

The background of the slide is a complex, futuristic circuit board design. It features a mix of blue and purple hues with glowing white and yellow points of light. The circuit lines are intricate, with some forming circular patterns and others branching out. The overall aesthetic is high-tech and digital.

Recommender systems

RECOMMENDATION PROBLEM

Nguyen Ngoc Thao
nnthao@fit.hcmus.edu.vn

Content outline

- The recommendation problem
- Content-based recommendation
- Collaborative filtering

Recommender systems

- **Recommender systems (RS)** are widely used on the Web for suggesting products and services to users.

Frequently bought together



This item: SteelSeries Aerox 5 Wireless - Gaming Mouse - 18000 CPI -- TrueMove Air Optical Sens...



SteelSeries Apex Pro TKL Wireless HyperMagnetic Gaming Keyboard — World's Fastest Keyboard —...



SteelSeries QcK Gaming Mouse Pad - XL RGB Prism Cloth - Sized to Cover Desks

Image credit: [Amazon](https://www.amazon.com)

Products related to this item

Sponsored ⓘ



Razer DeathAdder V3 Pro Wireless Gaming Mouse: 64g Ultra Lightweight - Focus...
★★★★☆ 911



Razer DeathAdder V3 Wired Gaming Mouse: 59g Ultra Lightweight - Pro 30K Optical Sen...
★★★★☆ 272



ASUS ROG Spatha X Wireless Gaming Mouse (Magnetic Charging Stand, 12...
★★★★☆ 711



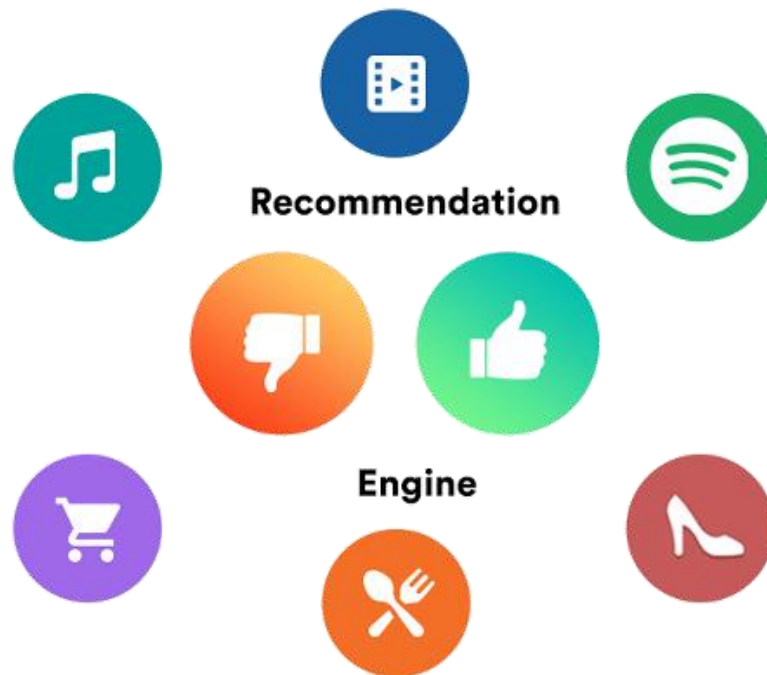
ROCCAT Kone XP Air – Wireless Customizable Ergonomic RGB Gaming Mouse, 19K DPI Opti...
★★★★☆ 589



Alienware AW720M Tri-Mode Wireless Gaming Mouse - 2.4GHz Wireless, Bluetooth...
★★★★☆ 1,350

Recommender systems

- RS are used in a wide variety of domains and industries.
- They help enhance user experiences, increase engagement, and improve business outcomes.



NETFLIX



facebook

ZARA

GrabFood

amazon

Avengers: Endgame

2019 · PG-13 · 3h 1m

IMDb RATING

★ 8.4/10
1.2M

YOUR



More like this



Action Adventure Drama

After the devastating events of Avengers: Infinity War, the remaining allies, the Avengers assemble once more to reverse the damage.



★ 7.8 ☆

Captain America: Civil War

+ Watchlist



★ 7.9 ☆

Thor: Ragnarok

+ Watchlist



★ 8.4 ☆

Joker

+ Watchlist



★ 8.4 ☆

The Dark Knight Rises

+ Watchlist



5 out of 5 ⓘ

Based on the opinion of 9 people

Do you recommend Rocky Mountain Foot and Ankle?

Yes

No



Ratings and reviews have changed

Now it's easier to find great businesses with recommendations

Learn More



Recommended by 23 people

MOST HELPFUL

MOST RECENT



Sara Thomason Stephens reviewed Rocky Mountain Foot and Ankle —

5★

December 19, 2016 · 🌐

Love this place. Dr. Ericson is fantastic. His receptionist so are great as is his Aides. Very compassionate group of people.



1



Like



Comment



Share



Buffer



Nicole Stevenson ★ recommends Rocky Mountain Foot and Ankle.

November 5, 2018 · 🌐

the staff is amazing!! lucky to have been accepted as a walk in. I was in major pain from an ingrown toenail!! they were all amazing and took care of me promptly!!



You have a new friend suggestion: ...

6 Jan at 6:53 am

You have a new friend suggestion: ...

5 Jan at 10:23 am

mentioned you in a comment.

2 Jan at 8:48 pm

People You May Know

See all friend recommendations



Grilled Cheese

72 mutual friends



Nicolas Cage

29 mutual friends



Sarah Michelle Gellar

74 mutual friends



Stephen King

13 mutual friends



MARKET SNAPSHOT

U.S.	EUROPE	ASIA
DJIA	13,342.50	+19.11 0.14%
S&P 500	1,436.88	+3.32 0.23%
NASDAQ	3,110.06	+5.53 0.18%

Canon P-215

Scan-tini
mobile document scanner

buy now
* get a free carrying case
promotion ends 12/31/12

Bloomberg

Our Company | Professional | Anywhere

Search News, Quotes and Opinion

HOME QUICK **NEWS** OPINION MARKET DATA PERSONAL FINANCE TECH POLITICS SUSTAINABILITY TV VIDEO RADIO

U.S. Stocks Rise as Investors Weigh Stimulus Prospects

By Inyoung Hwang - Sep 12, 2012 11:14 AM ET

f t in + 7 COMMENTS

QUEUE

Get The Market Now
newsletter. Learn more >

SUBSCRIBE

HEADLINES MOST POPULAR **RECOMMENDED**

Based on your reading history you may be interested in:

Glenn Beck Returns to Television With Dish Network Agreement



The Top Ten Stocks for Wednesday, September 12

U.S. [stocks](#) rose, with benchmark indexes trading near four-year highs, as a German court cleared the way for Europe's bailout fund and investors weighed prospects for stimulus measures from the [Federal Reserve](#).

