NATIONAL RESEARCH
UNIVERSITY

School of Data Analysis and Artificial
Intelligence Department of Computer Science

DATA SCIENCE FOR BUSINESS

Lecture 6. Personalization. Recommender systems

Moscow, May 20, 2022.

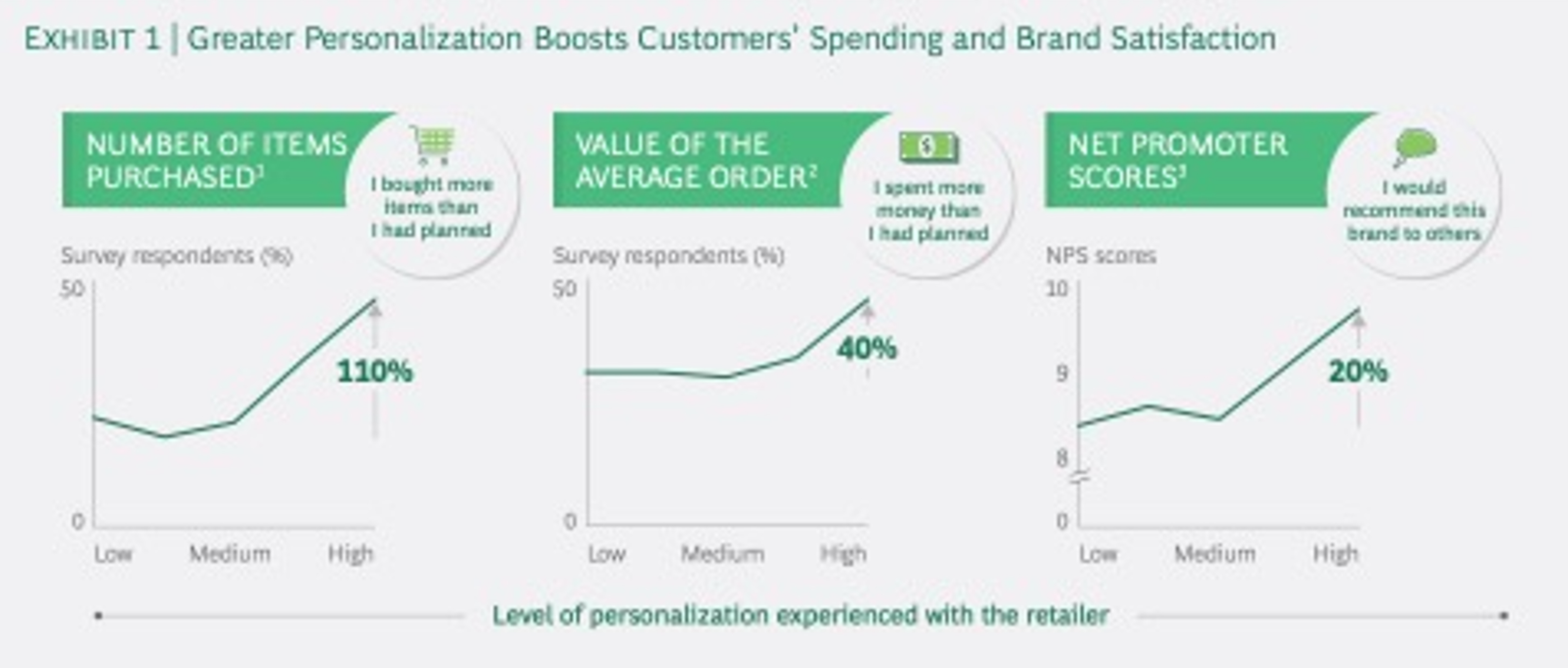
lzhukov@hse.ru

PERSONALIZATION

Personalization - delivering right experience to the right customer at the right time

PERSONALIZATION IN RETAIL

Personalization - continually tailoring the shopping experience to individual customers



*BCG report "Next level of personalization in retail"

MARKETING TECHNOLOGY STACK

*BCG report "The building blocks of personalization"

MARKET BASKET ANALYSIS

Market Basket Analysis is a technique which identifies the strength of association between pairs of products purchased together and creating association rules: (if A => then B)

Usage:

- Cross selling on the items
- Product store placement
- Affinity promotion

Data needed: Point-of-sale data – transactions (items bought in a single purchase)

