Overview of Recommender Systems

Masoud Mansoury
AMLab, University of Amsterdam
Discovery Lab, Elsevier

Recommendation Problems

- **Prediction version of problem**
  - Predict the rating value for a user-item combination.
  - It is assumed that training data is available, indicating user preferences for items.
  - The missing (or unobserved) values are predicted using this training model.
  - Also referred to as the *matrix completion problem*.
- **Ranking version of problem**
  - Not necessary to predict the ratings of users for specific items in order to make recommendations to users.
  - Rather, recommend the *top-k* items for a particular user.
  - Also referred to as the *top-k recommendation problem*.

Recommendation Problems

- **Prediction version of problem**

	i1	i2	i3	i4	i5	i6
u1	7	?	7	?	?	4
u2	6	7	?	4	3	?
u3	?	3	?	1	1	?
u4	1	?	2	3	?	4
u5	1	?	1	?	3	3

Recommendation Model

	i1	i2	i3	i4	i5	i6
u1	7	6	7	4	5	4
u2	6	7	4	4	3	4
u3	3	3	3	1	1	2
u4	1	2	2	3	3	4
u5	1	1	1	2	3	3

- **Ranking version of problem**

	i1	i2	i3	i4	i5	i6
u1	7	?	7	?	?	4
u2	6	7	?	4	3	?
u3	?	3	?	1	1	?
u4	1	?	2	3	?	4
u5	1	?	1	2	3	3

Recommendation Model

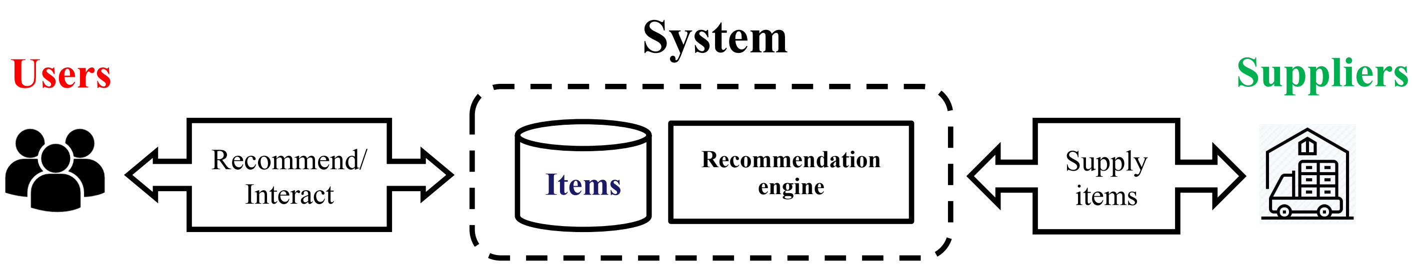
Top-2 recommendation lists

u1	i2
u1	i5
u2	i3
u2	i6
u3	i1
u3	i3
u4	i5
u4	i2
u5	i4
u5	i2

Goals of Recommender Systems

- **Increasing product sales is the primary goal of a recommender system.**

- **Recommender systems are, after all, utilized by merchants to increase their profit.**
  - By recommending carefully selected items to users, recommender systems bring relevant items to the attention of users.
  - This increases the sales volume and profits for the merchant.



Goals of Recommender Systems

- ***Business-centric*** **goal of increasing revenue:**
  - Relevance
    - The most obvious operational goal of a recommender system is to recommend items that are relevant to the user at hand.
    - Users are more likely to consume items they find interesting.
  - Novelty
    - Recommender systems are truly helpful when the recommended item is something that the user has not seen in the past.
    - For example, popular movies of a preferred genre would rarely be novel to the user.
    - Repeated recommendation of popular items can also lead to reduction in sales diversity

Goals of Recommender Systems

- Serendipity
  - The items recommended are somewhat unexpected.
  - Serendipity is different from novelty in that the recommendations are truly surprising to the user, rather than simply something they did not know about before.
  - If a new Indian restaurant opens in a neighborhood, the recommendation of that restaurant to a user who normally eats Indian food is novel but not necessarily serendipitous.
  - When Ethiopian food is recommended (unknown that such food might appeal to her), the recommendation is serendipitous.
  - Serendipity $\longrightarrow$ increasing sales diversity or beginning a new trend of interest in the user.

Goals of Recommender Systems

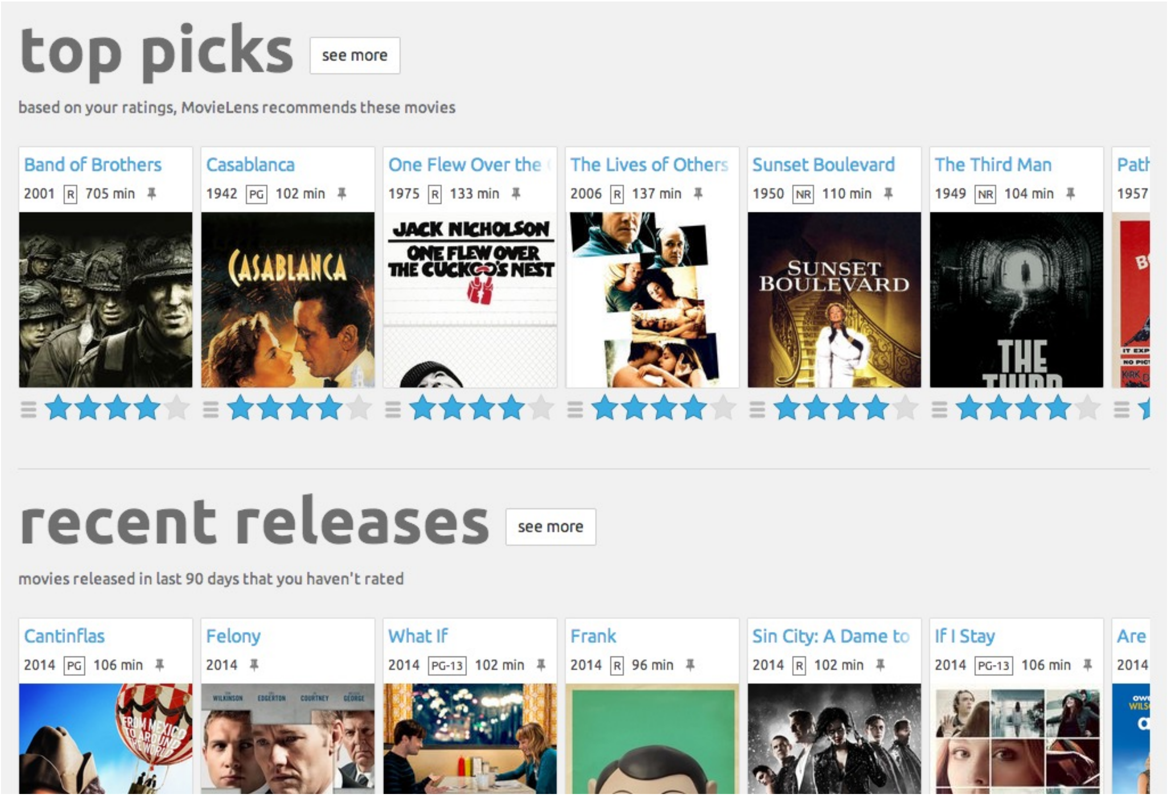
- Increasing recommendation diversity
  - Recommender systems typically suggest a list of top-k items.
  - When all these recommended items are very similar, it increases the risk that the user might not like any of these items.
  - On the other hand, when the recommended list contains items of different types, there is a greater chance that the user might like at least one of these items.
  - Diversity has the benefit of ensuring that the user does not get bored by repeated recommendation of similar items.

Goals of Recommender Systems

- **Soft goals**
  - From the perspective of the user,
    - recommendations can help improve overall user satisfaction.
    - Increased user loyalty and sale at the site, e.g., a user repeatedly receiving relevant recommendations from Amazon.com is more likely to use the site again.
  - At the merchant end,
    - the recommendation process can provide insights into the needs of the user and help customize the user experience further.
  - Providing the user an explanation for why a particular item is recommended is often useful.
    - For example, in the case of Netflix, recommendations are provided along with previously watched movies.

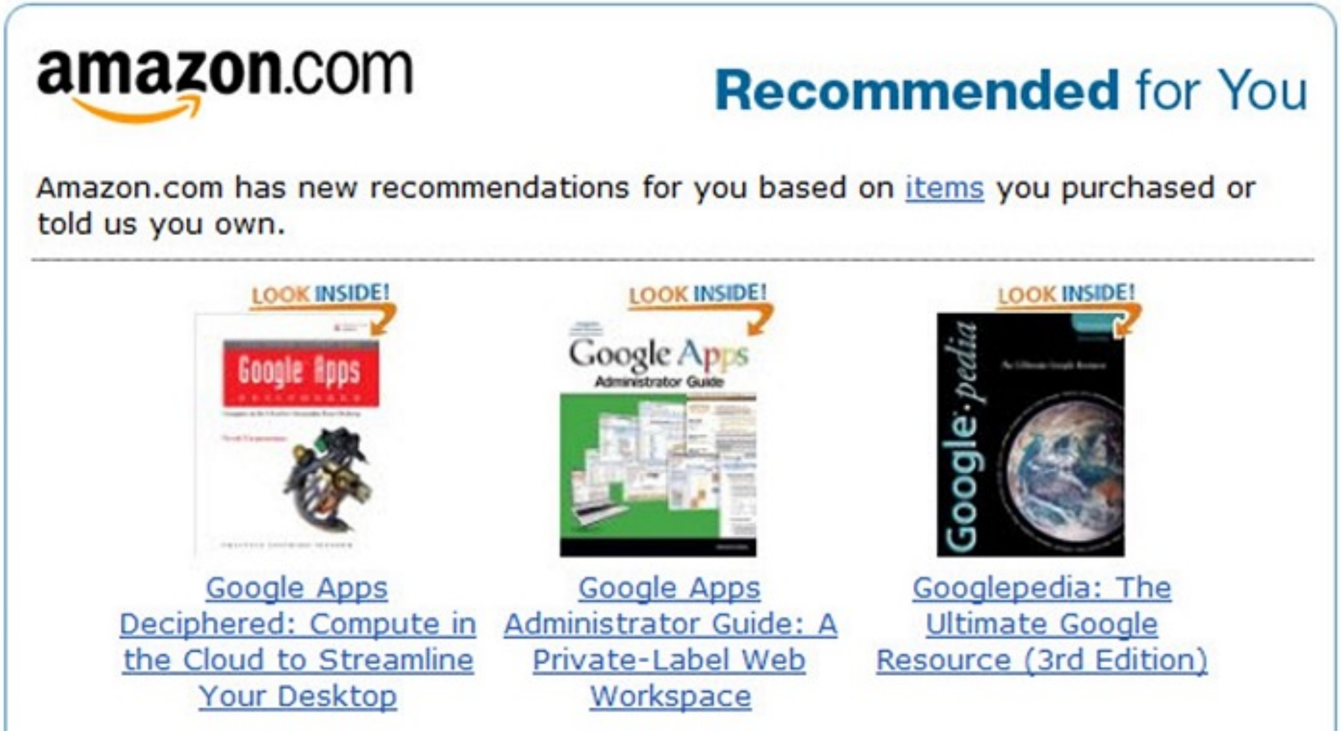
Examples of Existing Recommender Systems