JPMorgan Chase & Co. (JPM) and Travelers Cos. rose at least 1 percent, pacing gains among financial companies. PulteGroup Inc. (PHM) advanced 8.1 percent as homebuilders rallied.

U.S. Stocks Reverse Gains as Technology Shares Retreat

Euro, Stocks Climb as German Court Approves Bailout Funds

European Stocks Rise to 14-Month High After German Ruling

Ben & Jerry's Sues Porn Seller Over Flavor-Tied Titles

U.S. Stocks Extend Gain as France Said to Press Spain

Advertisement





Building an AI-Powered Outfit Recommendation System With Dataiku
(2021)

Recommender systems

- These systems **help users deal with the information overload** by **giving them personalized recommendations**.
 - E.g., given thousands of movies, a RS selects and recommends some movies to each user that he will most likely enjoy watching.



Vast number of users and items

The screenshot shows the Amazon homepage with the 'Shop by Department' dropdown menu open. The menu lists various categories, including 'All Departments', 'Amazon Home Services', 'Amazon Instant Video', 'Appliances', 'Apps & Games', 'Arts, Crafts & Sewing', 'Automotive', 'Baby', 'Beauty', 'Books', 'CDs & Vinyl', 'Cell Phones & Accessories', 'Clothing, Shoes & Jewelry', 'Electronics', 'Gift Cards', 'Grocery & Gourmet Food', 'Health & Personal Care', 'Home & Kitchen', 'Industrial & Scientific', 'Kindle Store', 'Luggage & Travel Gear', 'Magazine Subscriptions', 'Movies & TV', 'Musical Instruments', 'Office Products', 'Patio, Lawn & Garden', 'Pet Supplies', 'Prime Pantry', 'Software', 'Sports & Outdoors', 'Tools & Home Improvement', 'Toys & Games', 'Video Games', and 'Wine'. The 'All Departments' option is highlighted in blue. The 'Shop by Department' dropdown is located on the left side of the page, and the search results for 'departments' are displayed on the right.

amazon
Try Prime

Shop by
Department ▾

1-16 of 897,416 results for "departments"

Show results for

Movies & TV >
Mystery & Thrillers
Documentary

CDs & Vinyl >
Soundtracks
Alternative Rock
New Wave

Books >
Mysteries

All ▾ departments

All Departments

Amazon Home Services
Amazon Instant Video
Appliances
Apps & Games
Arts, Crafts & Sewing
Automotive
Baby
Beauty
Books
CDs & Vinyl
Cell Phones & Accessories
Clothing, Shoes & Jewelry
Electronics
Gift Cards
Grocery & Gourmet Food
Health & Personal Care
Home & Kitchen
Industrial & Scientific
Kindle Store
Luggage & Travel Gear

Collectibles & Fine Art
Computers
Credit and Payment Cards
Digital Music
Magazine Subscriptions
Movies & TV
Musical Instruments
Office Products
Patio, Lawn & Garden
Pet Supplies
Prime Pantry
Software
Sports & Outdoors
Tools & Home Improvement
Toys & Games
Video Games
Wine

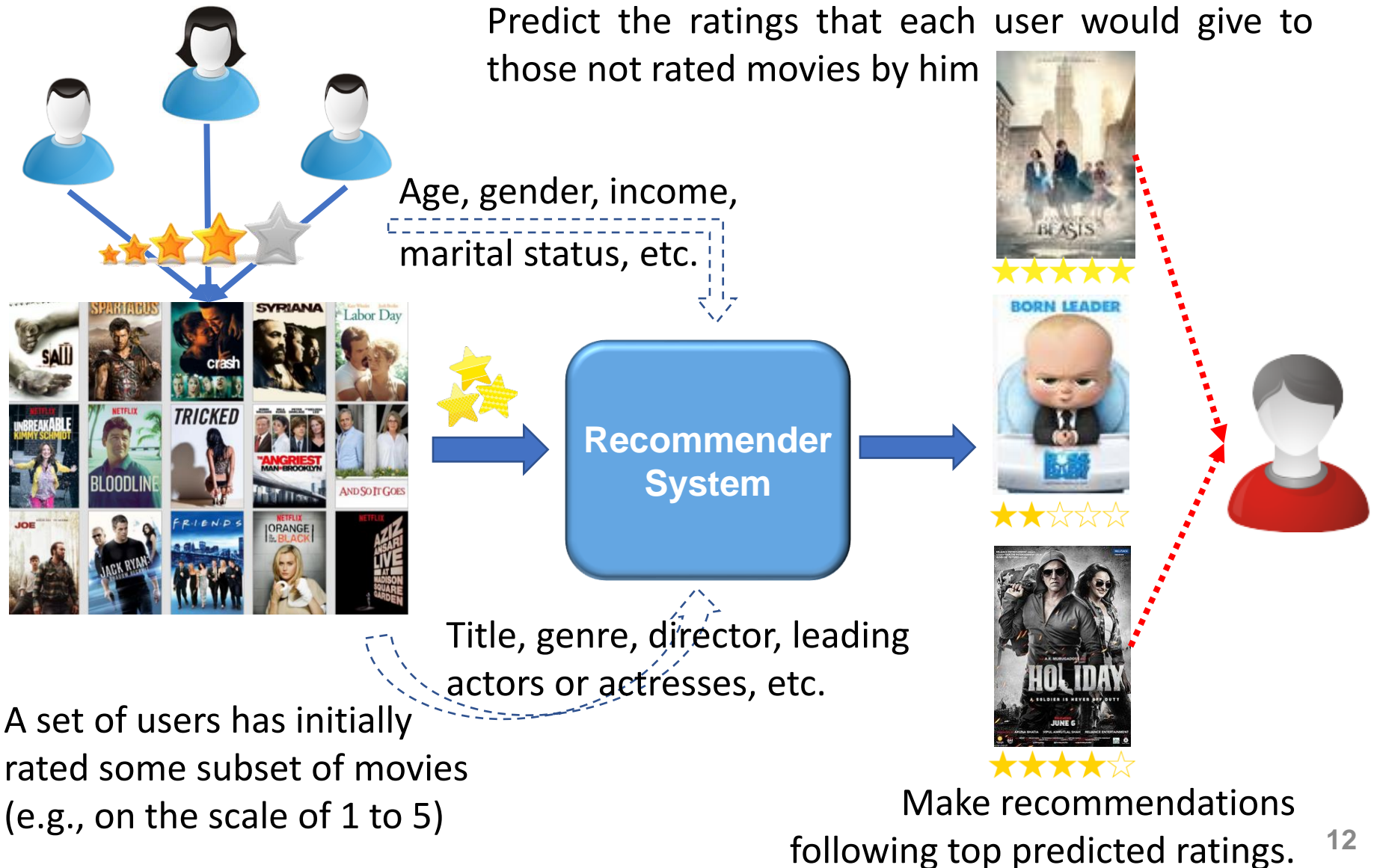
Amazon serves **600 million items** across 33 main categories and other services. There are over **180.1 million Prime subscribers** (2024).