card_number	store_number	transaction_number	sku	quantity	price	date	hour	month	item_catgry	item_sub_catgry	item_famly_name	item_size	wday
6083024086777210	346941	97729055	1000705	1	2.79	10/7/15	Afternoon	10	Beverages	Iced Coffee	Iced Coffee	Large	Weekday
6083024086777210	346941	97729055	1003180	2	2.18	10/7/15	Afternoon	10	Food - Bakery	Donut Varieties	Traditional Donut	1 Donut	Weekday
6083023014191450	304658	97711690	1001090	1	2.89	12/14/15	Night	12	Beverages	Hot Espresso	Latte	Medium	Weekday
6083023014191450	304658	97711070	1001090	1	2.89	12/1/15	Night	12	Beverages	Hot Espresso	Latte	Medium	Weekday
6083023014191450	304658	97711070	1002829	1	0.99	12/1/15	Night	12	Food - Bakery	Muffin	Muffin	1 Muffin	Weekday
6045176722097870	310010	97708983	1000704	1	2.39	10/12/15	Morning	10	Beverages	Iced Coffee	Iced Coffee	Medium	Weekday
6083017633636840	341519	97707246	1000704	1	1.49	9/8/15	Night	9	Beverages	Iced Coffee	Iced Coffee	Medium	Weekday
6076782313502050	351119	97689200	1000701	1	0.99	1/14/16	Lunch	1	Beverages	Hot Coffee	Hot Coffee	Large	Weekday

ASSOCIATION RULES

A rule is given by an implication X => Y (if X, then Y)

Support - frequency of itemset in the data

$$\mathrm{supp}(X) = \frac{|\{t \in T; X \subseteq t\}|}{|T|} = P(X)$$

Confidence - how often the rule is true in the data

$$\mathrm{conf}(X \Rightarrow Y) = \mathrm{supp}(X \cup Y)/\mathrm{supp}(X) = P(X, Y) / P(X) = P(Y|X)$$

Lift – how presence of one affects another

$$\mathrm{lift}(X \Rightarrow Y) = \frac{\mathrm{supp}(X \cup Y)}{\mathrm{supp}(X) \times \mathrm{supp}(Y)} = P(X, Y) / P(X)\, P(Y)$$

Example database with 5 transactions and 5 items

transaction ID	milk	bread	butter	beer	diapers
1	1	1	0	0	0
2	0	0	1	0	0
3	0	0	0	1	1
4	1	1	1	0	0
5	0	1	0	0	0

rule: $\{\mathrm{milk}, \mathrm{bread}\} \Rightarrow \{\mathrm{butter}\}$

supp $(\{\mathrm{butter}\}) =$ 2/5 = 0.4

supp $(\{\mathrm{milk}, \mathrm{bread}\}) =$ 2/5 = 0.4

supp $(\{\mathrm{milk}, \mathrm{bread}\}\{\mathrm{butter}\}) =$ 1/5 = 0.2

conf $(\{\mathrm{milk}, \mathrm{bread}\} \Rightarrow \{\mathrm{butter}\}) =$ 0.2/0.4 = 0.5

lift $(\{\mathrm{milk}, \mathrm{bread}\} \Rightarrow \{\mathrm{butter}\}) = \frac{0.2}{0.4 \times 0.4} = 1.25$

School of Data Analysis and Artificial Intelligence
Department of Computer Science

ASSOCIATION RULES MINING

Frequent item sets – Apriori algorithm (Agrawal and Srikant,1994)

