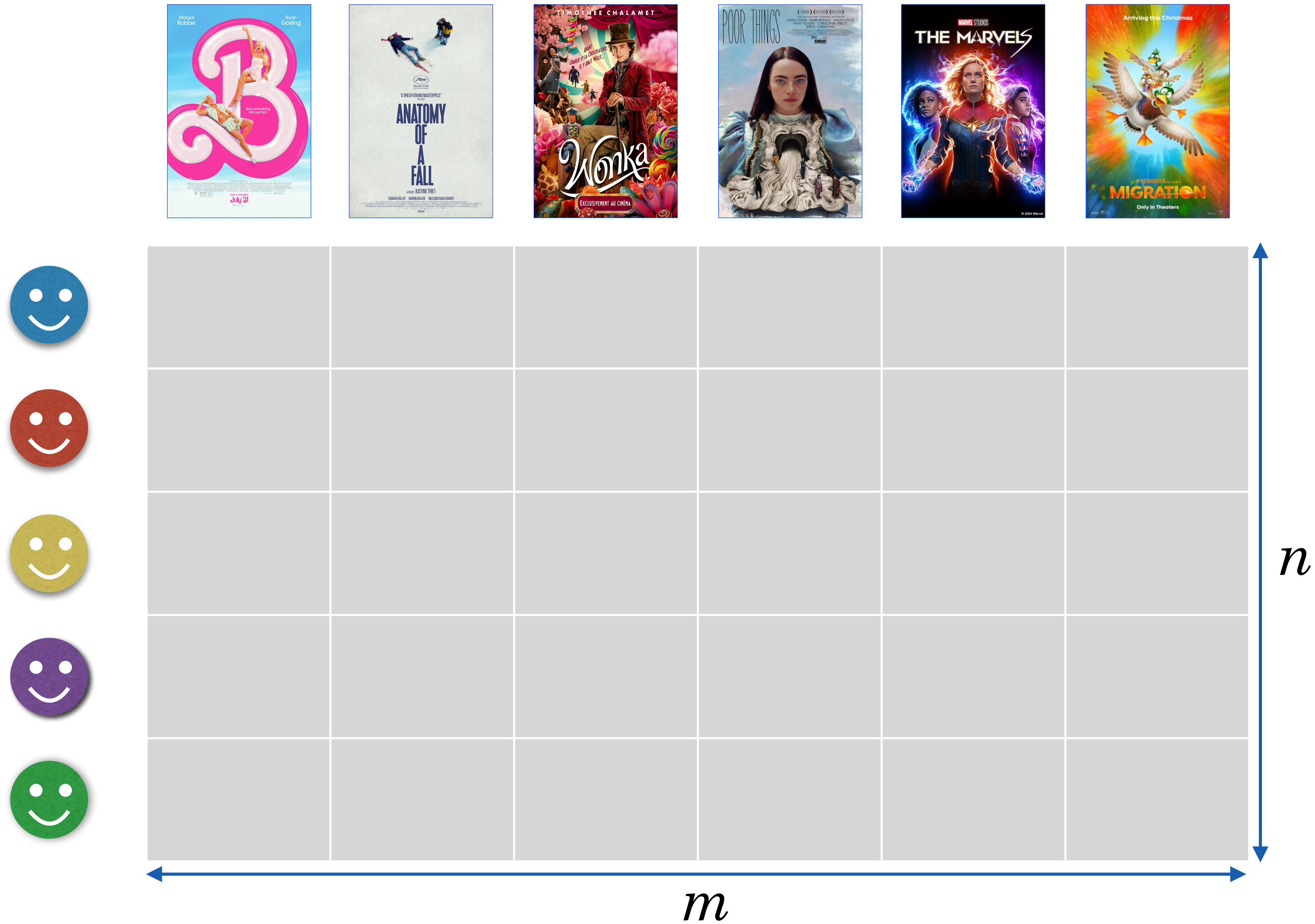
- Users preferences and item perceptions are changing over time
  - Reasons: emergence of new items, seasonality factors, etc.



Feedback matrix - Notations

Set of items: $\mathcal{I}$, set of items rated by user u: $\mathcal{I}_u$

Set of users: $\mathcal{U}$



Feedback matrix: R

Rating: r_{ui}, predicted rating: $\hat{r}_{ui}$

Evaluation of RS

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior
- User study
  - Recruit a small number of users and ask them to interact with the RS

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior
- User study
  - Recruit a small number of users and ask them to interact with the RS
  - Pros: Provides more insights than offline evaluation,

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior
- User study
  - Recruit a small number of users and ask them to interact with the RS
  - Pros: Provides more insights than offline evaluation,
  - Cons: May introduce biases in the collected data

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior
- User study
  - Recruit a small number of users and ask them to interact with the RS
  - Pros: Provides more insights than offline evaluation,
  - Cons: May introduce biases in the collected data
- Online evaluation
  - A/B testing on a real RS deployed online

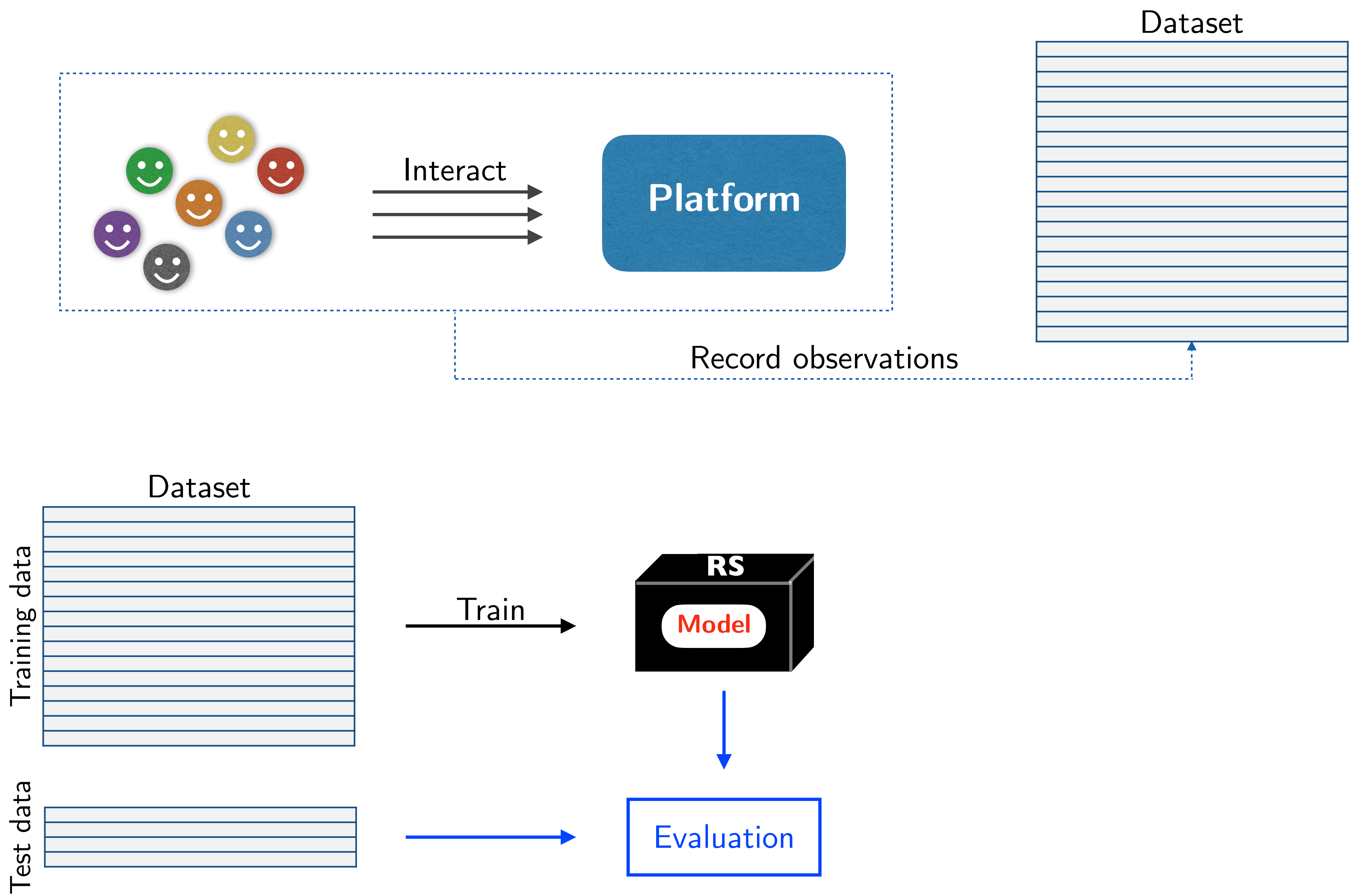
Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior
- User study
  - Recruit a small number of users and ask them to interact with the RS
  - Pros: Provides more insights than offline evaluation,
  - Cons: May introduce biases in the collected data
- Online evaluation
  - A/B testing on a real RS deployed online
  - Pros: Provides the strongest evidence,