Vast number of users and items

IMDb has roughly **22 million titles** (including episodes) and **14 million personalities**, and **192 million registered users**. (2024)



Movie recommendation





Recommendation problem

Users and items

- Let U be the set of users and S be the set of items to be recommended to the users.
- Each user $u \in U$ is defined with a **user profile** including various user characteristics.
 - E.g., UserID, age, gender, marital status, income, preferences, needs, usage behaviors, etc.
- Each item $s \in S$ is characterized by a set of features.
 - E.g., movie: FilmID, title, genre, director, year of release, actors, etc.
- In most applications, the spaces of U and S can be **huge**.

Users and items


+ Harry Potter and the Sorcerer's Stone (2001)

PG | 2h 32min Adventure , Family , Fantasy 16 November 2001 (USA)


★ 7.6 /10
540,052

☆ Rate This

Genres



Save



Item Characteristics

An orphaned boy enrolls in a school of wizardry, where he learns the truth about himself, his family and the terrible evil that haunts the magical world.

Director: Chris Columbus

Writers: J.K. Rowling (novel), Steve Kloves (screenplay)

Stars: Daniel Radcliffe, Rupert Grint, Richard Harris | See full cast & crew »

The recommendation problem

- Let $p : U \times S \rightarrow R$ be a utility function that measures the usefulness of item s to user u .
 - R is a totally ordered set (e.g., non-negative integers or real numbers within a certain range).
- Learn the utility function p
 - The objective function can be arbitrary and application-dependent, e.g., user satisfaction or seller profitability.
- Use p to predict the utility value of each item s ($\in S$) to each user u ($\in U$) and recommend the top- k items to user u
 - Except items that already have utility values for u from the input data

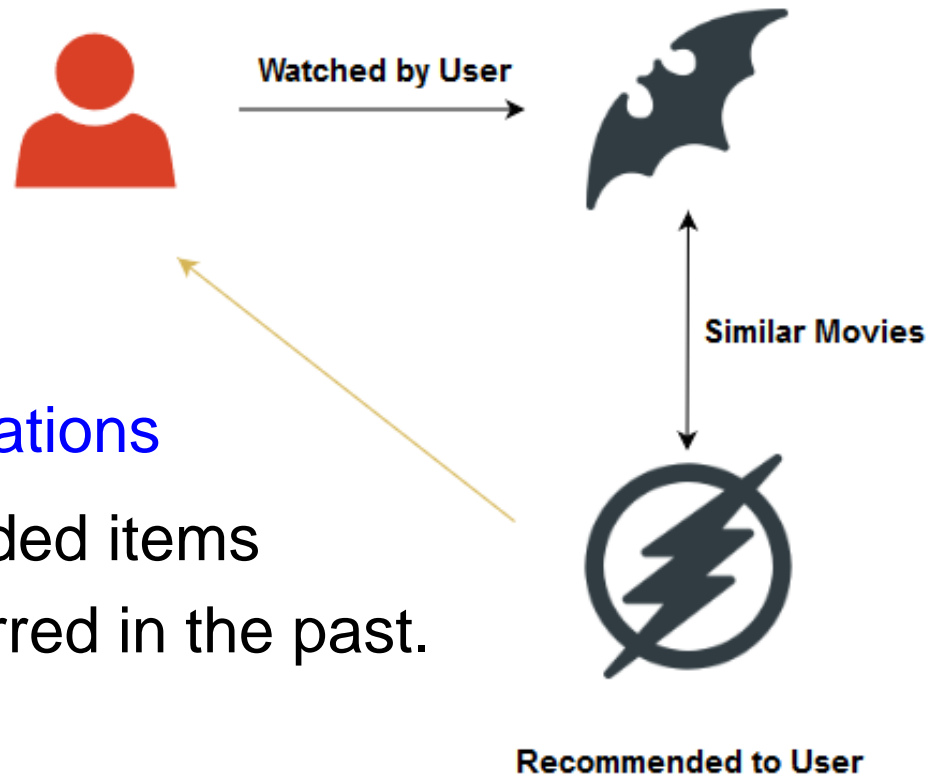
The recommendation problem

- **Rating prediction:** forecast the rating score a user is likely to give to an item that he has not seen or used before
 - The utility of item s to user u is the rating given to item s by user u .
- **Item prediction:** conclude a ranked list of items that a user is likely to buy or use
 - The utility of item s to user u is typically expressed as the probability that user u will buy or use item s .
- User-item interactions are typically binary or multi-scaled (yet concrete values are unconcerned).

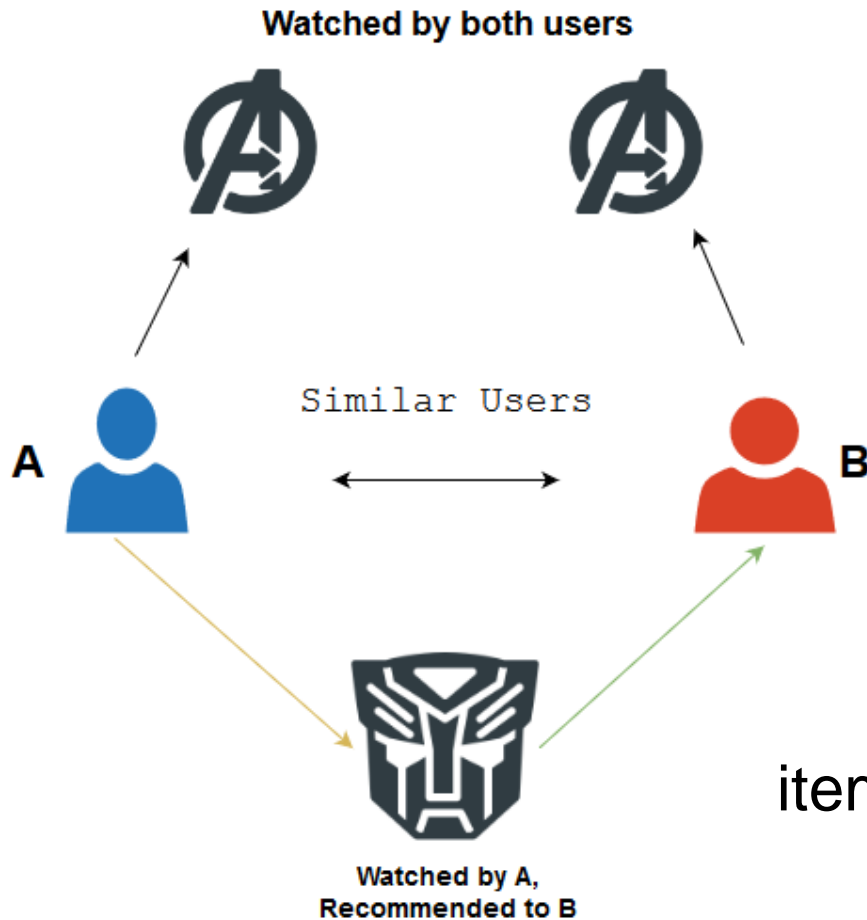
Approaches to recommendation

Content-based recommendations

The user will be recommended items similar to the ones he preferred in the past.