- **GroupLens Recommender System**
  - A research prototype for recommendation of Usenet news, and later MovieLens and BookLens



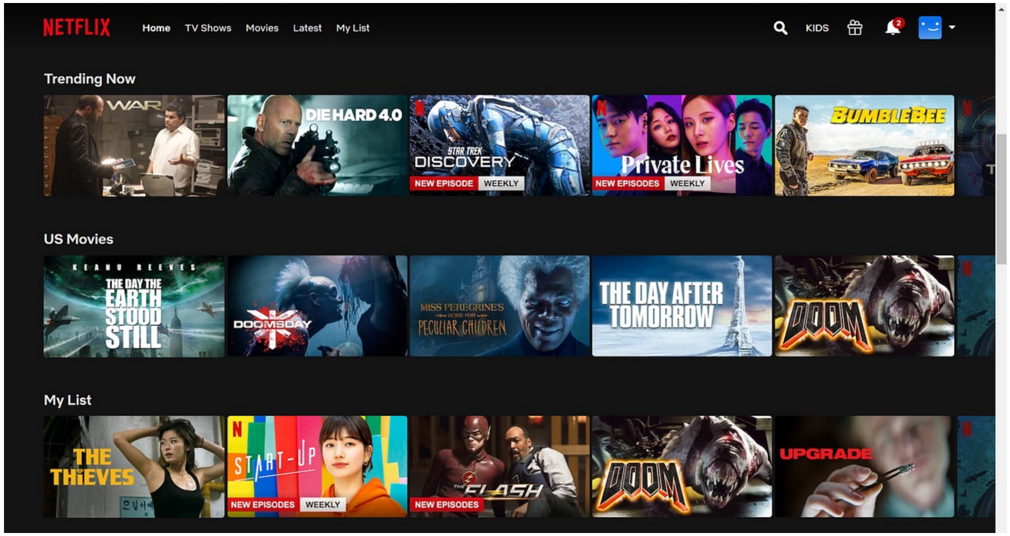
Examples of Existing Recommender Systems

- **Amazon.com Recommender System**
  - Product recommendation, e.g., books, software, electronics, etc.
  - On the basis of explicitly provided ratings, buying behavior, and browsing behavior.



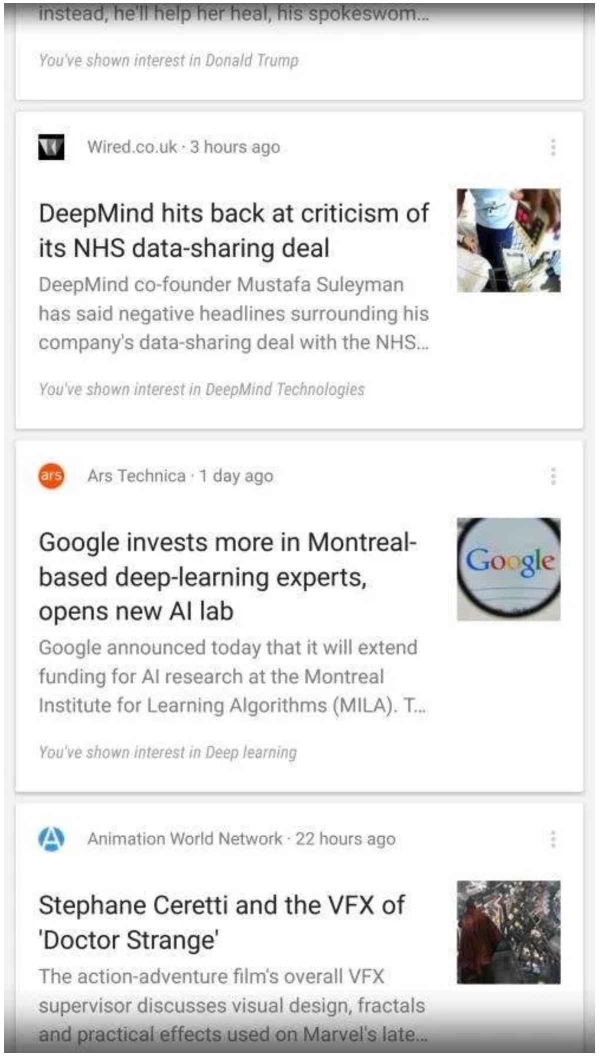
Examples of Existing Recommender Systems

- **Netflix Movie Recommender System**
  - Netflix has contributed significantly to the research community as a result of the *Netflix Prize contest*.
  - This contest was designed to provide a forum for competition among various collaborative filtering algorithms contributed by contestants.
  - A data set of Netflix movie ratings was released, and the task was to predict ratings of particular user-item combinations.



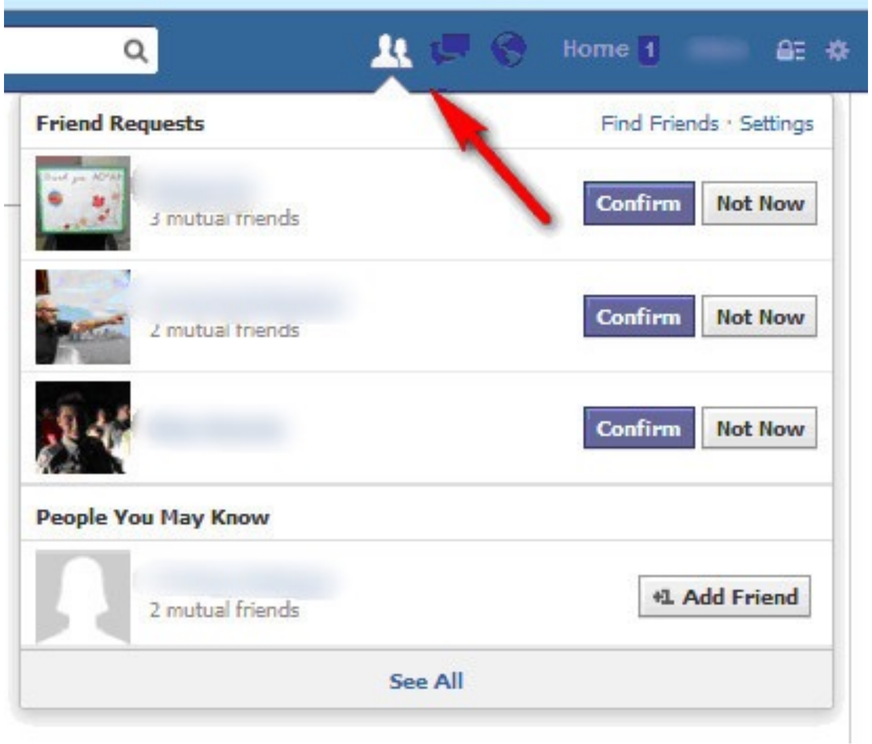
Examples of Existing Recommender Systems

- **Google News Recommender System**
  - New recommendation to users based on their history of clicks
  - *unary ratings*, in which a user can express her affinity for an item, but no mechanism exists for her to show dislike.
  - The ratings are *implicit*, because they are inferred from user actions rather than being explicitly specified by the user.



Examples of Existing Recommender Systems

- **Facebook Friend Recommender System**
  - Goals are different than product recommendations
  - Social networks are heavily dependent on the growth of the network to increase their advertising revenues.
  - Therefore, the recommendation of potential friends (or links) enables better growth and connectivity of the network.
  - Referred to as *link prediction* in the field of social network analysis.
  - Recommendations based on *structural relationships* than ratings data.