Evaluation methods

- Offline evaluation
  - Performed on previously collected datasets and simulate the user behavior
  - Pros: Does not require an interaction with real users, allows replicability,
  - Cons: Cannot measure the influence of the RS on the user behavior
- User study
  - Recruit a small number of users and ask them to interact with the RS
  - Pros: Provides more insights than offline evaluation,
  - Cons: May introduce biases in the collected data
- Online evaluation
  - A/B testing on a real RS deployed online
  - Pros: Provides the strongest evidence,
  - Cons: Risk of negatively affecting the real users's experience

Offline evaluation



Offline evaluation

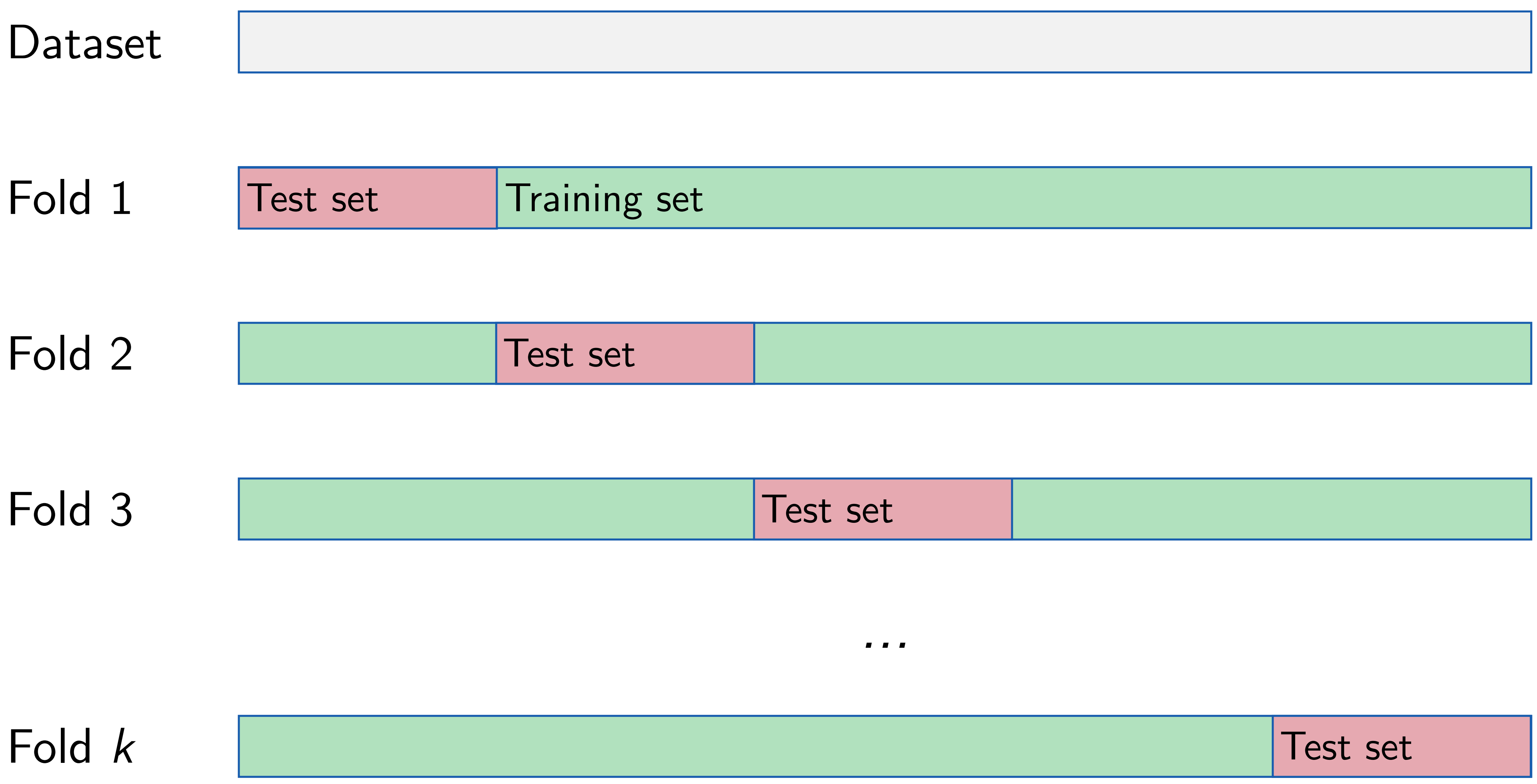
- Random split

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

Training set   Test set

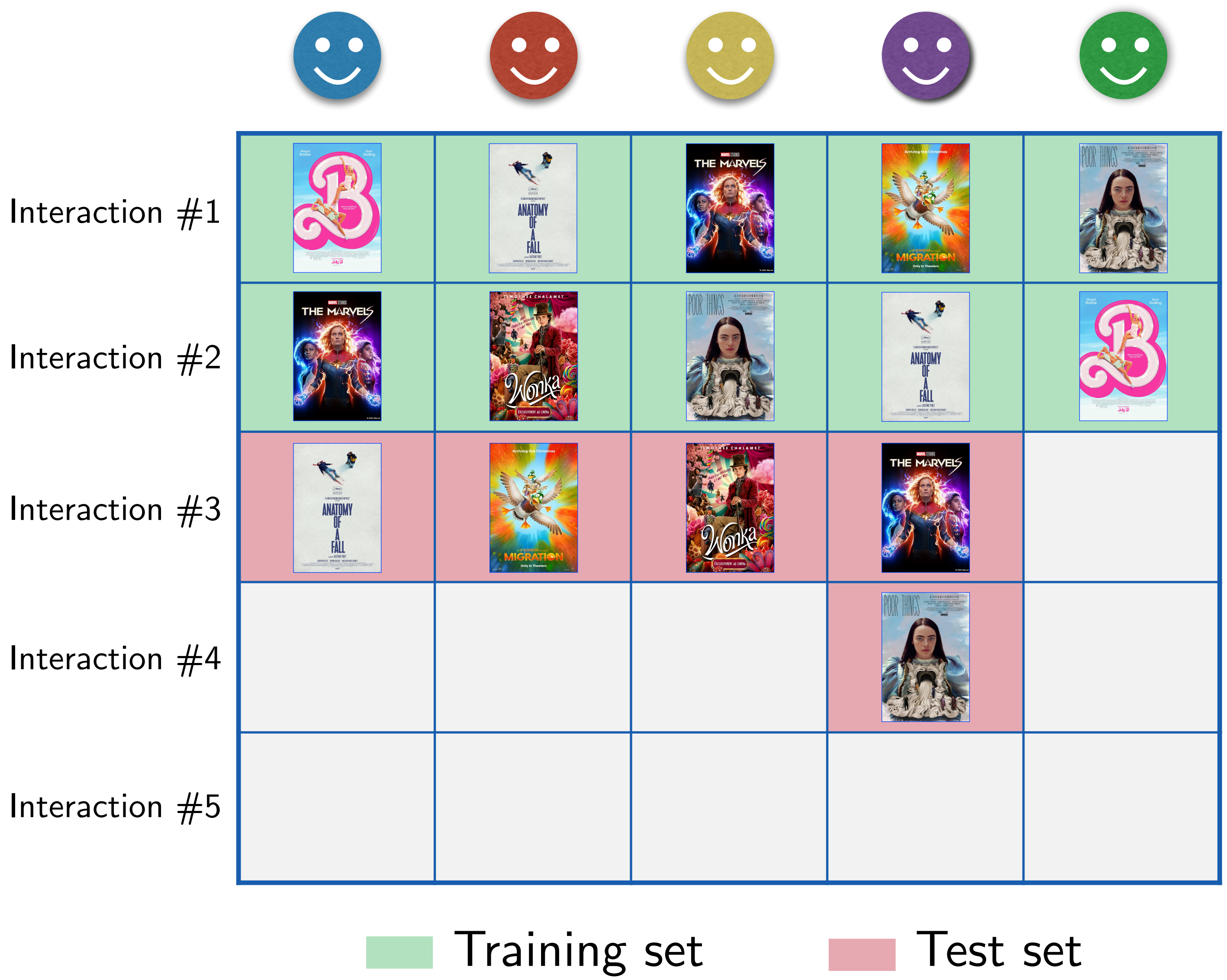
Offline evaluation

- Random split, k-fold cross validation



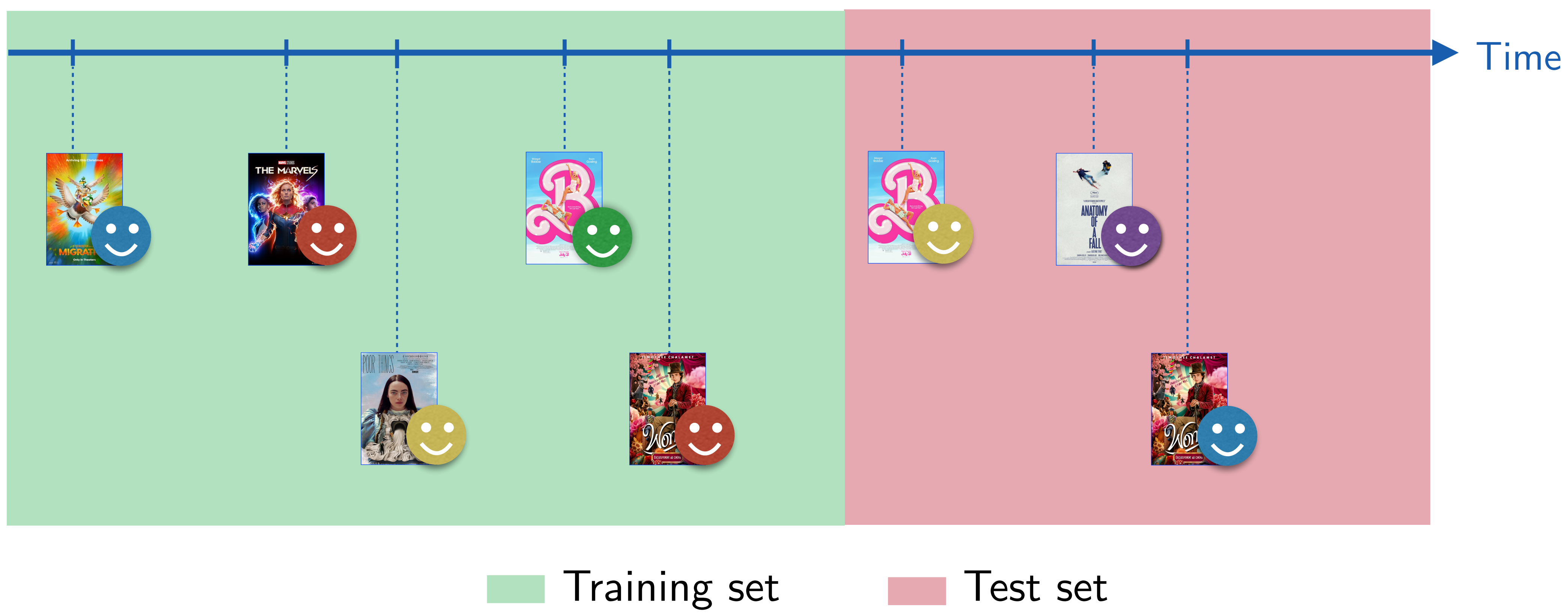
Offline evaluation

- Given-n split



Offline evaluation

- Chronological split



Evaluation criteria

- Accuracy
  - How accurate are the predictions of utility of items?

Evaluation criteria

- Accuracy
  - How accurate are the predictions of utility of items?
- Coverage
  - How many items from the catalog are the recommendations covering?

Evaluation criteria

- Accuracy
  - How accurate are the predictions of utility of items?
- Coverage