Approaches to recommendation

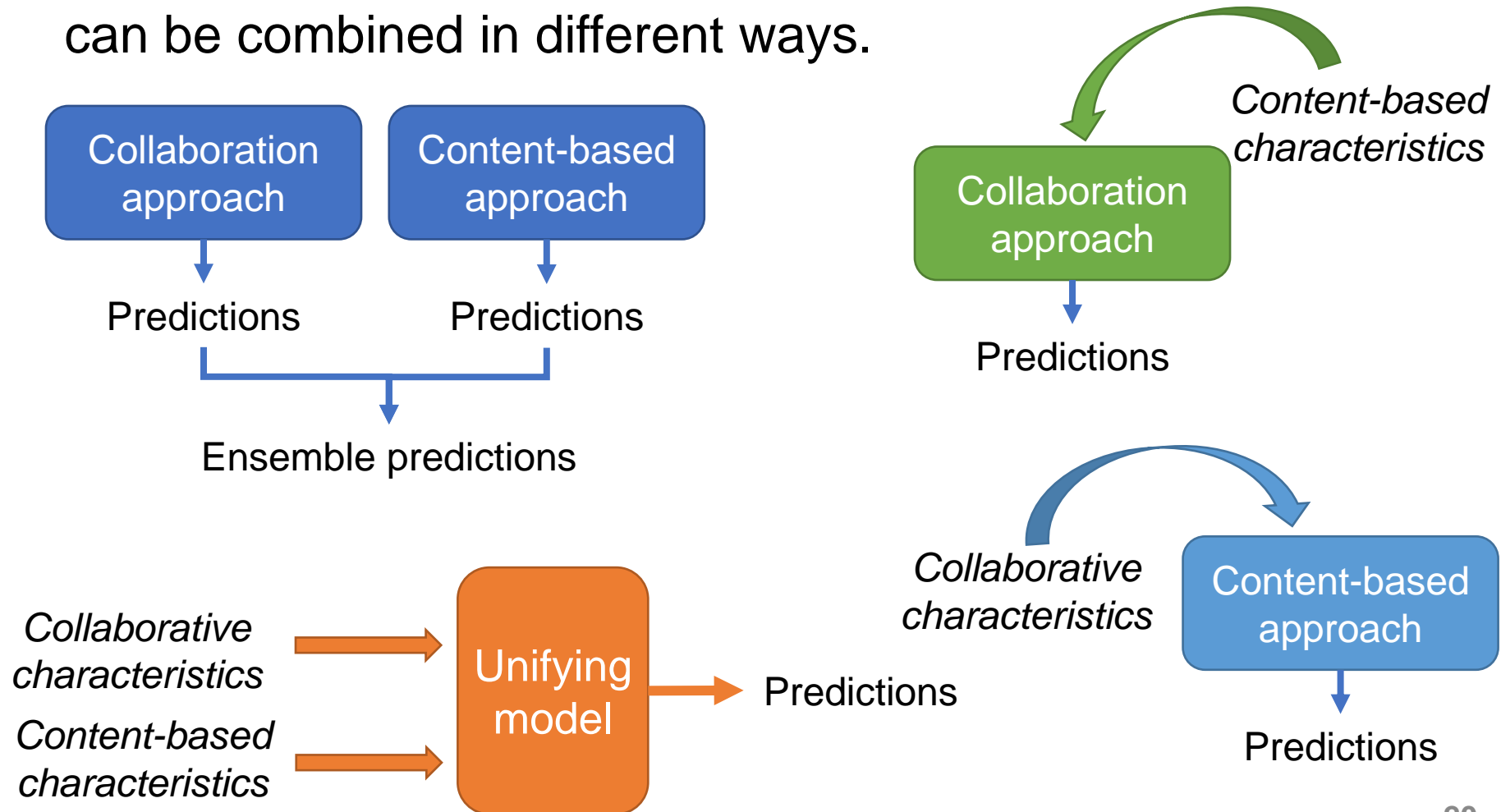


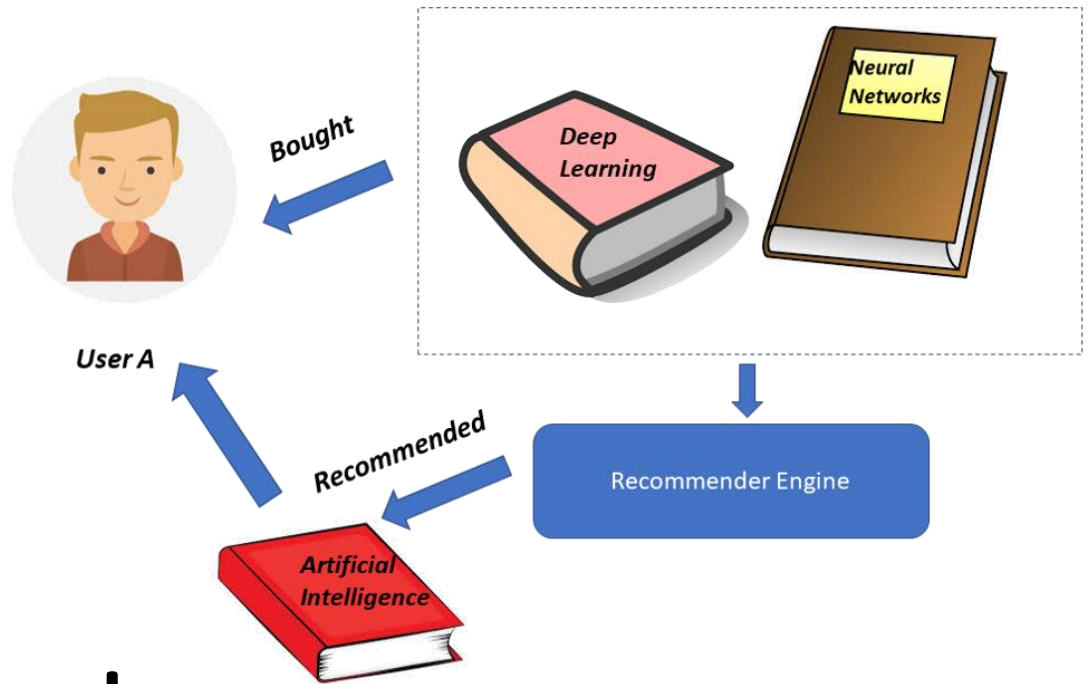
Collaborative filtering

The user will be recommended items that people with similar tastes liked previously

Hybrid approaches

- Collaborative filtering and content-based recommendation can be combined in different ways.