Examples of Existing Recommender Systems

System	Product Goal
Amazon.com	Books and other products
Netflix	DVDs, Streaming Videos
Jester	Jokes
GroupLens	News
MovieLens	Movies
Last.fm	Music
Google News	News
Google Search	Advertisements
Facebook	Friends, Advertisements
Pandora	Music
YouTube	Online Videos
TripAdvisor	Travel products
IMDb	Movies

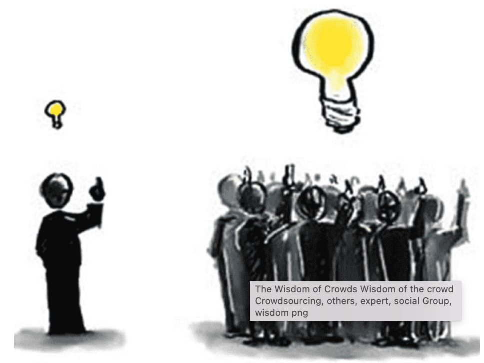
Recommendation Approaches

- **Collaborative Filtering Models**
  - Memory-based mothods (*neighborhood-based algorithms*)
    - User-based collaborative filtering
    - Item-based collaborative filtering
  - Model-based methods
- **Content-based Recommender Systems**
- **Knowledge-based Recommender Systems**
- **Demographic Recommender Systems**
- **Hybrid and Ensemble-based Recommender Systems**

Collaborative Filtering (CF)

Collaborative Filtering (CF)

- **The most prominent approach to generate recommendations**
  - Used by large, commenrcial e-commerce sites
  - Well-understood, various algorithms and variations exist
  - Applicable in many domains (book, movie, …)



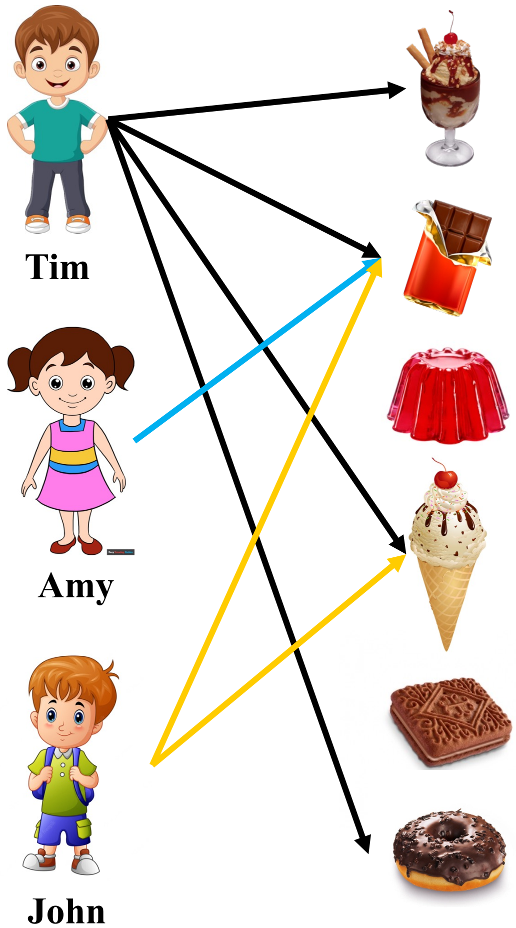
- **Approach**
  - Use the “wisdom of the crowd” to recommend items

- **Basic assumption and ideas**
  - Users give ratings to catalog items (implicitly or explicitly)
  - Customers who had similar tastes in the past, will have similar tastes in the future

Neighborhood-based CF

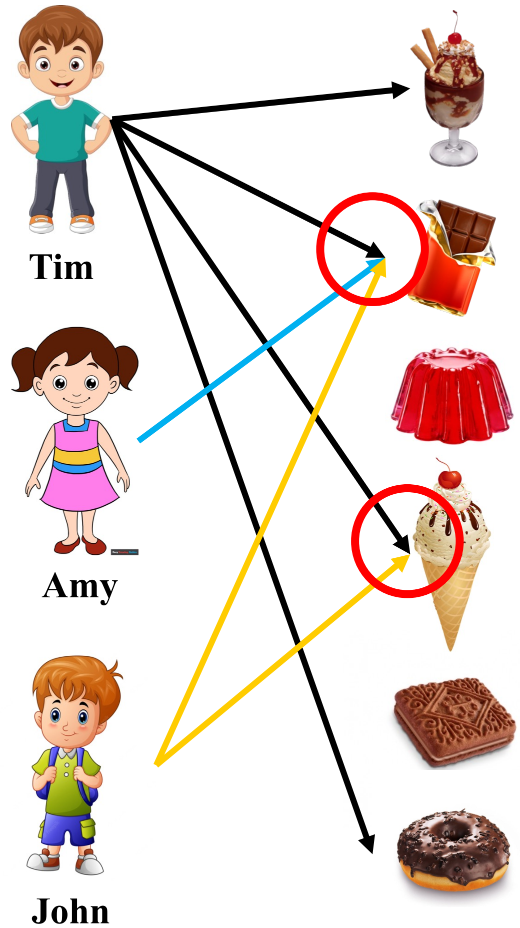
- **User-based Collaborative Filtering**
  - Ratings provided by like-minded users of a target user are used to make the recommendations.
  - The basic idea is to determine users, who are similar to the target user A, and recommend ratings for the unobserved ratings of A by computing weighted averages of the ratings of this peer group.
  - e.g., if Alice and Bob have rated movies in a similar way in the past, then Alice's observed ratings on the movie Terminator can be used to predict Bob's unobserved ratings on this movie.
- **Item-based Collaborative Filtering**
  - To make the rating predictions for target item B by user A, the first step is to determine a set S of items that are most similar to target item B.
  - The ratings in item set S, which are specified by A, are used to predict whether the user A will like item B.
  - e.g., Bob's ratings on similar science fiction movies like Alien and Predator can be used to predict his rating on Terminator.