  - How many items from the catalog are the recommendations covering?
- Diversity
  - How diverse are the computed recommendations?

Evaluation criteria

- Accuracy
  - How accurate are the predictions of utility of items?
- Coverage
  - How many items from the catalog are the recommendations covering?
- Diversity
  - How diverse are the computed recommendations?
- Serendipity
  - How surprising are the relevant computed recommendations?

Evaluation criteria

- Accuracy
  - How accurate are the predictions of utility of items?
- Coverage
  - How many items from the catalog are the recommendations covering?
- Diversity
  - How diverse are the computed recommendations?
- Serendipity
  - How surprising are the relevant computed recommendations?
- Explainability
  - Are we able to explain and justify the recommendations?

Evaluation metrics for accuracy

Measuring the accuracy of the predicted ratings.

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{|\mathcal{T}|} \sum_{(u,i) \in \mathcal{T}} |r_{ui} - \hat{r}_{ui}|$$

Evaluation metrics for accuracy

Measuring the accuracy of the predicted ratings.

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{|\mathcal{T}|} \sum_{(u,i) \in \mathcal{T}} |r_{ui} - \hat{r}_{ui}|$$

User	Actual rating	Predicted rating	Error
	★★★★★	★★★★☆	★
	★☆☆☆☆	★★★★☆	★★★
	★★★☆☆	★★★★★	★★
	★★★★☆	★★★★☆	
	★★☆☆☆	★☆☆☆☆	★

$$MAE = 1.4$$

Evaluation metrics for accuracy

Measuring the accuracy of the predicted ratings.