Database TDB

Tid	Items
10	A, C, D
20	B, C, E
30	A, B, C, E
40	B, E

1st scan → C_1

Itemset	sup
{A}	2
{B}	3
{C}	3
{D}	1
{E}	3

→ L_1

Itemset	sup
{A}	2
{B}	3
{C}	3
{E}	3

C_2

Itemset
{A, B}
{A, C}
{A, E}
{B, C}
{B, E}
{C, E}

2nd scan → C_2

Itemset	sup
{A, B}	1
{A, C}	2
{A, E}	1
{B, C}	2
{B, E}	3
{C, E}	2

→ L_2

Itemset	sup
{A, C}	2
{B, C}	2
{B, E}	3
{C, E}	2

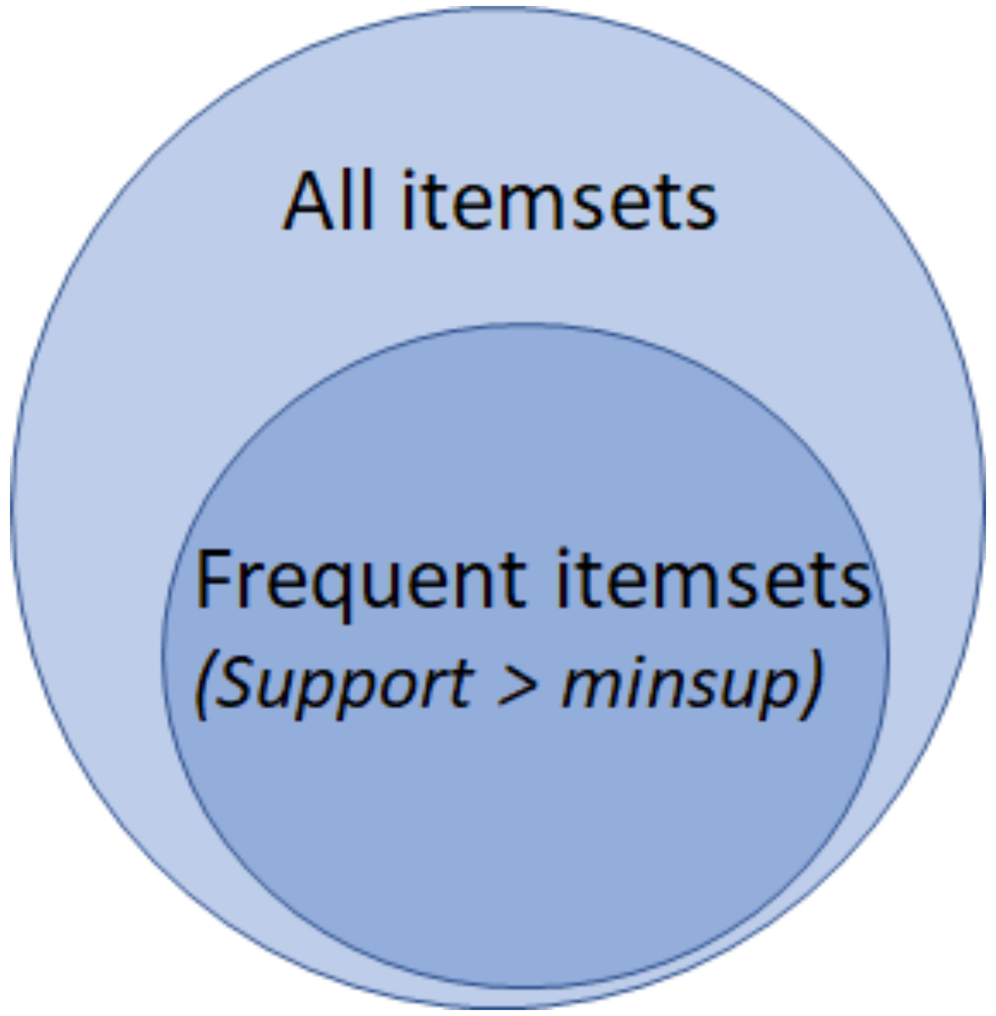
C_3

Itemset
{B, C, E}

3rd scan → L_3

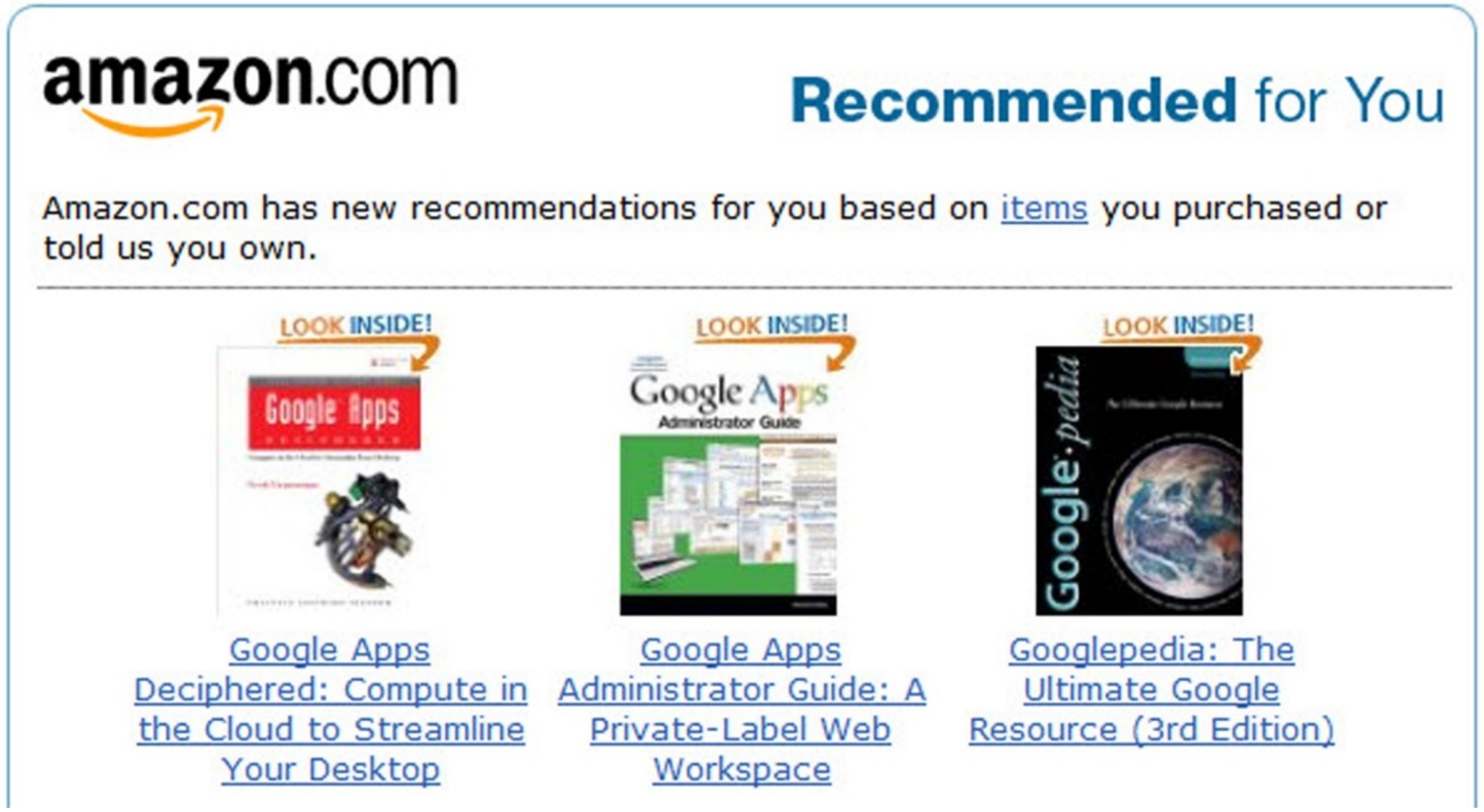
Itemset	sup
{B, C, E}	2

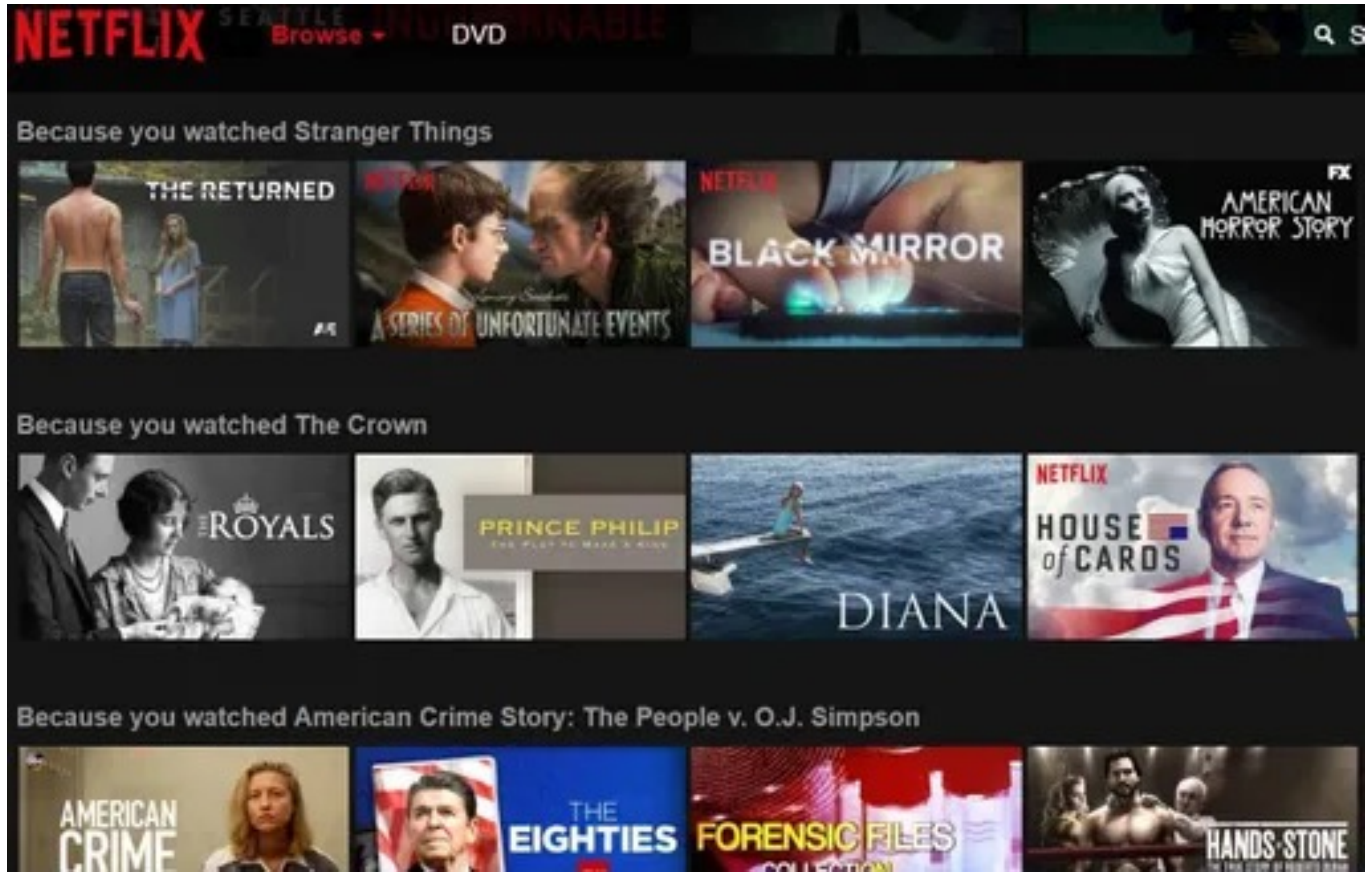
All non-empty subset of frequent itemset must be frequent.