User-Based VS. Item-Based CF



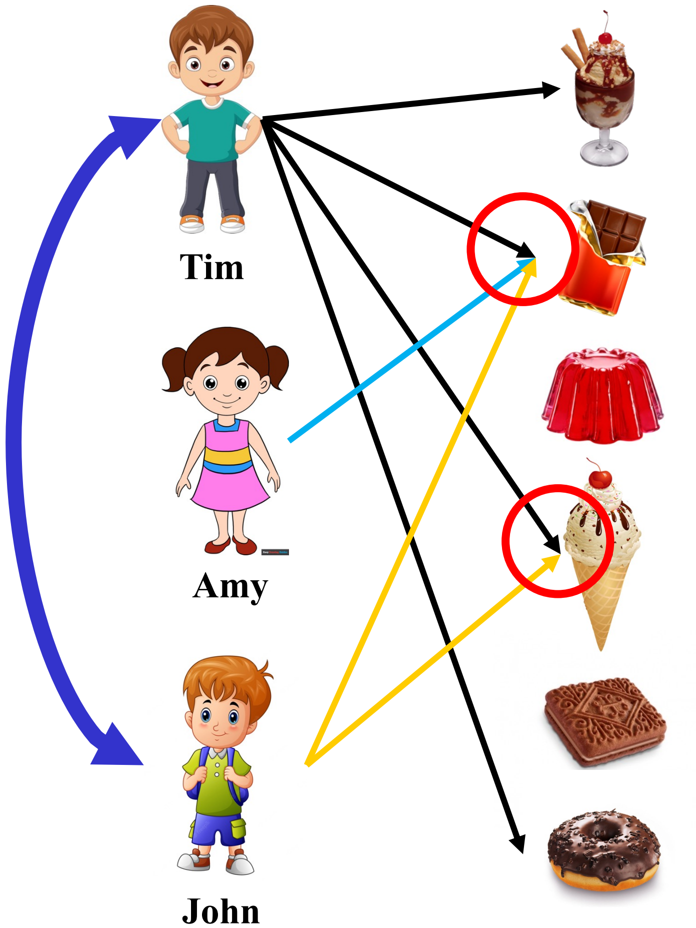
User-based Collaborative Filtering

User-Based VS. Item-Based CF



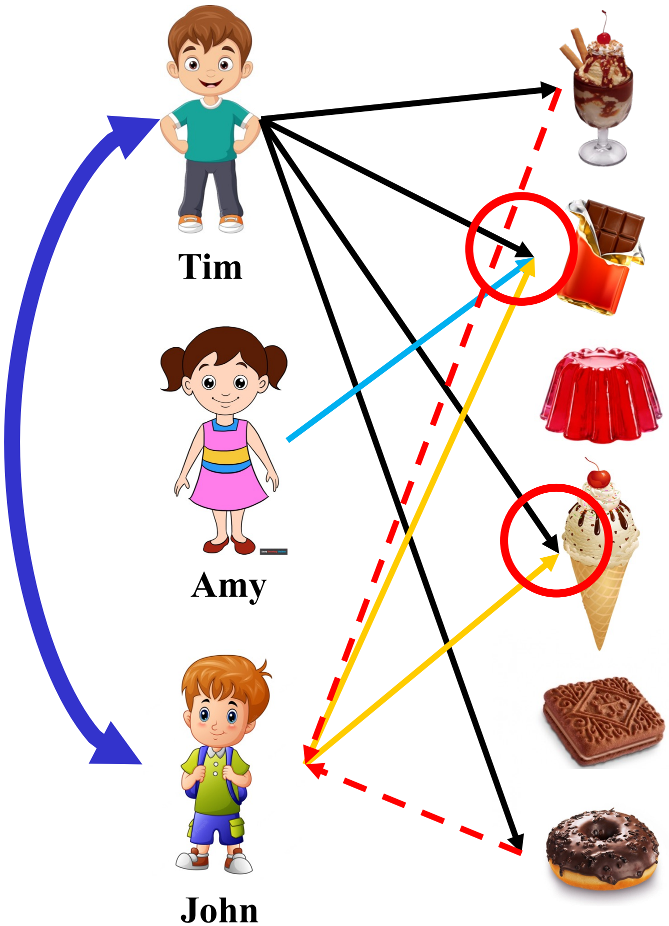
User-based Collaborative Filtering

User-Based VS. Item-Based CF



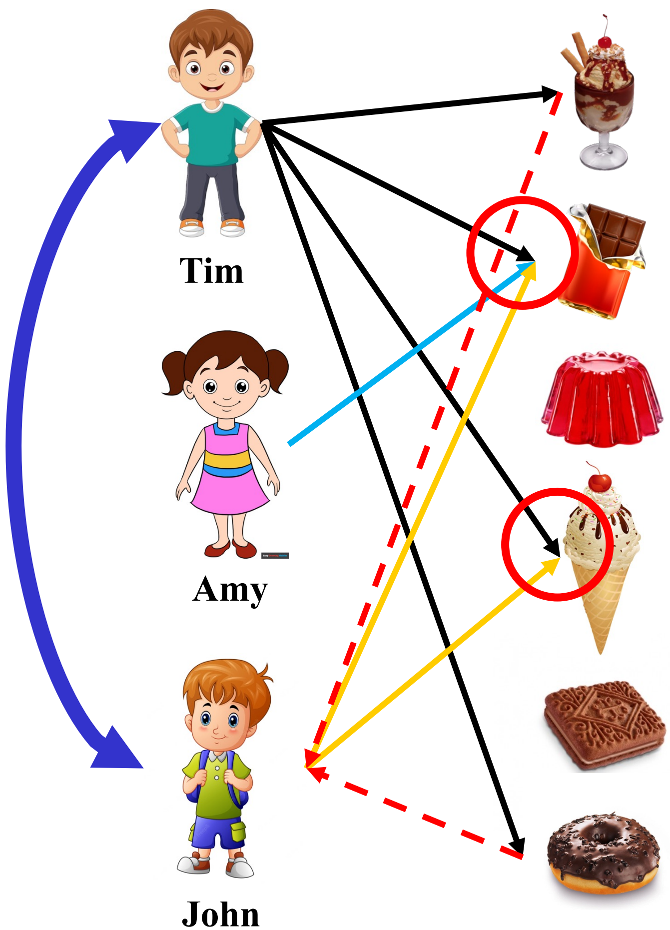
User-based Collaborative Filtering

User-Based VS. Item-Based CF

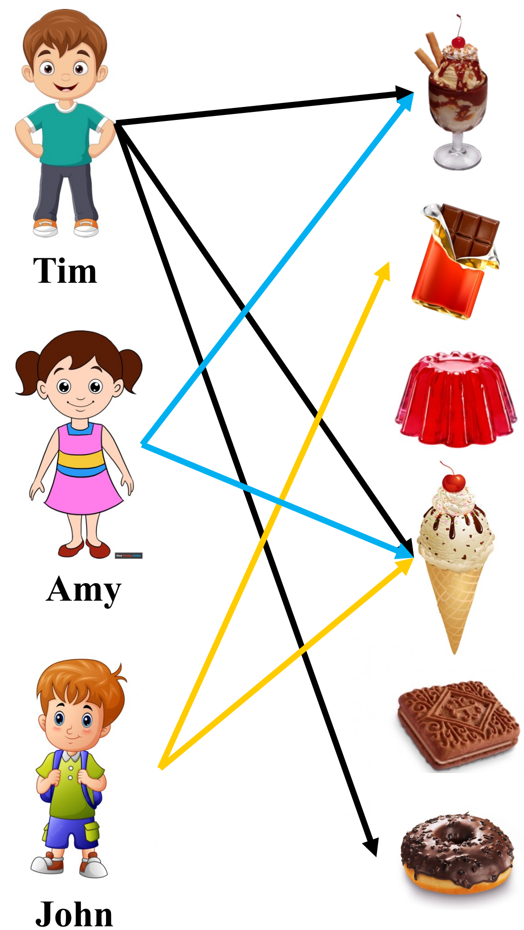


User-based Collaborative Filtering

User-Based VS. Item-Based CF

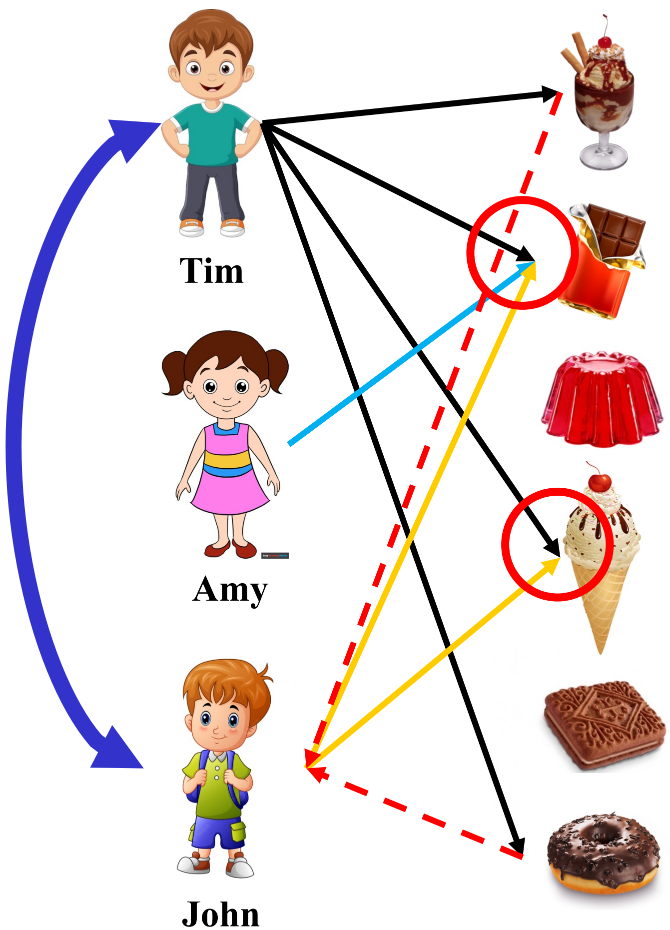


User-based Collaborative Filtering

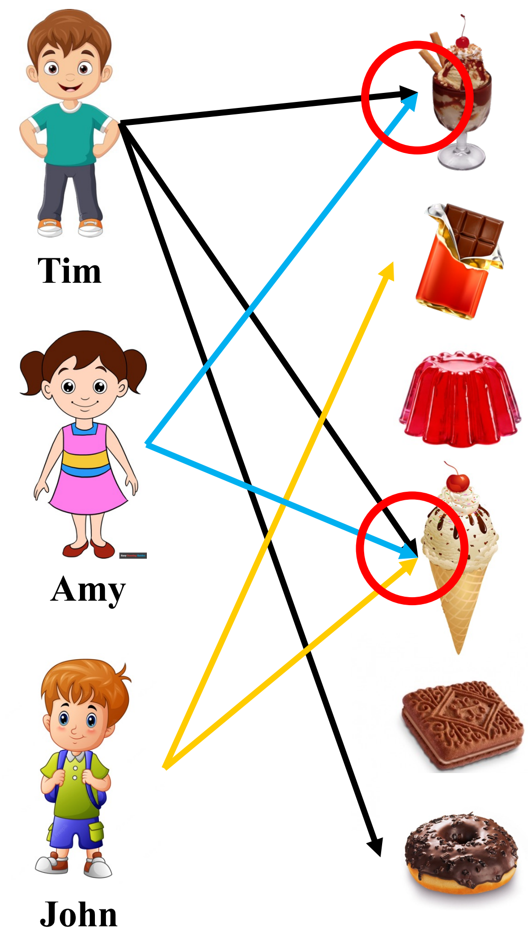


Item-based Collaborative Filtering

User-Based VS. Item-Based CF

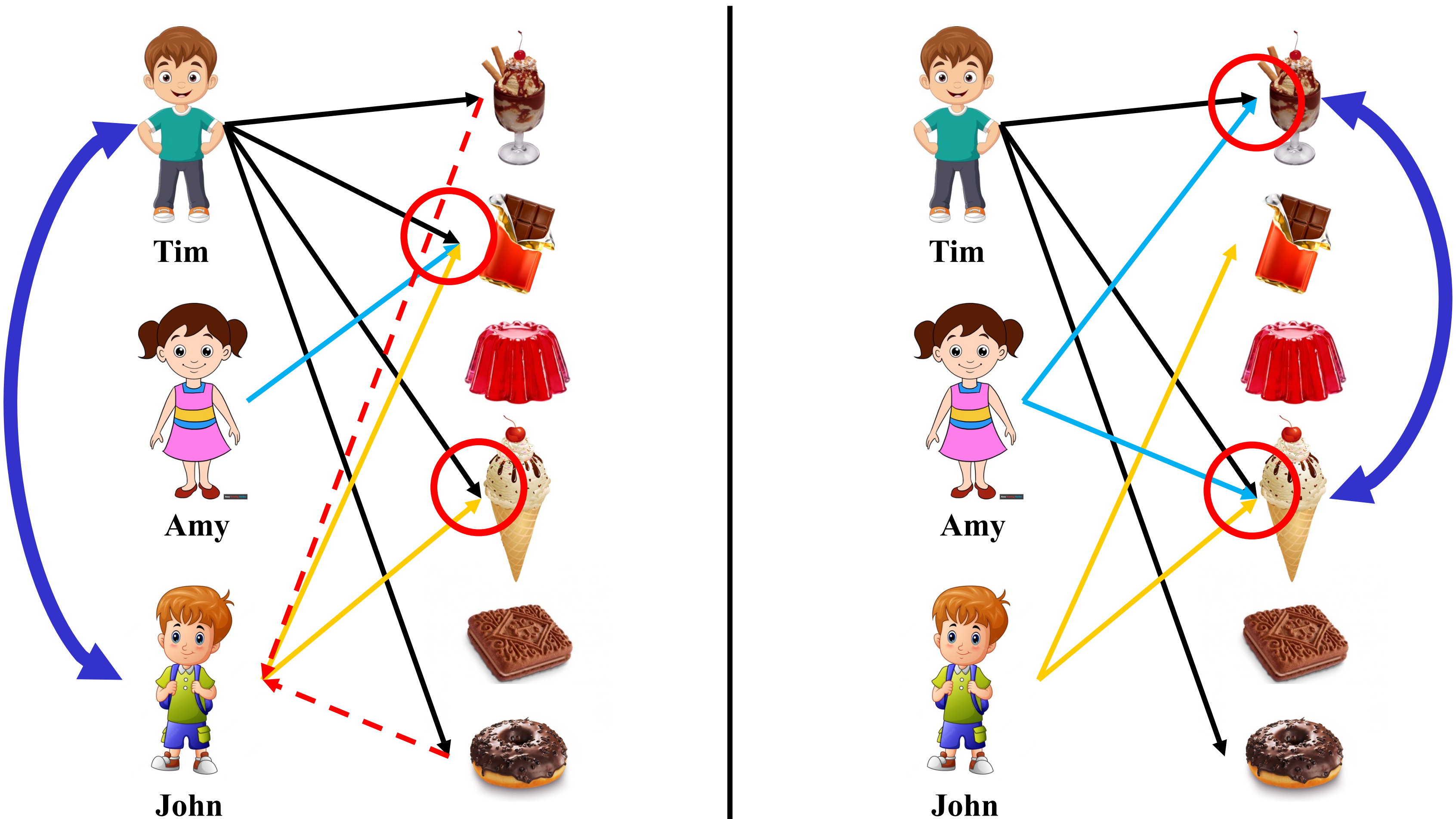


User-based Collaborative Filtering



Item-based Collaborative Filtering

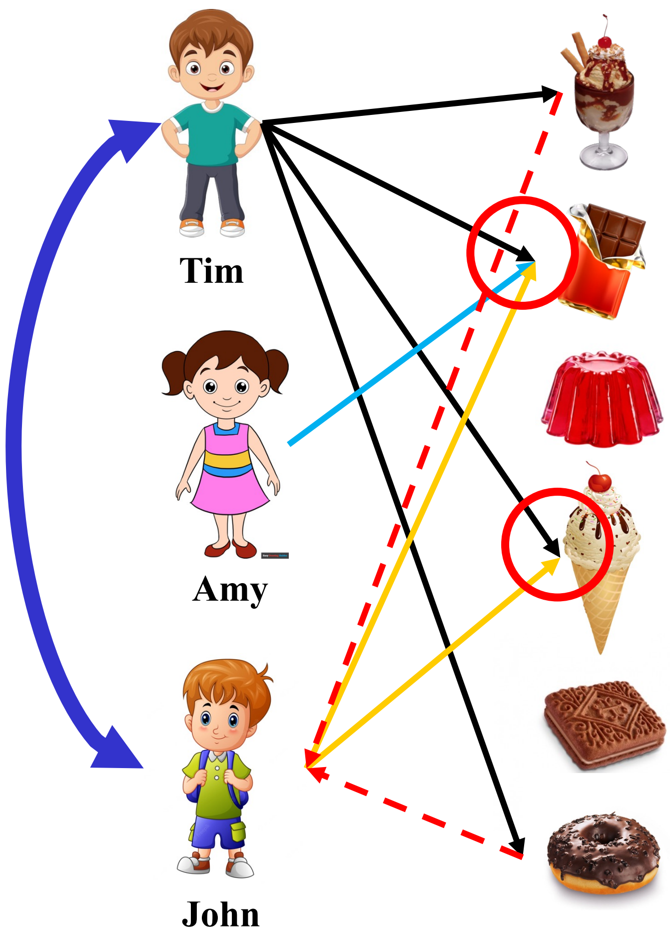
User-Based VS. Item-Based CF



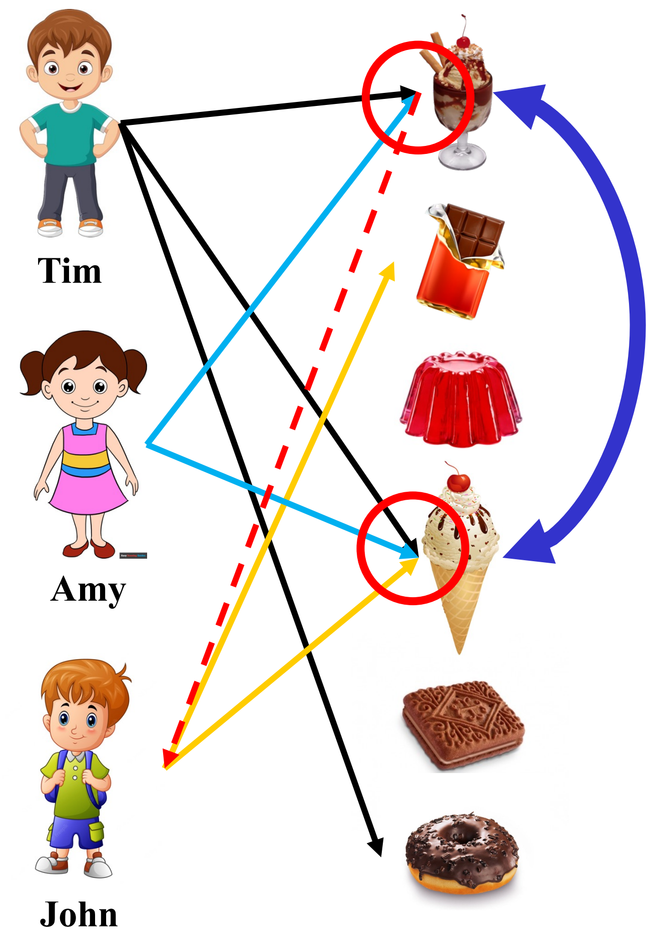
User-based Collaborative Filtering

Item-based Collaborative Filtering

User-Based VS. Item-Based CF



User-based Collaborative Filtering



Item-based Collaborative Filtering

Model-based CF

- **In model-based CF, machine learning and data mining methods are used in the context of predictive models.**