- Root Mean Squarred Error (RMSE)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{|\mathcal{T}|} \sum_{(u,i) \in \mathcal{T}} (r_{ui} - \hat{r}_{ui})^2}$$

Evaluation metrics for accuracy

Measuring the accuracy of the predicted ratings.

- Root Mean Squarred Error (RMSE)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{|\mathcal{T}|} \sum_{(u,i) \in \mathcal{T}} (r_{ui} - \hat{r}_{ui})^2}$$

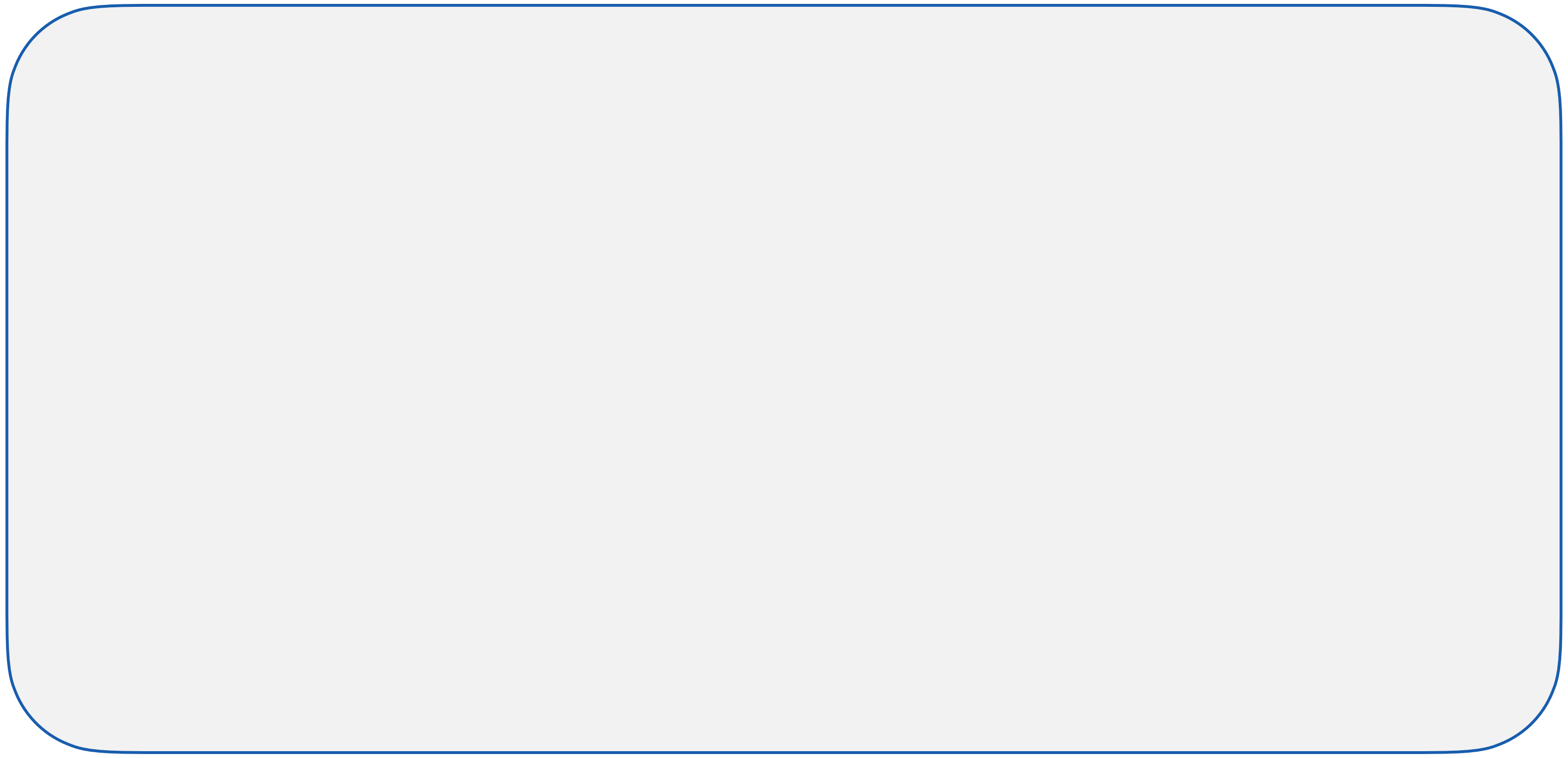
$$RMSE = 3$$

Evaluation metrics for accuracy

Measuring the accuracy of top-N recommendations

Evaluation metrics for accuracy

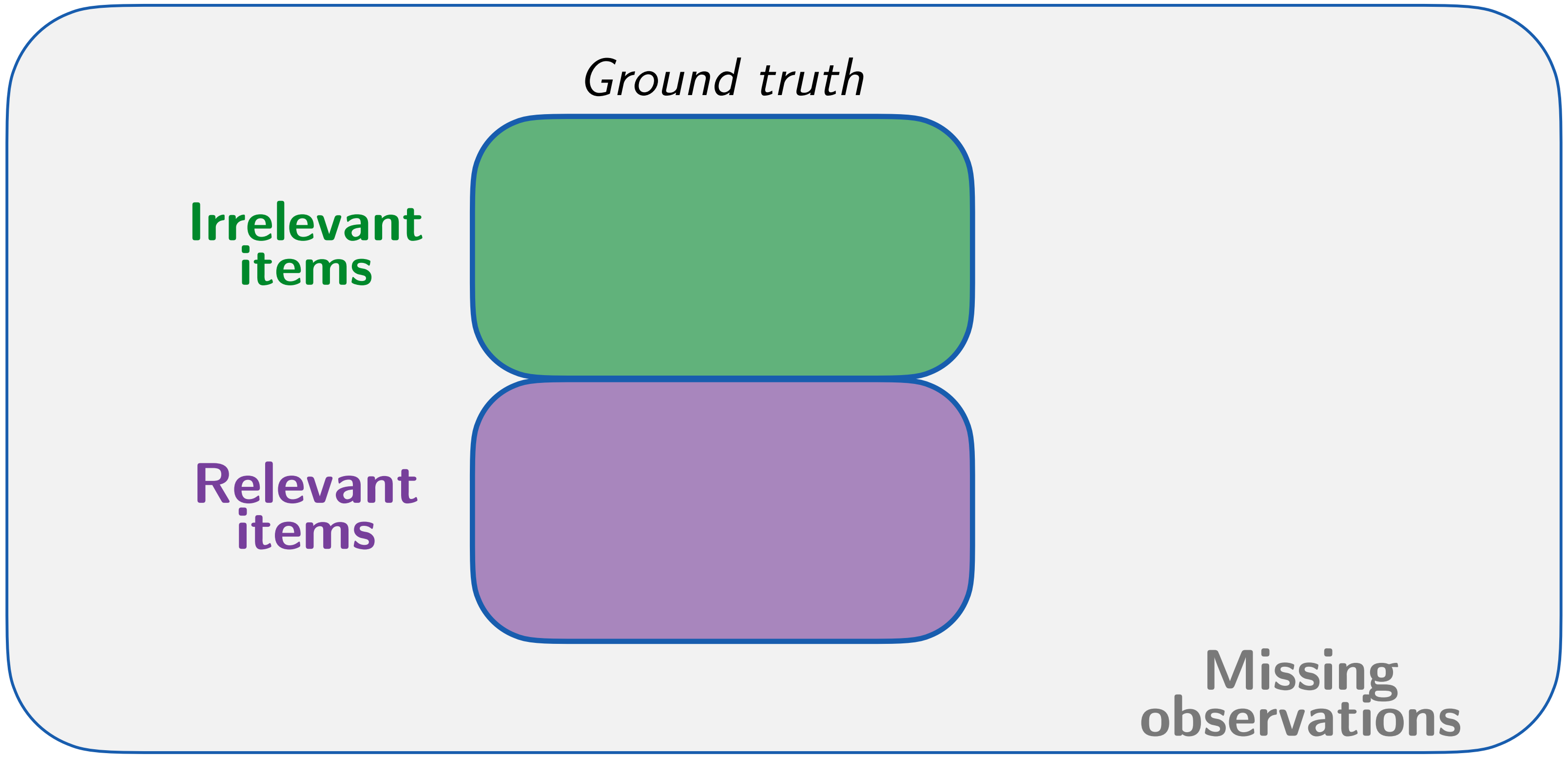
Measuring the accuracy of top-N recommendations