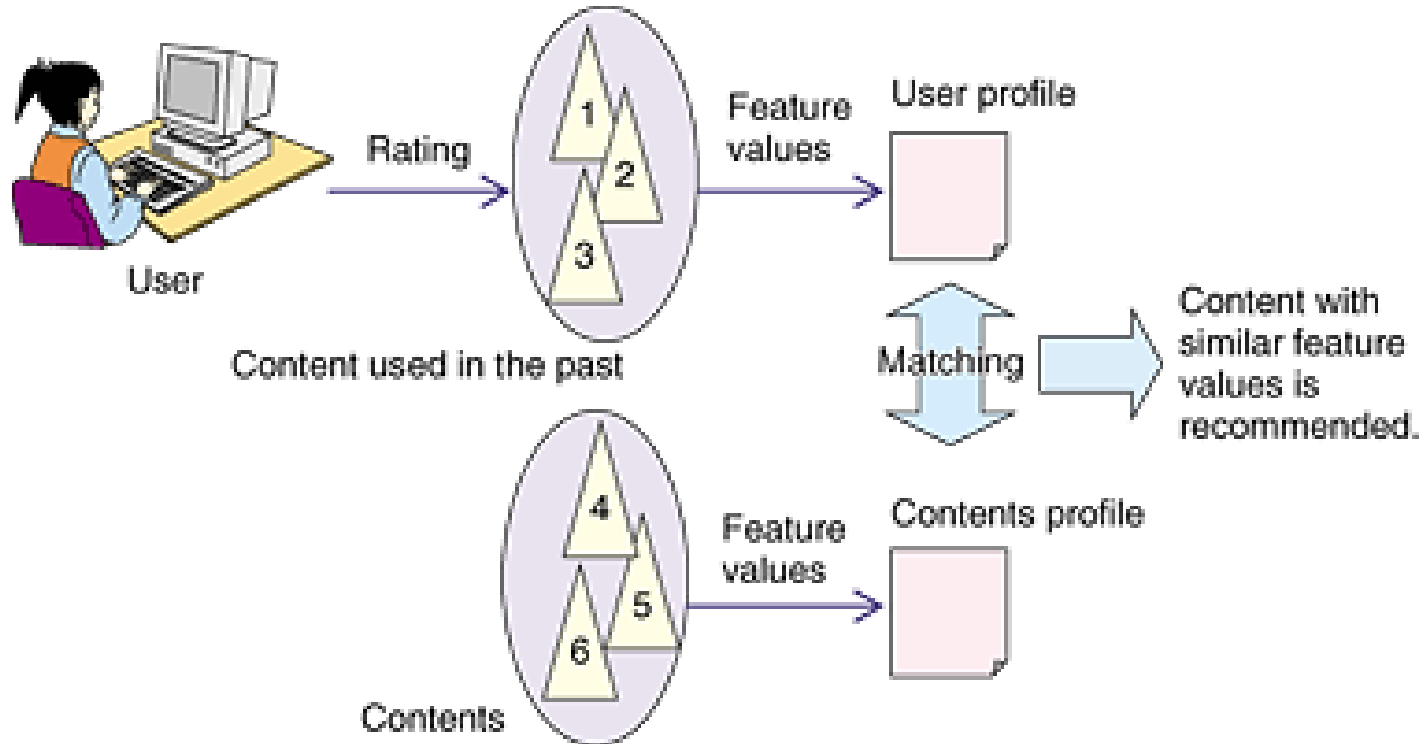




Content-based recommendation

Content-based recommendation

- Predict the utility of an item for a particular user based on how “similar” the item is to those that he liked in the past

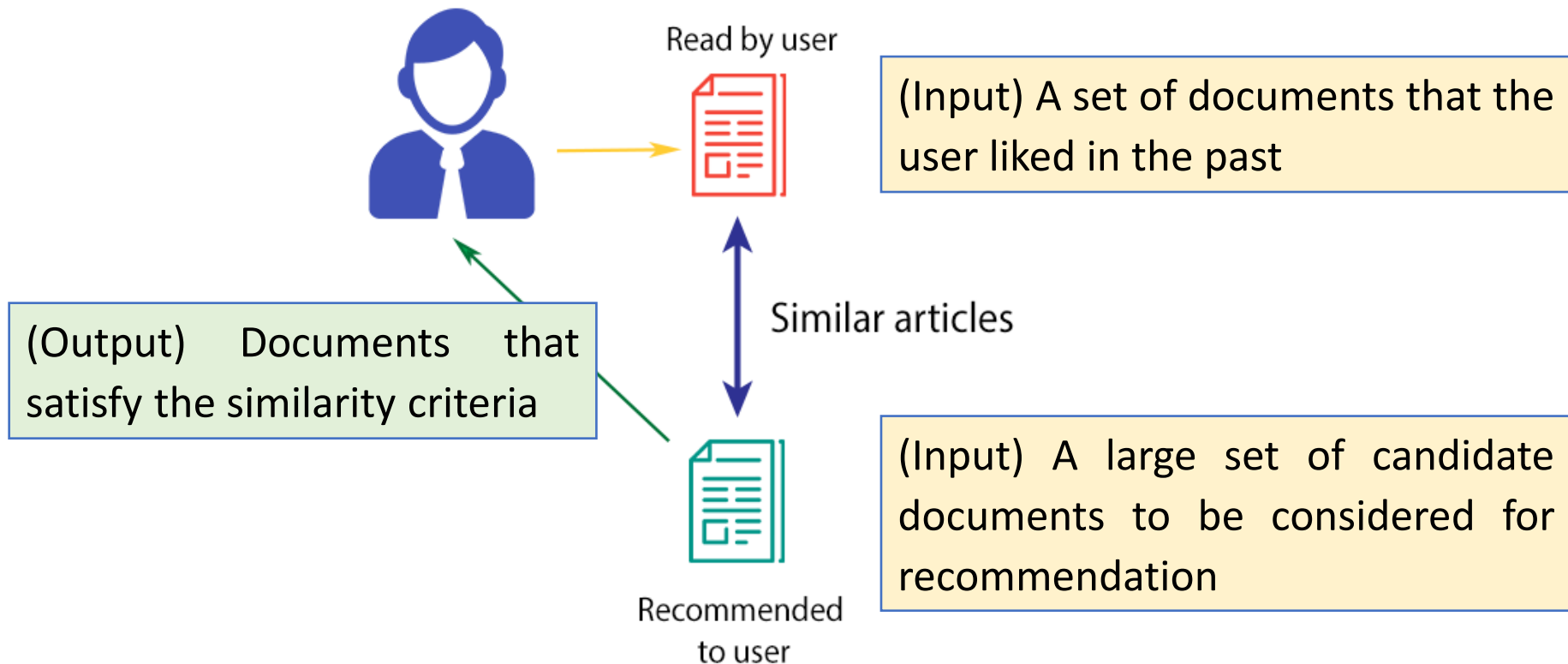


Users and items

- An item is usually represented by a set of features.
 - E.g., a movie – actors, director, genre, subject matter, etc.
- A user profile is also defined by the **same set of features**.
 - This profile can be **learnt explicitly** from the user (e.g., through questionnaires) or **implicitly** from his usage behavior over time.
- **Top- k similar items** are presented to the user after **matching his profile and candidate items** on the **same set of features**.