- **In cases where the model is parameterized, the parameters of this model are learned within the context of an optimization framework.**

- **Some examples of such model-based methods include decision trees, rule-based models, Bayesian methods and latent factor models.**

- **Many of these methods, such as latent factor models, have a high level of coverage even for *sparse ratings* matrices.**

Model-based CF

	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
	✓		✓	✓	
		✓			✓
	✓	✓	✓		
				✓	✓

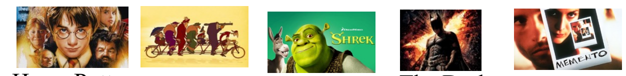
Model-based CF

	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
	✓		✓	✓	
		✓			✓
	✓	✓	✓		
				✓	✓

≈

1	.1
-1	0
.2	-1
.1	1

×



Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
.9	-1	1	1	-.9
-.2	-.8	-1	.9	1

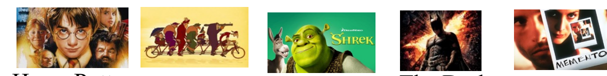
Model-based CF

Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
✓		✓	✓	
	✓			✓
✓	✓	✓		
			✓	✓

$\approx$

Users hidden feature 1	Users hidden feature 2
1	.1
-1	0
.2	-1
.1	1

$\times$

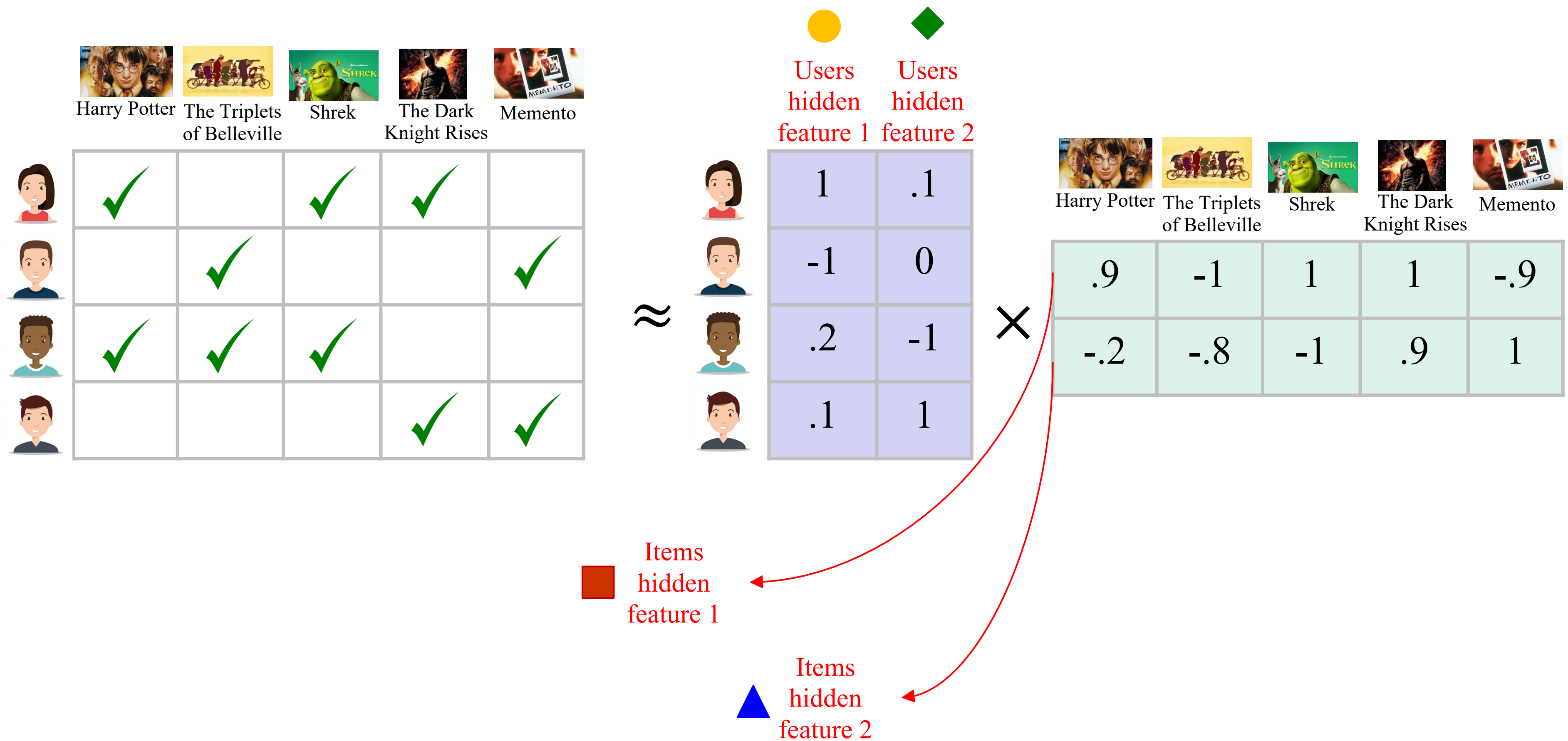


Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
.9	-1	1	1	-.9
-.2	-.8	-1	.9	1