Complete dataset

Evaluation metrics for accuracy

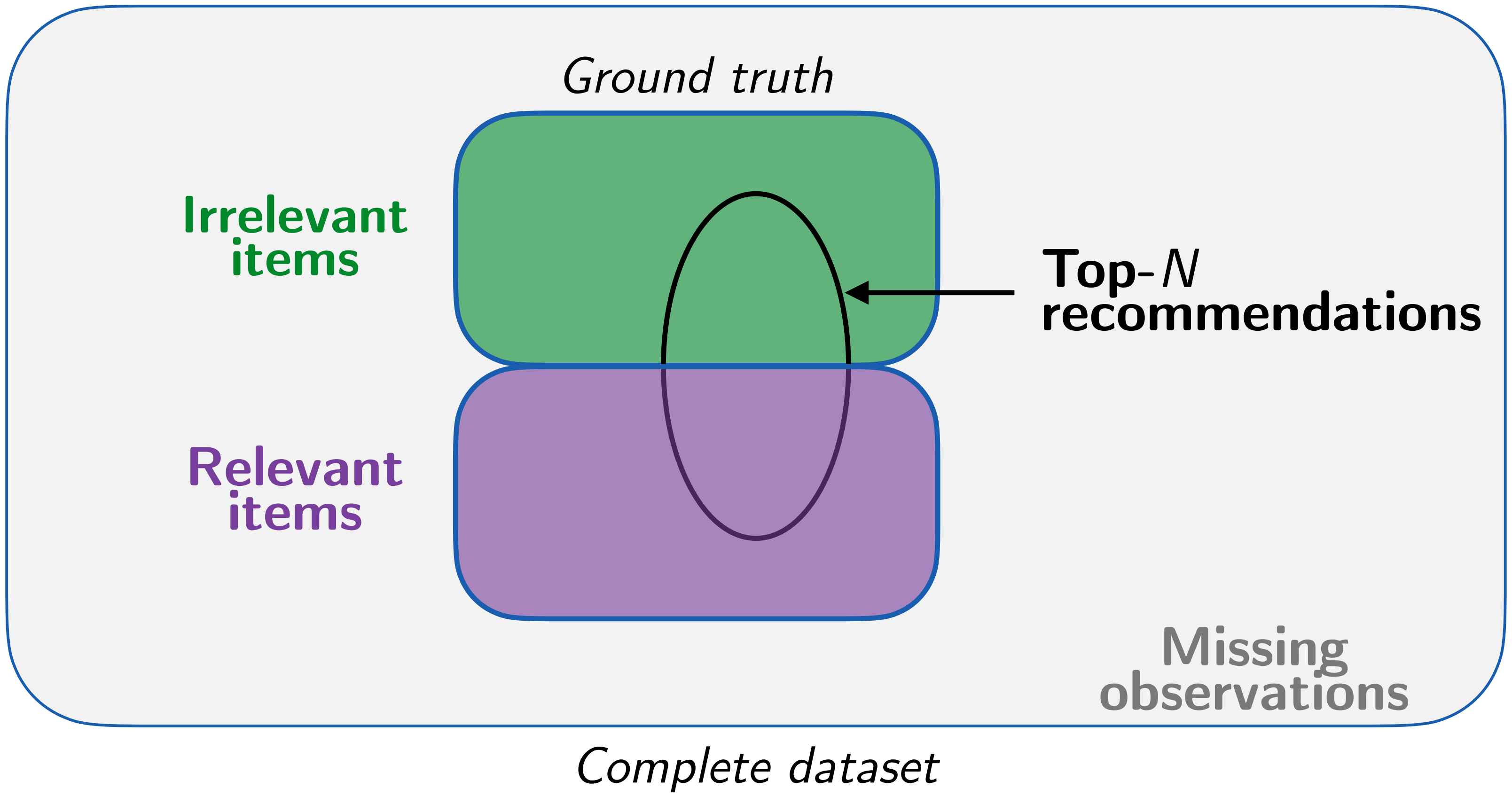
Measuring the accuracy of top-N recommendations