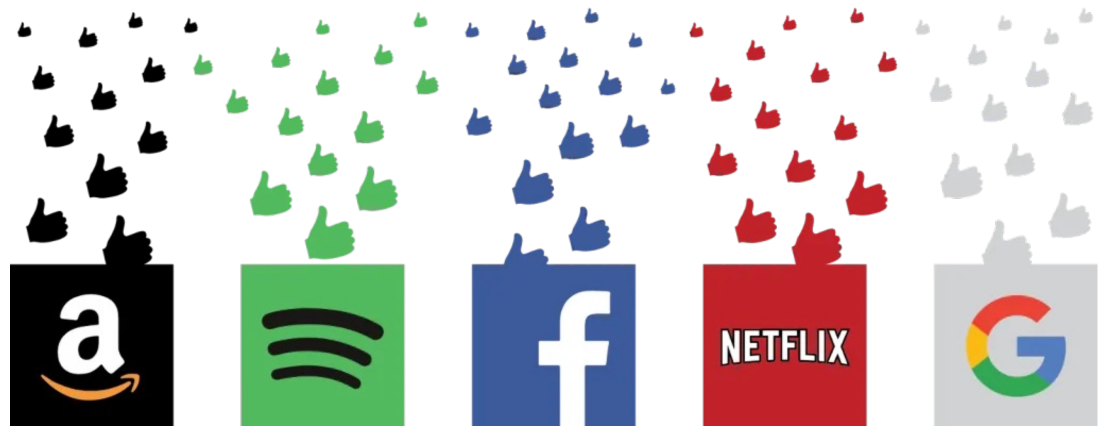


RECOMMENDER SYSTEMS

Recommendation system predicts a rating or preference (buying) a user would give to an item