Items hidden feature 1

Items hidden feature 2

Model-based CF



arthouse <-> blockbuster

children's <-> adult's

preference for arthouse <-> blockbuster

preference for children's <-> adult's

Model-based CF

	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
User 1	✓		✓	✓	
User 2		✓			✓
User 3	✓	✓	✓		
User 4				✓	✓

≈

	●	◆
User 1	1	.1
User 2	-1	0
User 3	.2	-1
User 4	.1	1

	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
■	.9	-1	1	1	-.9
▲	-.2	-.8	-1	.9	1

	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
User 1	.88	-1.08	0.9	1.09	-0.8
User 2	-0.9	1.0	-1.0	-1.0	0.9
User 3	0.38	0.6	1.2	-0.7	-1.18
User 4	-0.11	-0.9	-0.9	1.0	0.91

■ arthouse <-> blockbuster

● preference for arthouse <-> blockbuster

▲ children's <-> adult's

◆ preference for children's <-> adult's

Content-Based Recommender Systems

Content-Based Recommender Systems

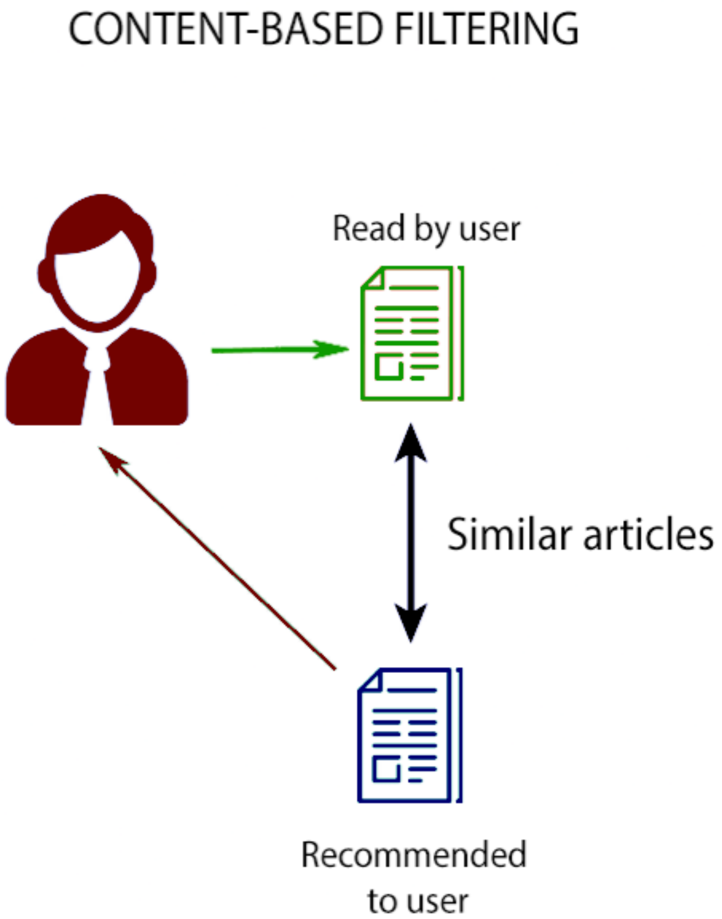
- **The descriptive attributes of items are used to make recommendations.**
  - The term "content" refers to these descriptions.

- **The ratings and buying behavior of users are combined with the content information available in the items.**

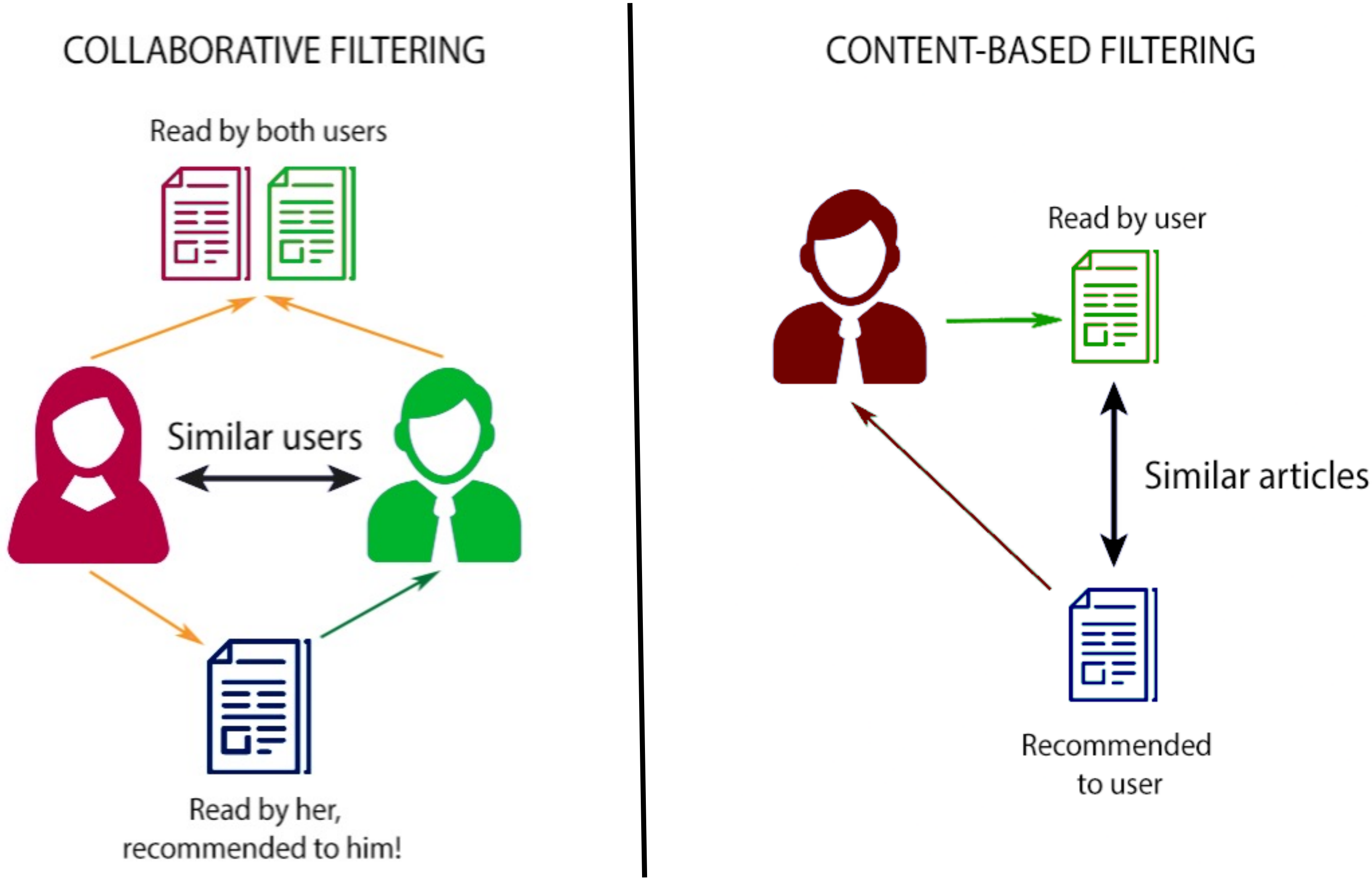
Collaborative Filtering cannot be used as the ratings of other users are not available

Content-based RS is a good choice in this situation because the item description of Terminator contains similar genre keywords as other science fiction movies, such as Alien and Predator.

Content-Based Recommender Systems



Content-based Vs. Collaborative Filtering



Content-Based Recommender Systems

- **+ This method is able to make recommendation for *new items*, when sufficient rating data are not available for that item.**

- **− This method is not effective at providing recommendations for *new users*.**
  - Enough rating data is not available to make robust recommendations.
- **− In many cases, content-based methods provide *obvious* recommendations because of the use of keywords or content.**
  - For example, if a user has never consumed an item with a particular set of keywords, such an item has no chance of being recommended.
  - This is because the constructed model is specific to the user at hand, and the community knowledge from similar users is not leveraged.
  - This phenomenon tends to reduce the diversity of the recommended items, which is undesirable.

Knowledge-Based Recommender Systems

Knowledge-Based Recommender Systems

- **Useful method when sufficient ratings are not available for the recommendation process, or when items are not purchased very often.**
  - Domains like real estate, automobiles, tourism requests, financial services, or expensive luxury goods.
  - It might be difficult to fully capture user interest with historical data.
    - Item domain might be complex in terms of its varied properties, and user interests may be regulated by a very specific combination of these options.
    - For example, cars may have several makes, models, colors, engine options, and interior options.

- **Customers want to define their requirements explicitly**
  - “the color of the car should be black

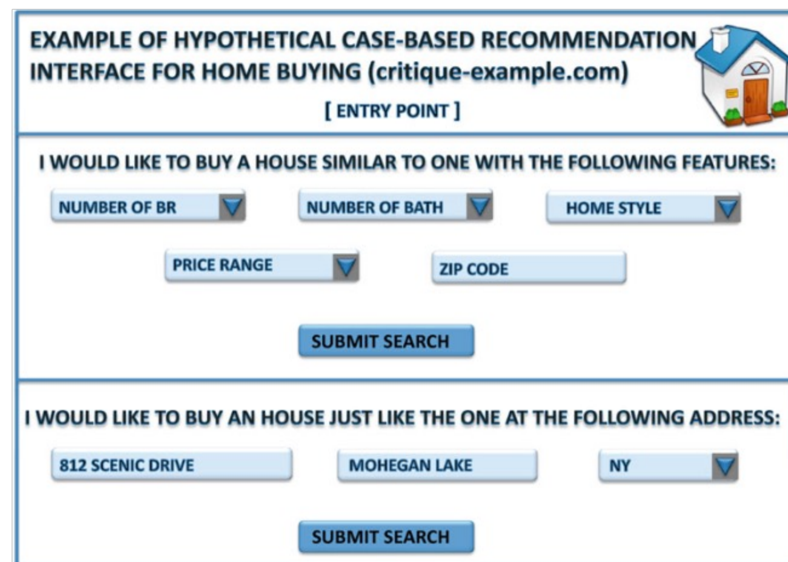
Knowledge-Based Recommender Systems

- **Knowledge-based RS can handle these cases as the recommendation process is performed on the basis of**
  - similarities between customer requirements and item descriptions, or
  - the use of constraints specifying user requirements.