Complete dataset

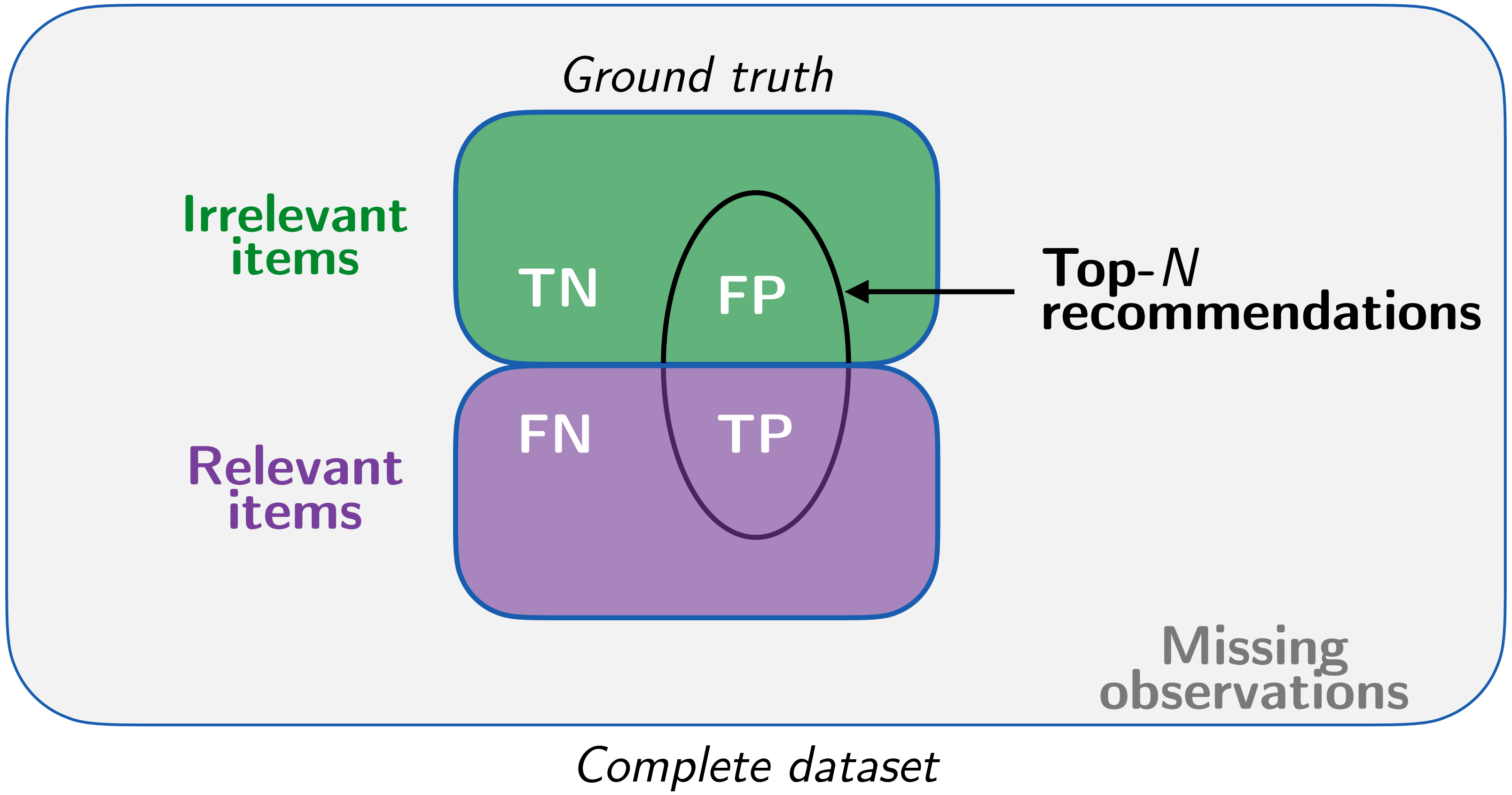
Evaluation metrics for accuracy

Measuring the accuracy of top-N recommendations



Evaluation metrics for accuracy

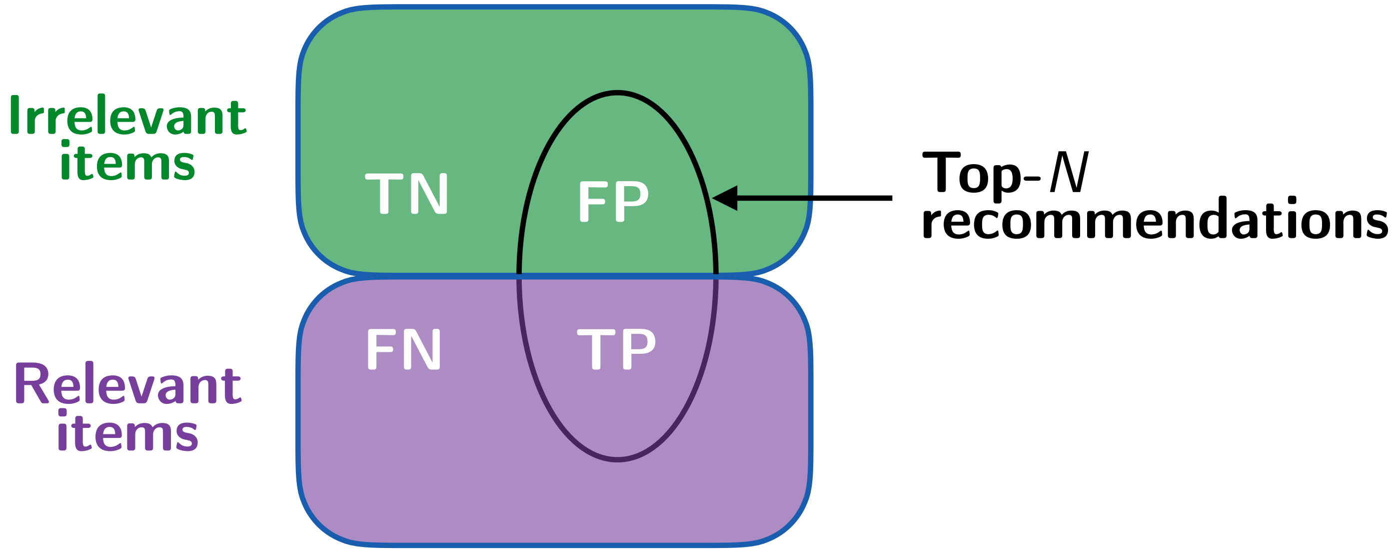
Measuring the accuracy of top-N recommendations



Evaluation metrics for accuracy

Measuring the accuracy of top-N recommendations

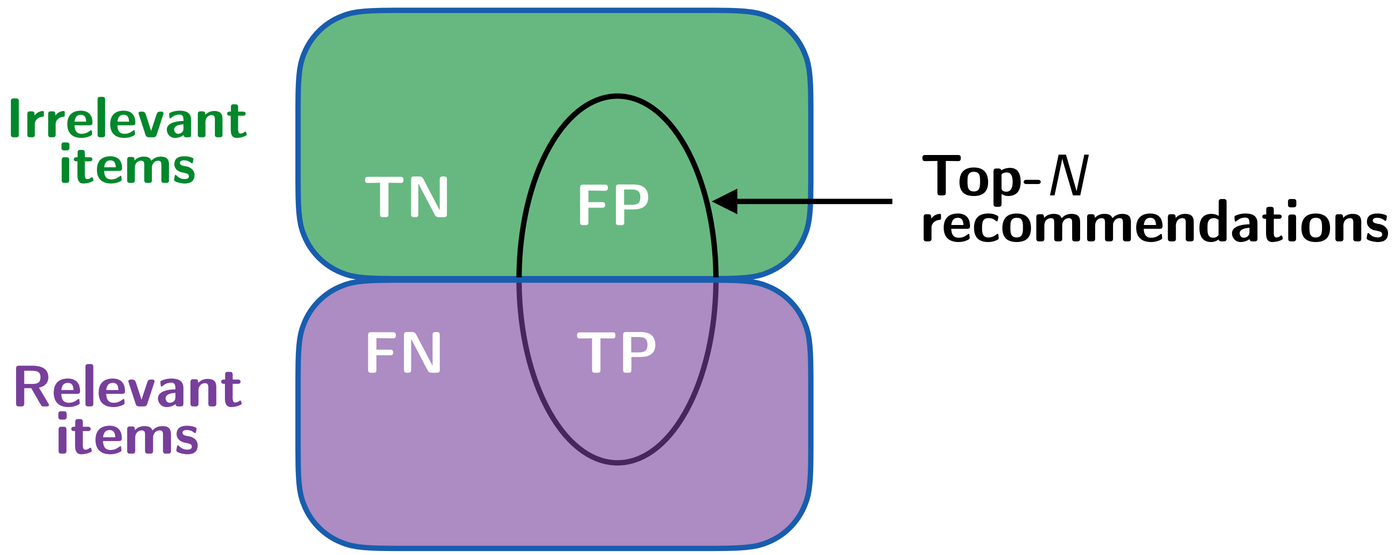
$$Recall_u@N = \frac{TP}{FN + TP} = \frac{\#relevant\ recommended\ items}{\#relevant\ items}$$