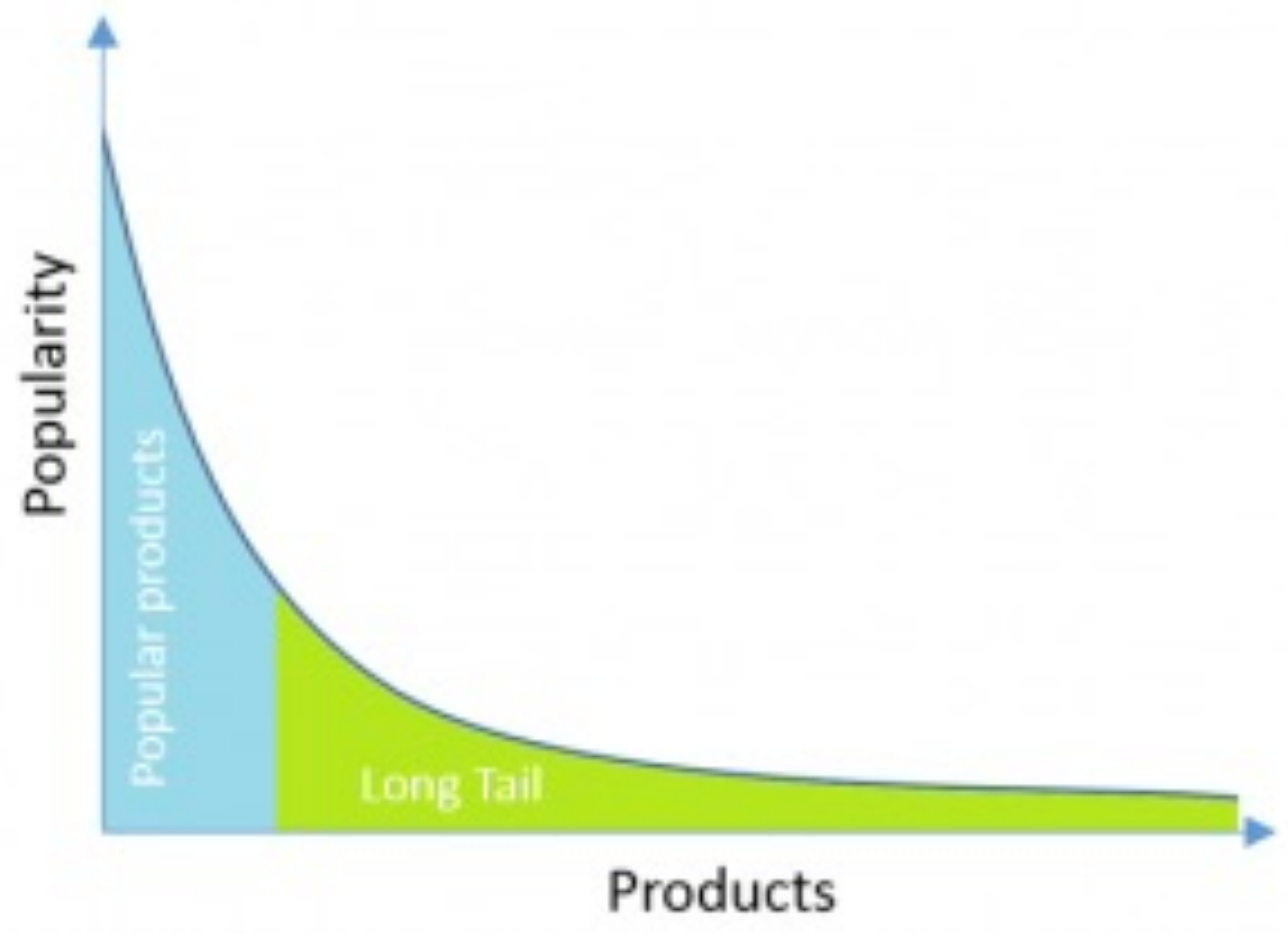




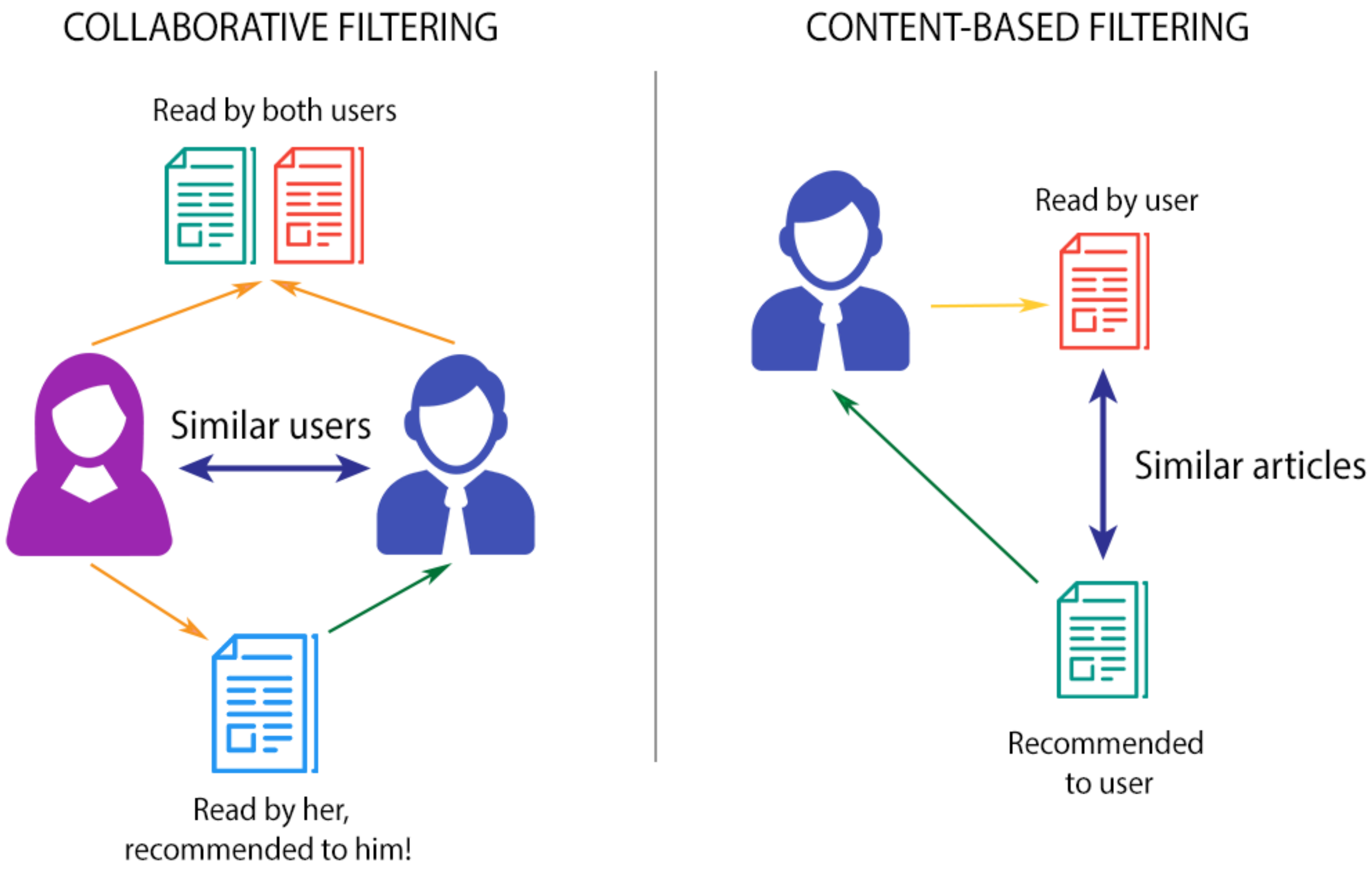
School of Data Analysis and Artificial Intelligence
Department of Computer Science
.

“LONG TALE”

Economics of product popularity – Pareto curve



RECOMMENDATION ENGINES



Industry Report

Amazon.com Recommendations

Item-to-Item Collaborative Filtering

Greg Linden, Brent Smith, and Jeremy York • *Amazon.com*

Recommendation algorithms are best known for their use on e-commerce Web sites,[1] where they use input about a customer's interests to generate a list of recommended items. Many applications use only the items that customers purchase and explicitly rate to represent their interests, but they can also use other attributes, including items viewed, demographic data, subject interests, and favorite artists.