Text document recommendation

- Content-based recommendation is primarily applied to the domain of text documents.
 - E.g., recommend news articles, Web pages, and blogs

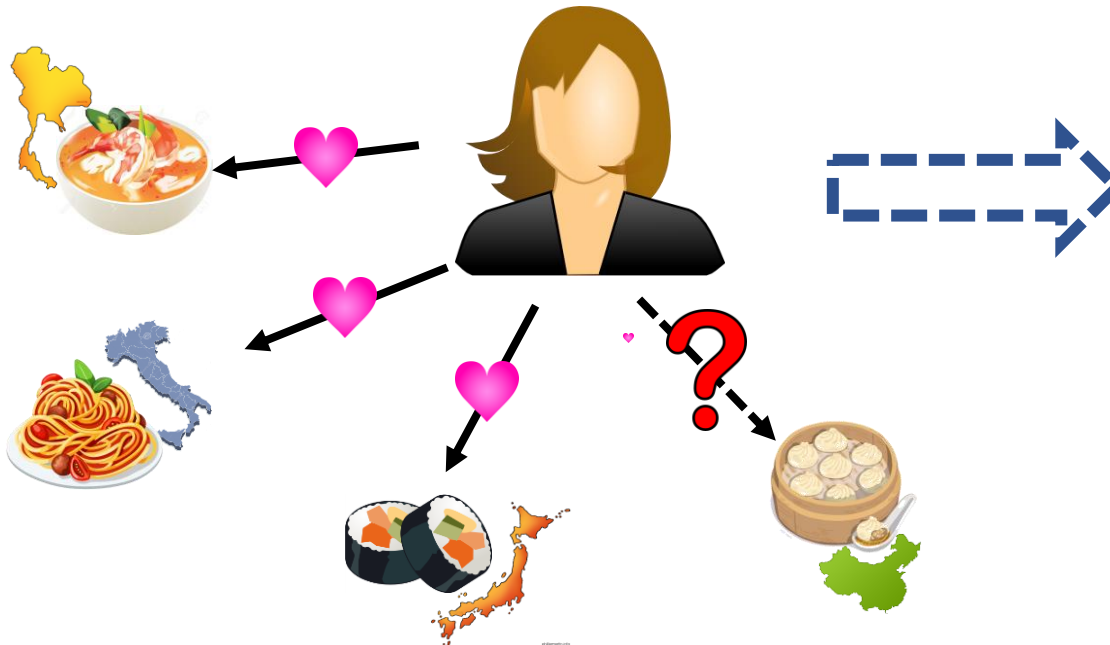


Text document recommendation

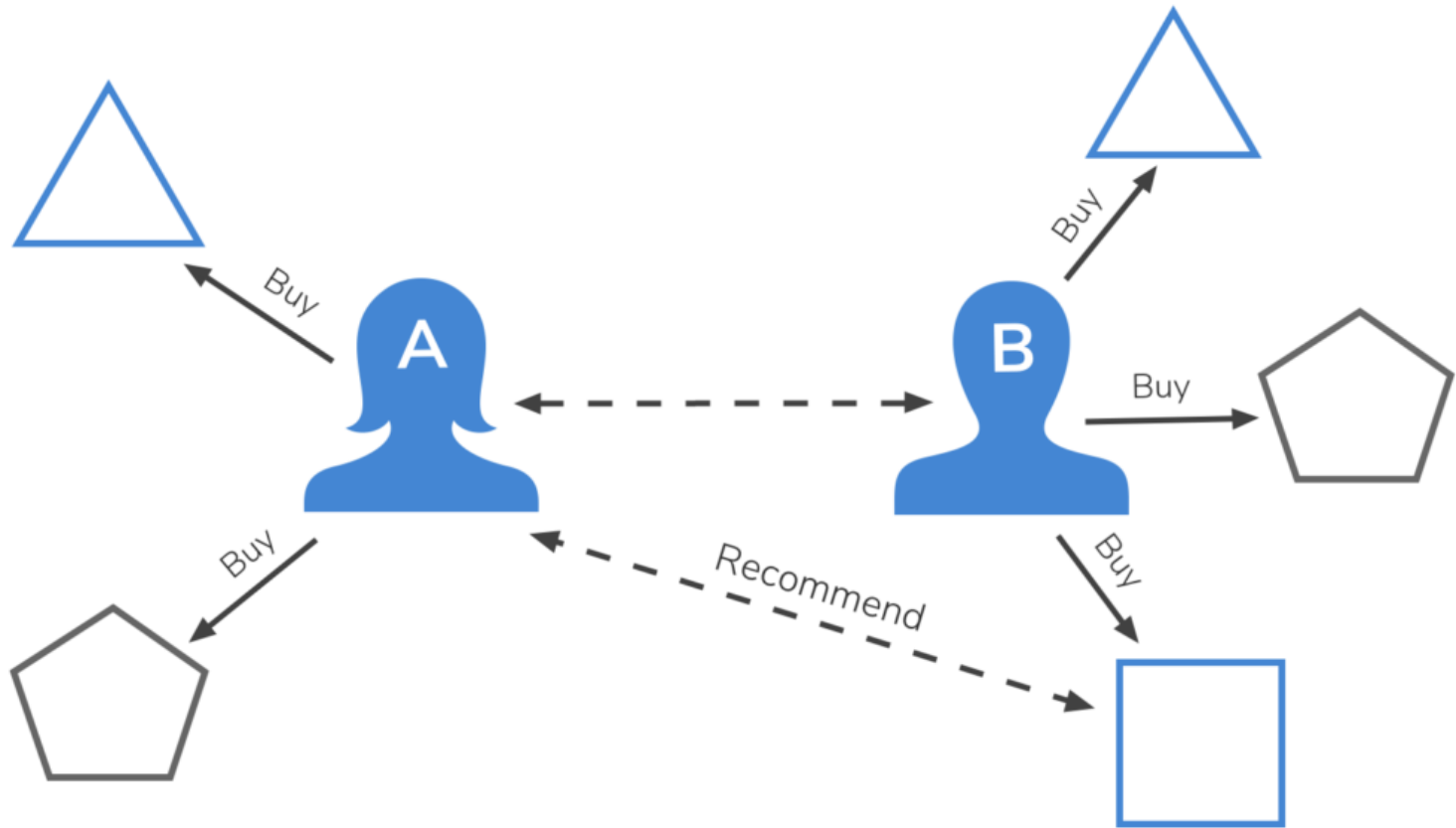
- Most techniques are based on those in information retrieval.
- We define each document by a feature vector and each user profile on the same set of features.
 - Document: a vector of keywords using the TF-IDF scheme in the vector space model
 - User profile: an “average” / “abstract” vector of relevant documents
- The system uses common similarity metrics to match the user profile with candidate documents.
 - E.g., Euclidean distance, cosine similarity, etc.