Evaluation metrics for accuracy

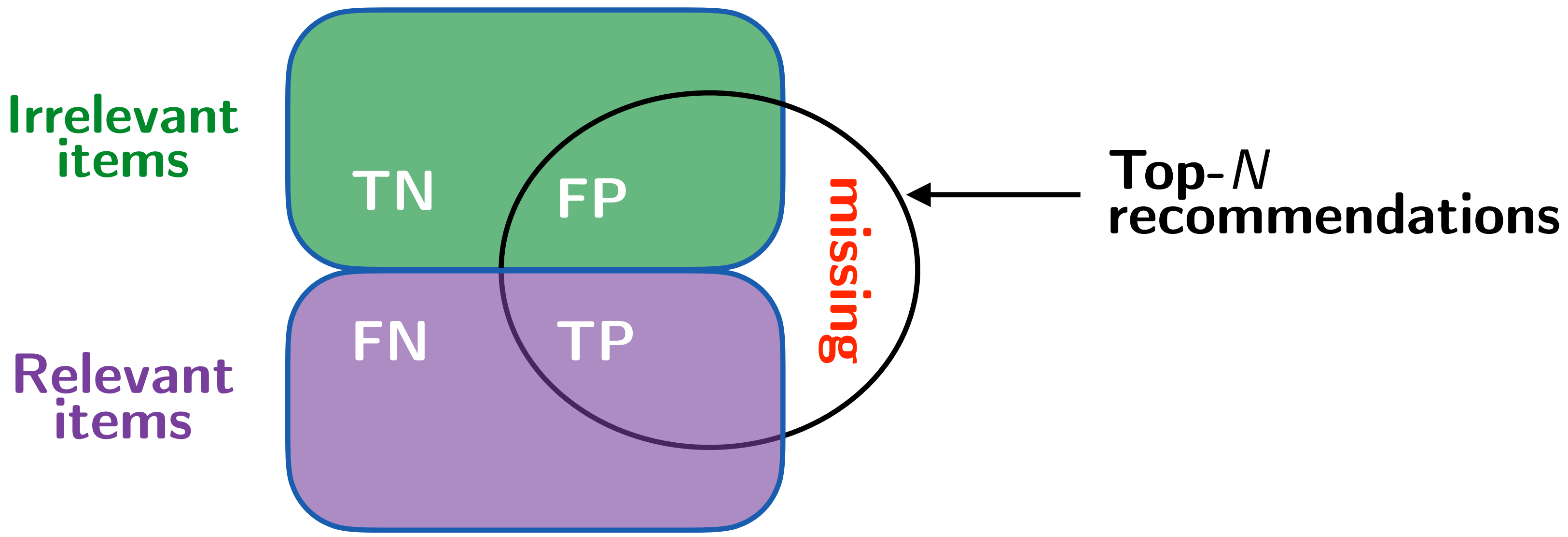
Measuring the accuracy of top-N recommendations

$$Precision_u@N = \frac{TP}{FP+TP} = \frac{\#relevant\ recommended\ items}{\#recommended\ items}$$



Evaluation metrics for accuracy

- Recall: *Missing Not As Random* assumption
  - Non-relevant items are more likely to be missing than relevant items
- Precision: *All Missing As Negative* assumption
  - All missing ratings are irrelevant



Evaluation metrics for ranking

Measuring the quality of ranking of top-N recommendations

- Discounted Cumulative Gain (DCG):

$$DCG_u@N = \sum_{i=1}^{N} \frac{rel_{ui}}{log_2(i+1)}$$

where rel_{ui} = *1* if item at rank i is relevant, *0* otherwise.

Evaluation metrics for ranking

Measuring the quality of ranking of top-N recommendations

- Discounted Cumulative Gain (DCG):

$$DCG_u@N = \sum_{i=1}^{N} \frac{rel_{ui}}{log_2(i+1)}$$

where $rel_{ui} = 1$ if item at rank i is relevant, 0 otherwise.

$$DCG_u \quad 1 + 0 + 0 + 0.43 + 0.38$$

Evaluation metrics for ranking

Measuring the quality of ranking of top-*N* recommendations

- Normalized Discounted Cumulative Gain (DCG):

$$NDCG@N = \frac{1}{|\mathcal{T}_u|} \sum_{u \in \mathcal{T}_u} \frac{DCG_u@N}{DCG_u^*@N}$$

where $DCG^*{}_u$ is the best possible DCG_u that can be obtained.

Evaluation metrics for ranking

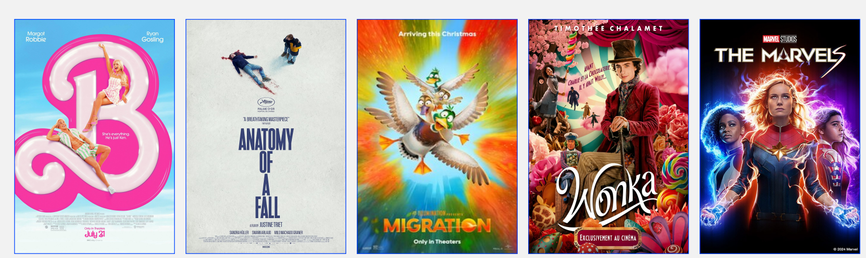
Measuring the quality of ranking of top-*N* recommendations

- Normalized Discounted Cumulative Gain (DCG):

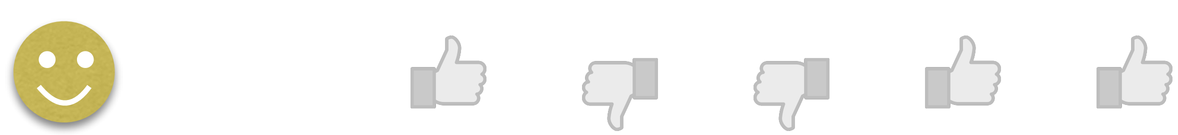
$$NDCG@N = \frac{1}{|\mathcal{T}_u|} \sum_{u \in \mathcal{T}_u} \frac{DCG_u@N}{DCG^*_u@N}$$

where $DCG^*{}_u$ is the best possible DCG_u that can be obtained.

Recommendation list



i = 1 i = 2 i = 3 i = 4 i = 5



$DCG^*{}_u$ $1 + 0.63 + 0.5 + 0.43 + 0.38$

List of other used metrics

- Mean Average Precision (MAP)
- Mean Reciprocal Rank (MRR)
- Area under the ROC Curve (AUC)
- Hit-Rank (HR)
- Average Reciprocal Hit-Rank (ARHR)

  ...