Knowledge-Based Recommender Systems

1. **Constraint-based recommender systems**
   - users typically specify requirements or constraints (e.g., lower or upper limits) on the item attributes.

2. **Case-based recommender systems**
   - users specify specific cases as targets or anchor points, and similarity metrics are defined on the item attributes to retrieve similar items to these cases.



Demographic Recommender Systems

Demographic Recommender Systems

- **The demographic information about the user is leveraged to learn classifiers that can map specific demographics to ratings or buying propensities.**
  - An early recommender system, referred to as Grundy, recommended the same books to all persons in the same group.

- **Not a good recommender system on a stand-alone basis, but**
  - Adds significantly to the power of other recommender systems as a component of hybrid or ensemble models.

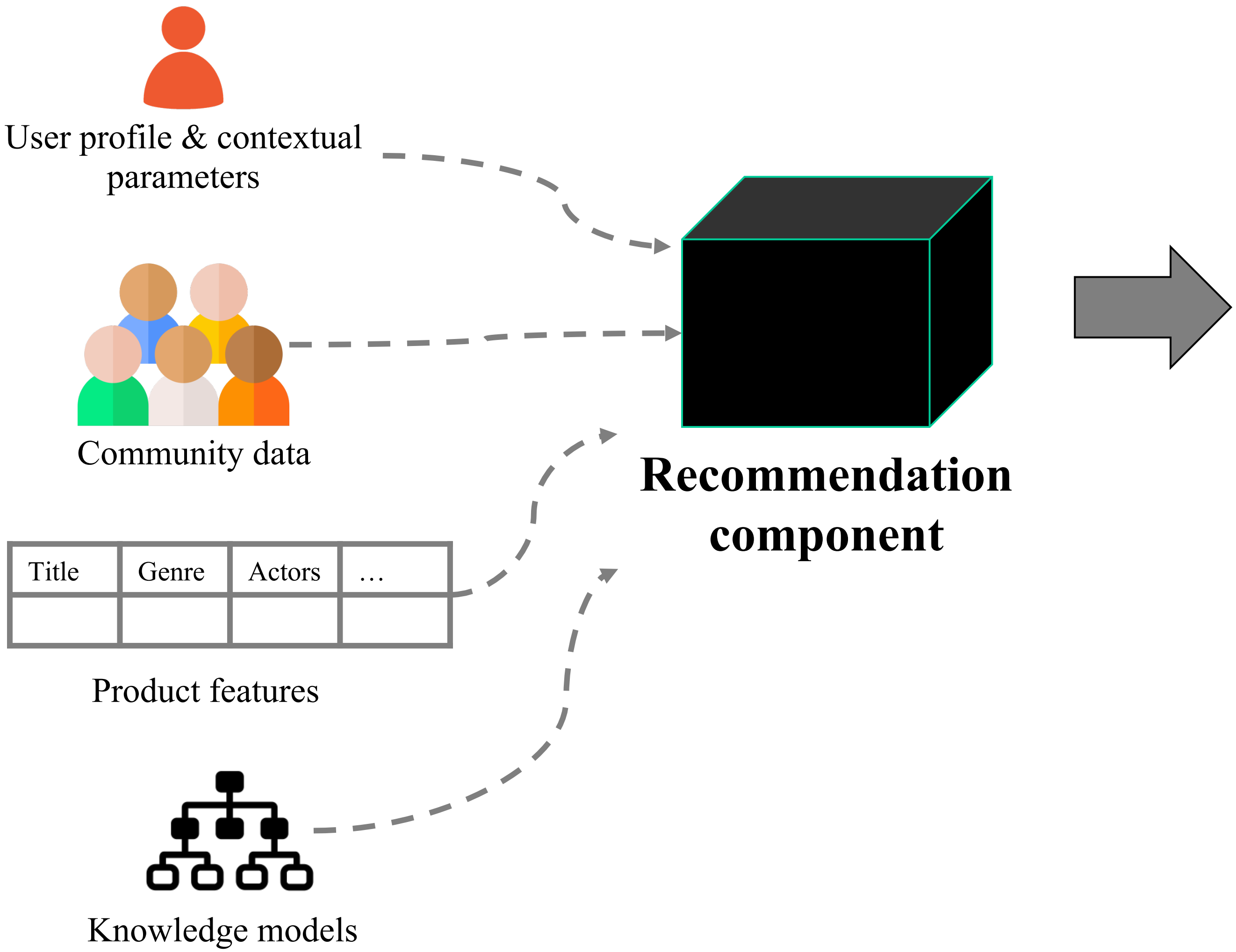
Hybrid Recommender Systems

Hybrid Recommender Systems

- **Exploits different sources of input, and they may work well in different scenarios.**
  - Collaborative filtering systems rely on community ratings,
  - Content-based methods rely on textual descriptions and the target user's own ratings,
  - Knowledge-based systems rely on interactions with the user in the context of knowledge bases.
  - Demographic systems use the demographic profiles of the users to make recommendations.

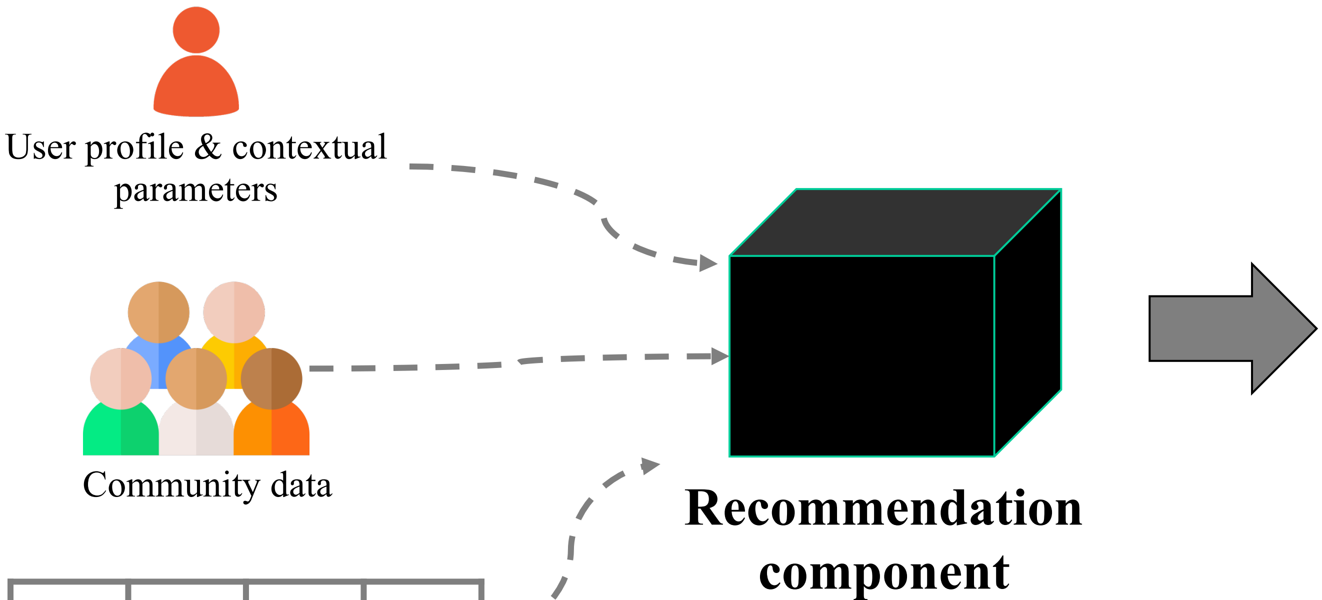
- **Combining different types of recommender systems for improving the performance**
  - In many cases where a wider variety of inputs is available, one has the flexibility of using different types of recommender systems for the same task.