Limitations

- Items dissimilar to those the user liked in the past are not recommended by the system.
- The user will never see anything completely novel but could be of interest → the business gains less profit from the user.

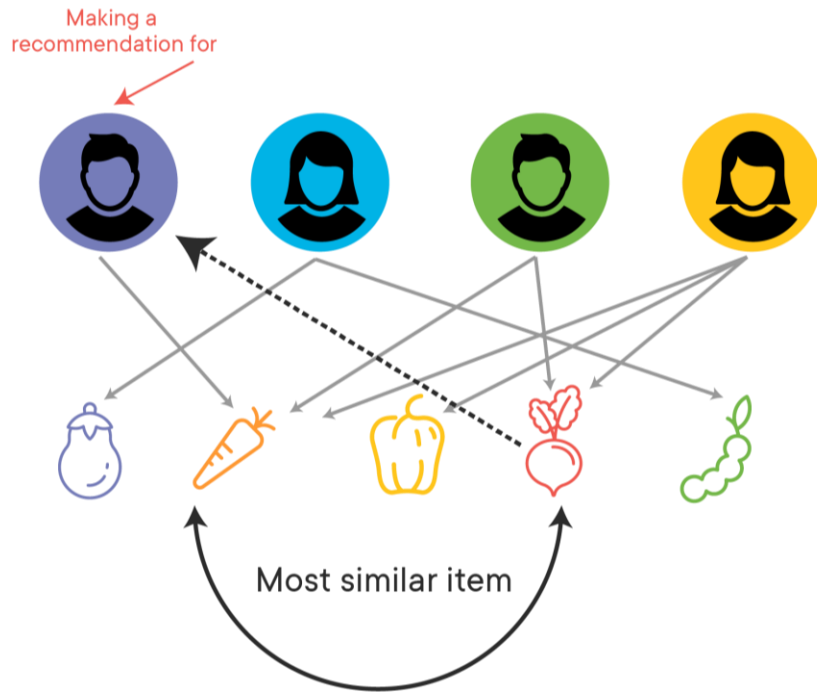


Collaborative Filtering

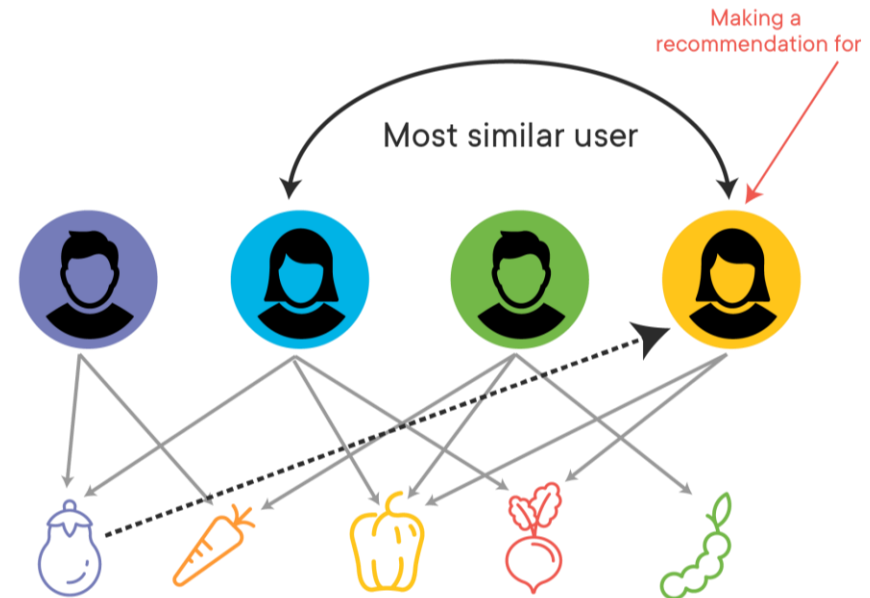


Collaborative filtering (CF)

- CF derives the utility of an item for a particular user from **the items previously considered by other like-minded users.**



Item-based collaborative filtering
(IBCF)