Non-personalized recommendations

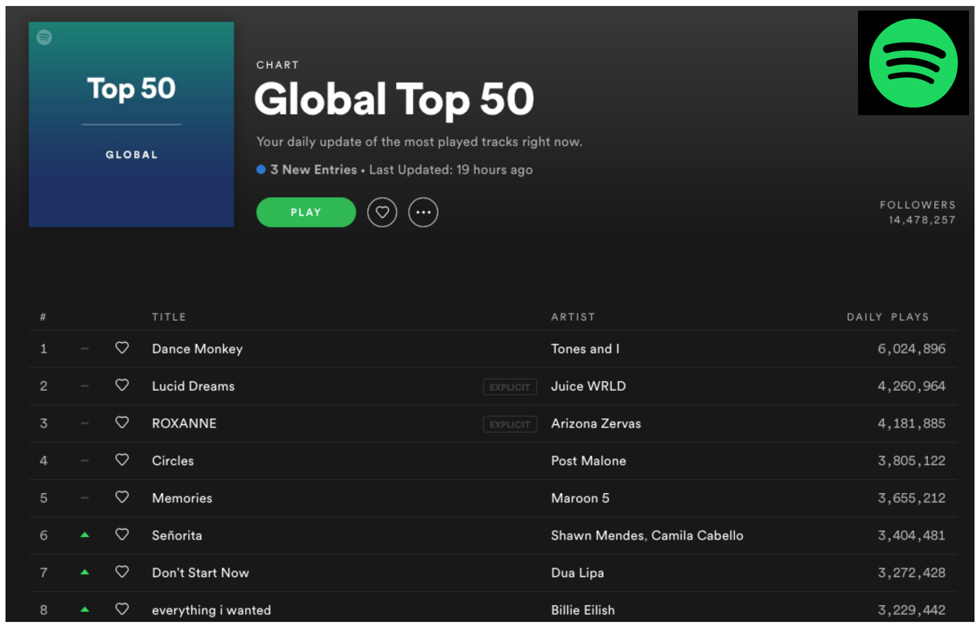
Non-personalized recommendations

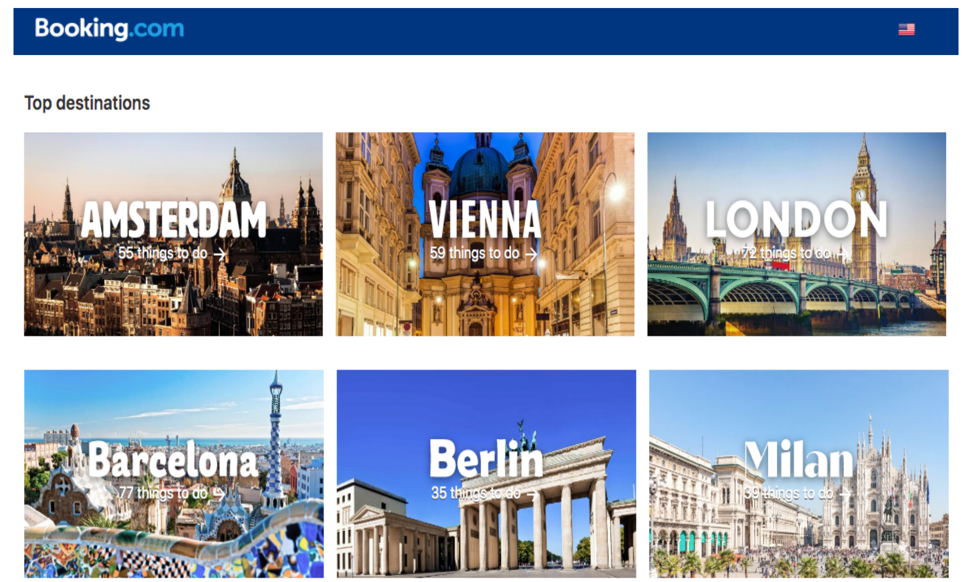
- Personalized recommendations
  - Tailor the recommendation list to a specific individual
  - Using information about the target user and similar users

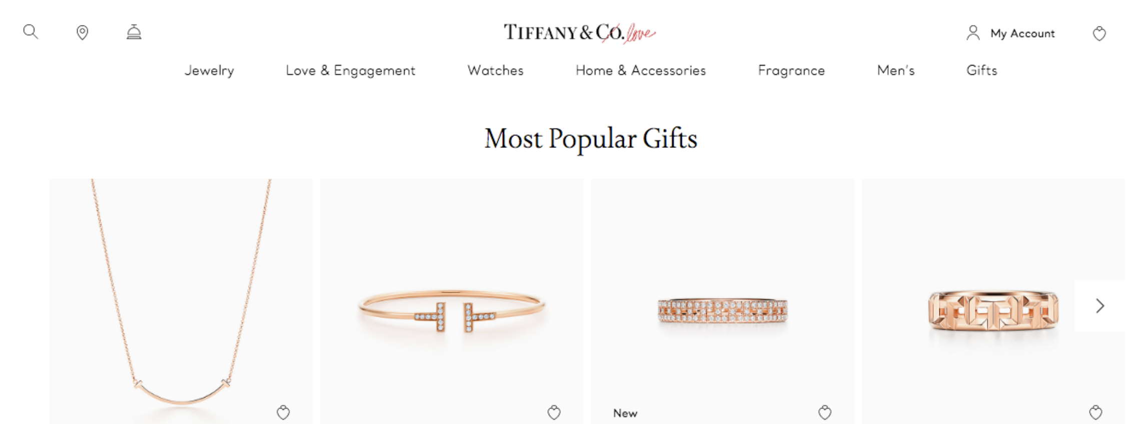
vs.

- Non-personalized recommendations
  - Do not require to know anything about the individual user
  - Using aggregated information about users

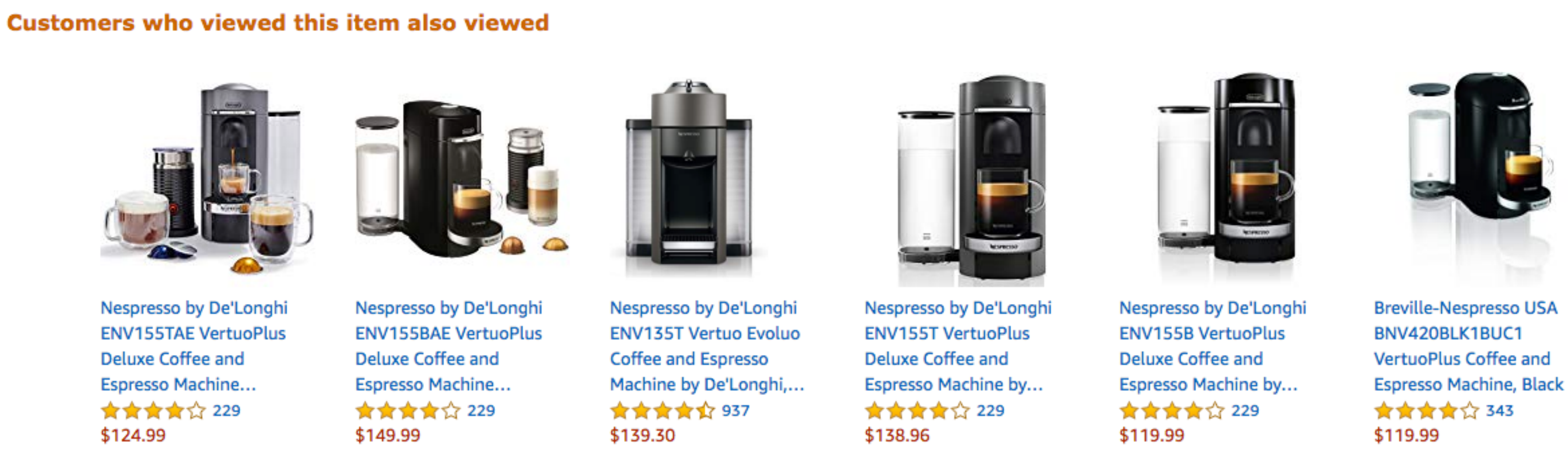
Most popular item recommendation







Association rules

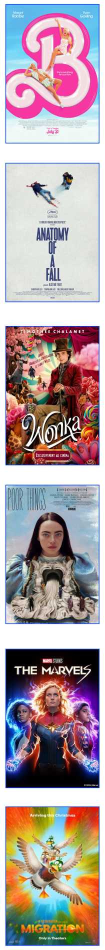


Association rules

1		1	1		
			1	1	1
	1	1	1		
1	1			1	1
	1		1		

Also watched...

Users who watched...



	1	1	1	1	1
1		1	2	1	1
1	1		2	0	0
1	2	2		1	1
1	1	0	1		2
1	1	0	1	2	

Confidence: $c(i \to j) = \dfrac{|\mathcal{U}_i \cap \mathcal{U}_j|}{|\mathcal{U}_i|}$

Support: $S(i \to j) = \dfrac{|\mathcal{U}_i \cap \mathcal{U}_j|}{|\mathcal{U}|}$

Towards personalized recommendations...

- Recommendations approaches
  - Collaborative Filtering approaches
  - Content-Based Filtering approaches
  - Hybrid approaches
  - Context-aware approaches