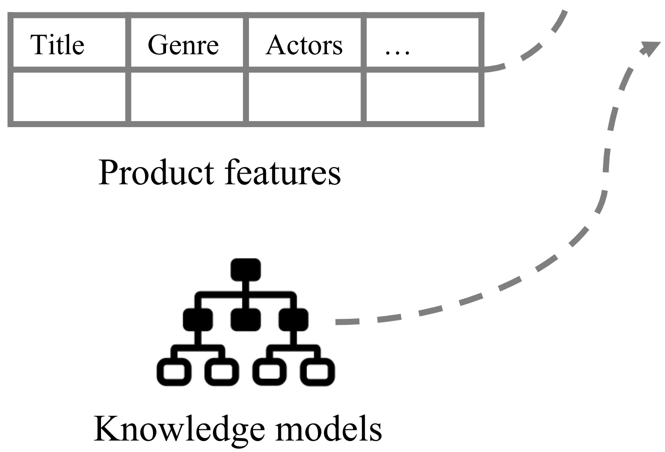
Hybrid Recommender Systems



item	score
$i1$	0.9
$i2$	1
$i3$	0.3
…	…

Hybrid Recommender Systems





item	score
$i1$	0.9
$i2$	1
$i3$	0.3
…	…

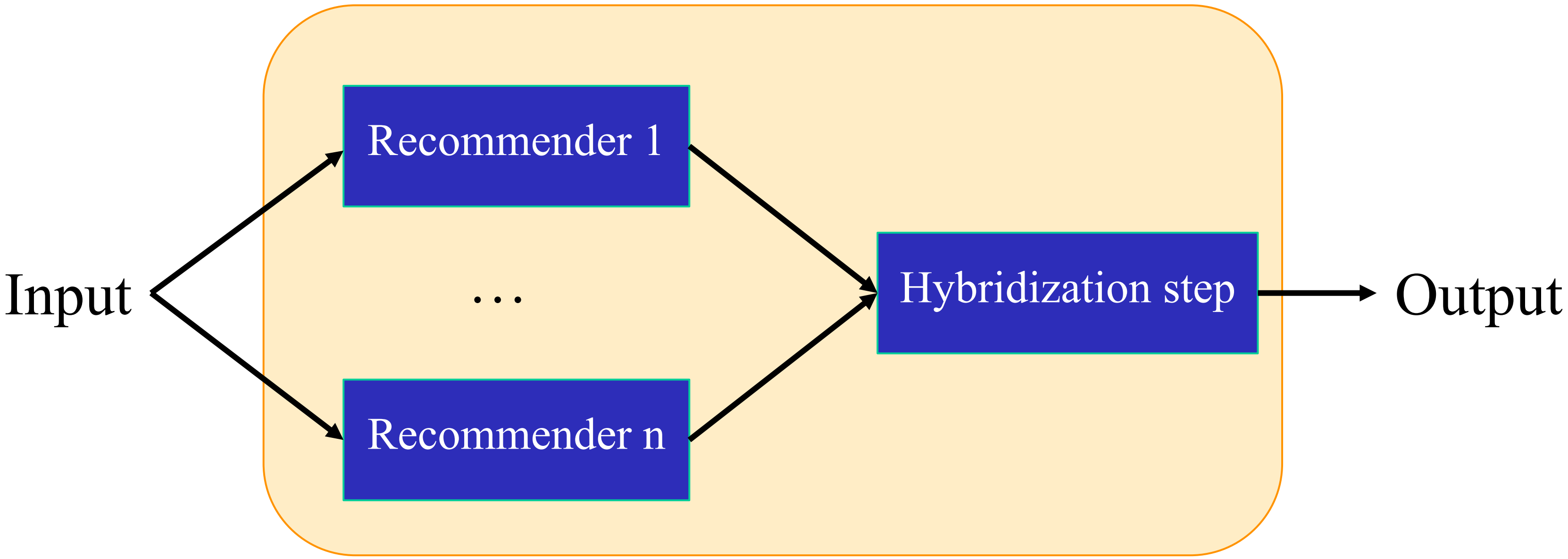
- **Collaborative:** "Tell me what's popular among my peers"
- **Content-based:** "Show me more of the same what I've liked"
- **Knowledge-based:** "Tell me what fits best on my needs"

Hybrid Recommender Systems

- **Different hybridization designs**
  - Parallel use of several systems
  - Monolithic exploiting different features
  - Pipelined invocation of different systems

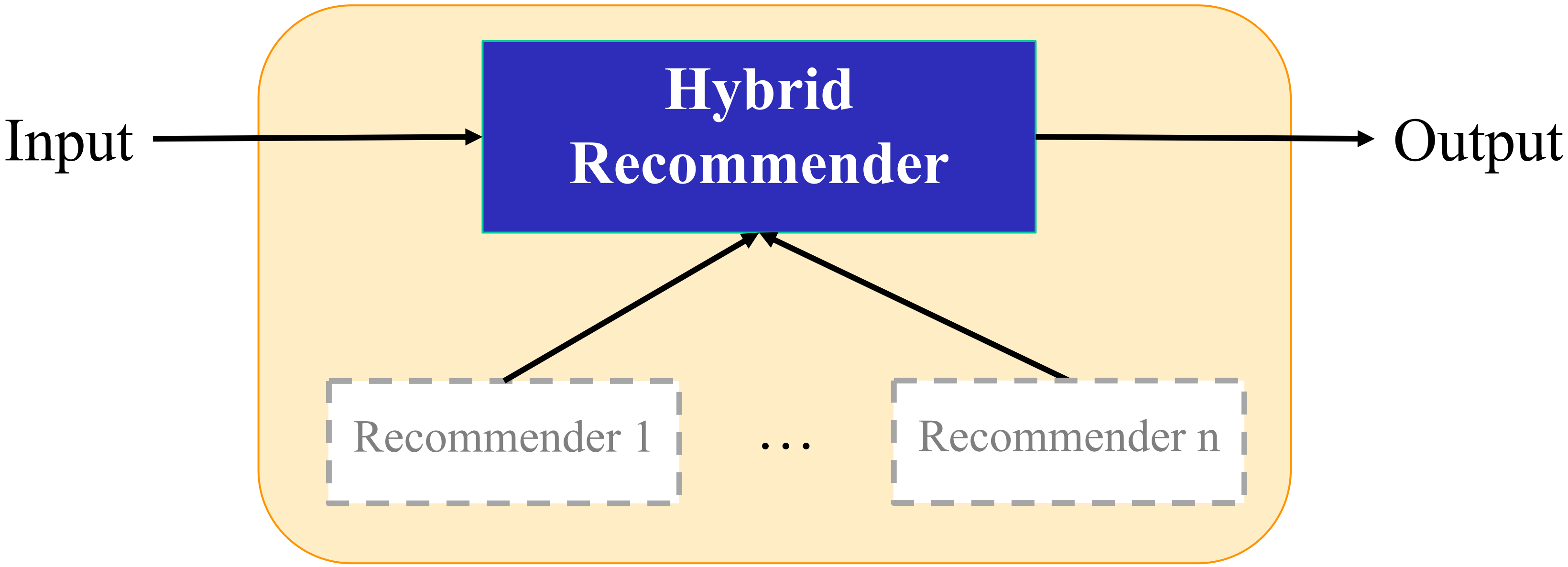
Hybrid Recommender Systems

- **Parallelized hybridization design**
  - Outputs of several existing recommendation models are combined.



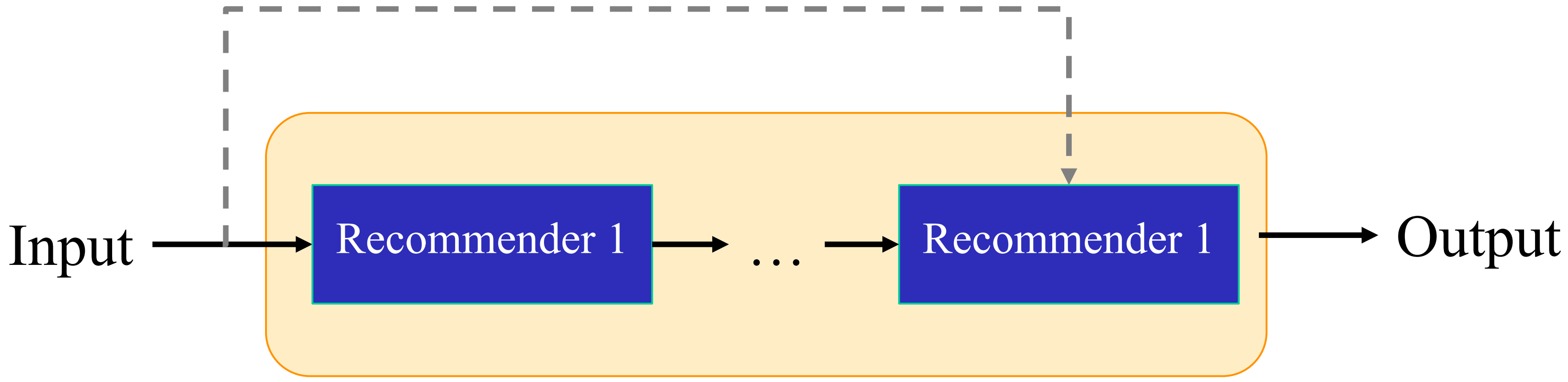
Hybrid Recommender Systems

- **Monolithic hybridization design**
  - Only a single recommendation component.

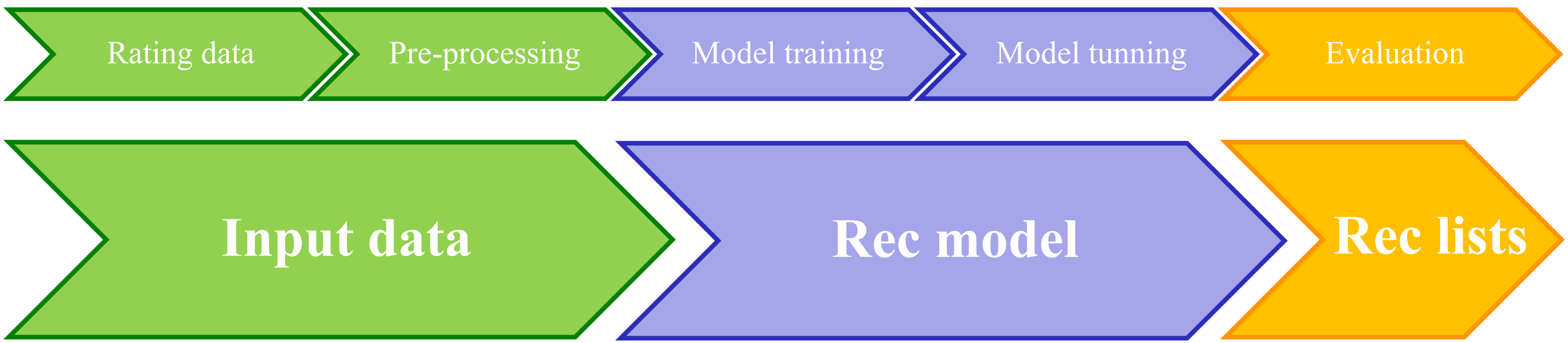


Hybrid Recommender Systems

- **Pipelined hybridization design**
  - One recommender system pre-processes some input for the subsequent one.



Where we are …



Ratings
Item contents
User attributes
User requirements
User targets (cases)
…

Collaborative Filtering
Content-based
Knowledge-based
Demographic
Hybrid
…

How to measure the quality of the recommendations?
How to evaluate the effectiveness of the recommendation model?

Overview of Recommender Systems

Masoud Mansoury
AMLab, University of Amsterdam
Discovery Lab, Elsevier