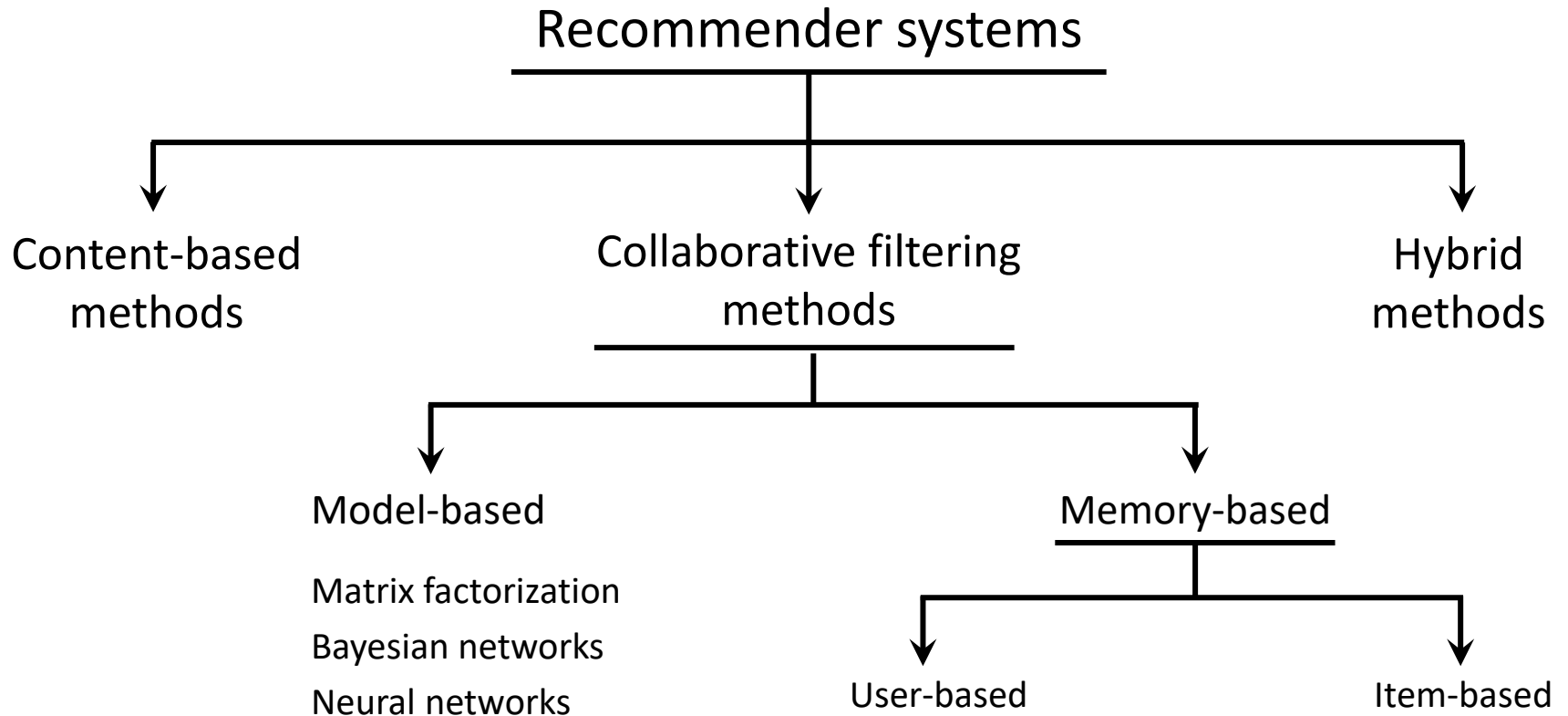


User-based collaborative filtering
(UBCF)

Collaborative filtering types

- A CF system often utilizes only the interaction between consumers and items while ignoring their attributes.
- **User-based collaborative filtering (UBCF)**
 - Find **users** that have similar tastes to that of the target user
 - Predict the item that the target user might like based on ratings given to that item by those “similar” users
- **Item-based collaborative filtering (IBCF)**
 - Find **items** that are “similar” to the candidate item, using the ratings given to each pair of items by users who have rated them both
 - Predict the likelihood that the target user might like the candidate item based on the ratings given to those “similar items” by him

Collaborative filtering types



Collaborative filtering types

Memory-based	Model-based
complete input data is required	abstraction (model) that represents input data
does not scale well	scales well
pre-computation not possible	pre-computation possible
relies on similarity metrics between users and items	relies on matrix factorization or machine learning

Feedback types

- Most CF algorithms are based on certain statistical models of user interests built from **user feedbacks**.



Explicit feedbacks

- Preferences given by the user directly to the item, using one or more ordinal / qualitative scales.
 - E.g., users giving ratings to movies
- More precise but more difficult to collect from users



Implicit feedbacks

- Something similar to the ratings a user would assign, generated from observations available to the system
 - E.g., views, clicks, purchases, likes, shares etc.
- Easier to collect though less accurate in reflecting user tastes

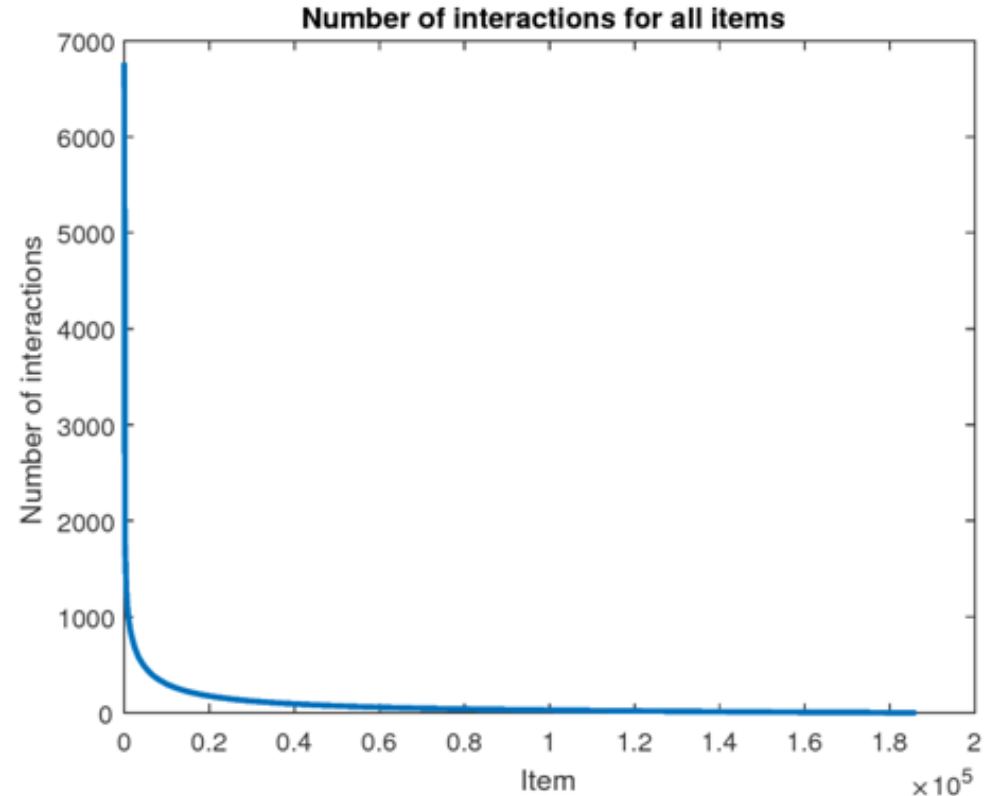
Cold-start problem

- The system cannot draw any inferences for users or items about which it has **not yet gathered sufficient information**.

Number of user interactions associated to each item in the Movielens dataset.

Few items have a very high number of interactions, more than 5000, while most of the others have less than 100.

Image credit: [Wikipedia](#)



Cold-start problem

- New community

- This refers to the start-up of the recommender, when, although a catalogue of items might exist, almost no users are present.

- New item

- A new item is added to the system, it might have some content information, but no interactions are present.

- New user

- A new user registers and has not provided any interaction yet.

Approaches to Collaborative filtering



k-Nearest Neighbors



Association rules



Matrix factorization



Deep learning

Q1. Recommendation strategies

- John wants to build a recommender system for a recently opened online bookstore. The bookstore has over 1 million titles, yet there are only 10,000 ratings. Which strategy will work best in this scenario, *content-based recommendation*, *user-based collaborative filtering*, or *item-based collaboration filtering*? Justify your answer.
- A user has rated 5/5 stars for both titles, “Linear Algebra” and “Differential Equations”. Which of the following titles has least possibility to be recommended by the above system? Justify your answer.
 - a) “Operating Systems”
 - b) “Convex Optimization”
 - c) “Harry Potter: The Goblet of Fire”
 - d) Not sure. It depends on the ratings of other users.

List of references



- Bing Liu. 2007. *Web Data Mining-Exploring Hyperlinks, Contents, and Usage Data*. Springer Series on Data-Centric Systems and Applications. **Chapter 12.4.**



THE END