At Amazon.com, we use recommendation algorithms to personalize the online store for each customer. The store radically changes based on customer interests, showing programming titles to a software engineer and baby toys to a new mother. The click-through and conversion rates – two important measures of Web-based and email advertising effectiveness – vastly exceed those of untargeted content such as banner advertisements and top-seller lists.

E-commerce recommendation algorithms often operate in a challenging environment. For example:

- A large retailer might have huge amounts of data, tens of millions of customers and millions of distinct catalog items.
- Many applications require the results set to be returned in realtime, in no more than half a second, while still producing high-quality recommendations.
- New customers typically have extremely limited information, based on only a few purchases or product ratings.
- Older customers can have a glut of information, based on thousands of purchases and ratings.
- Customer data is volatile: Each interaction provides valuable customer data, and the algorithm must respond immediately to new information.

There are three common approaches to solving the recommendation problem: traditional collaborative filtering, cluster models, and search-based methods. Here, we compare these methods with our algorithm, which we call *item-to-item collaborative filtering*. Unlike traditional collaborative filtering, our algorithm's online computation scales independently of the number of customers and number of items in the product catalog. Our algorithm produces recommendations in realtime, scales to massive data sets, and generates high-quality recommendations.

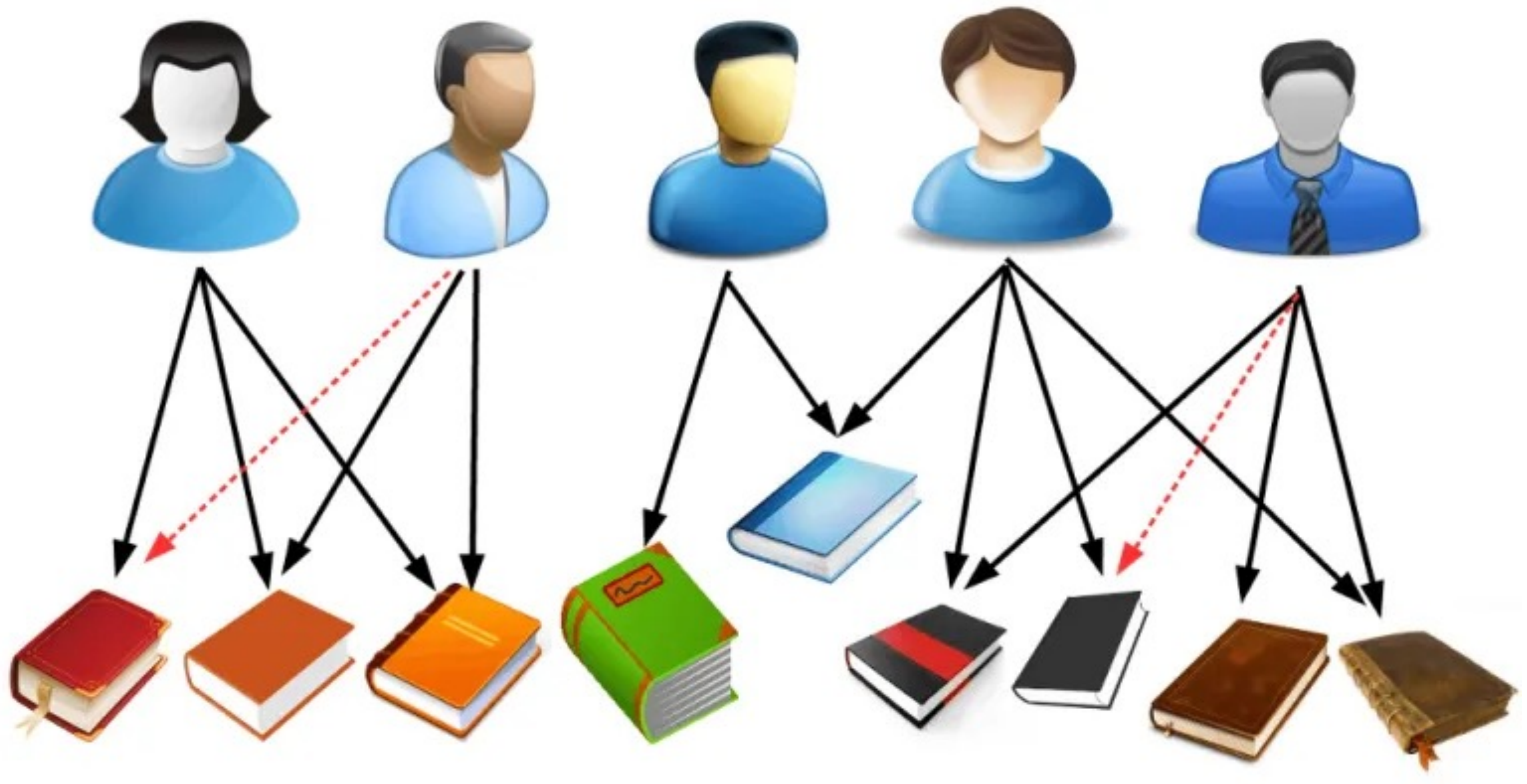
Recommendation Algorithms

Most recommendation algorithms start by finding a set of customers whose purchased and rated items overlap the user's purchased and rated items.[2] The algorithm aggregates items from these similar customers, eliminates items the user has already purchased or rated, and recommends the remaining items to the user. Two popular versions of these algorithms are *collaborative filtering* and *cluster models*. Other algorithms – including search-based methods and our own item-to-item collaborative filtering – focus on finding similar items, not similar customers. For each of the user's purchased and rated items, the algorithm attempts to find similar items. It then aggregates the similar items and recommends them.

Traditional Collaborative Filtering

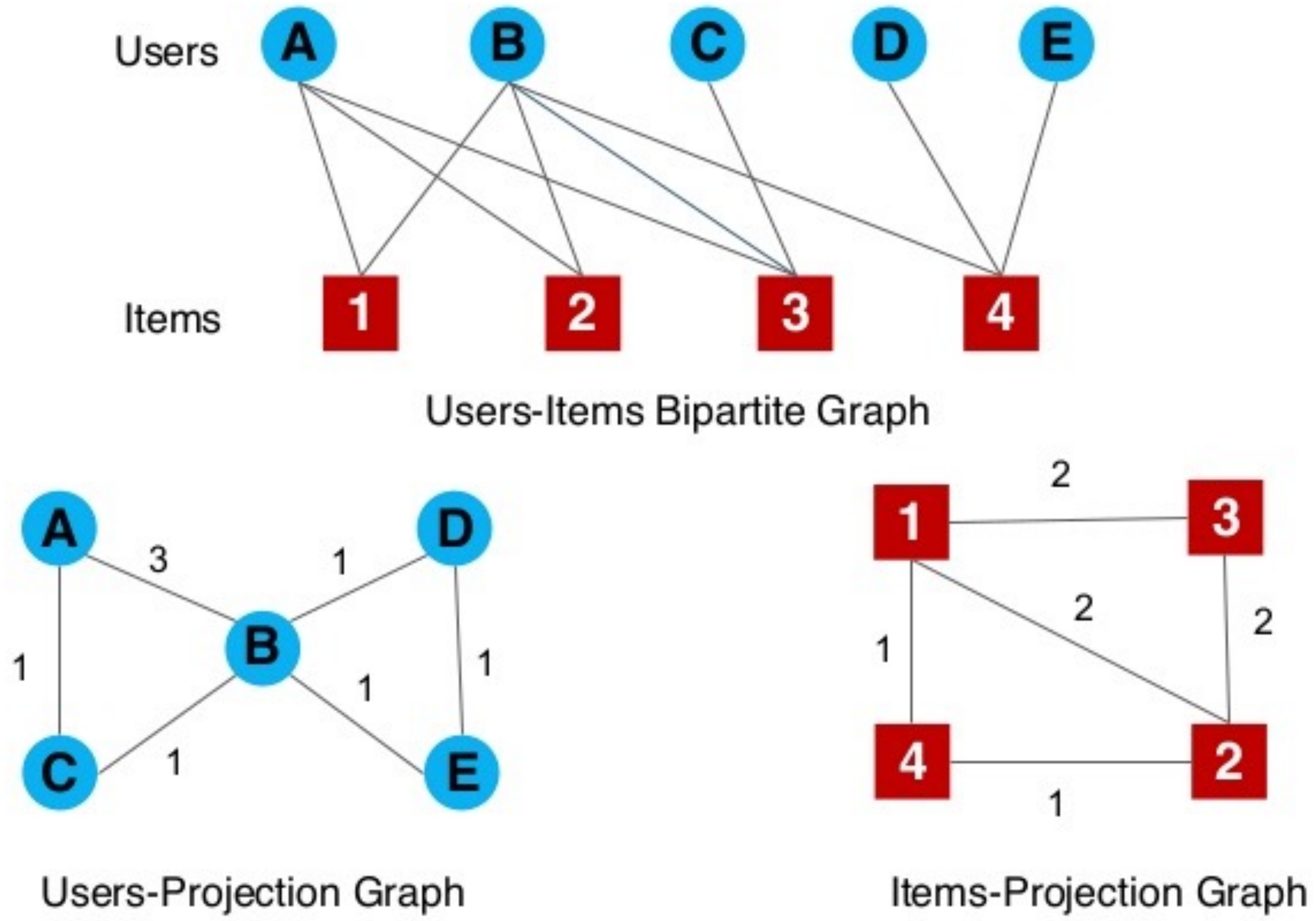
A traditional collaborative filtering algorithm represents a customer as an N-dimensional vector of items, where N is the number of distinct catalog items. The components of the vector are positive for purchased or positively rated items and negative for negatively rated items. To compensate for

COLLABORATIVE FILTERING



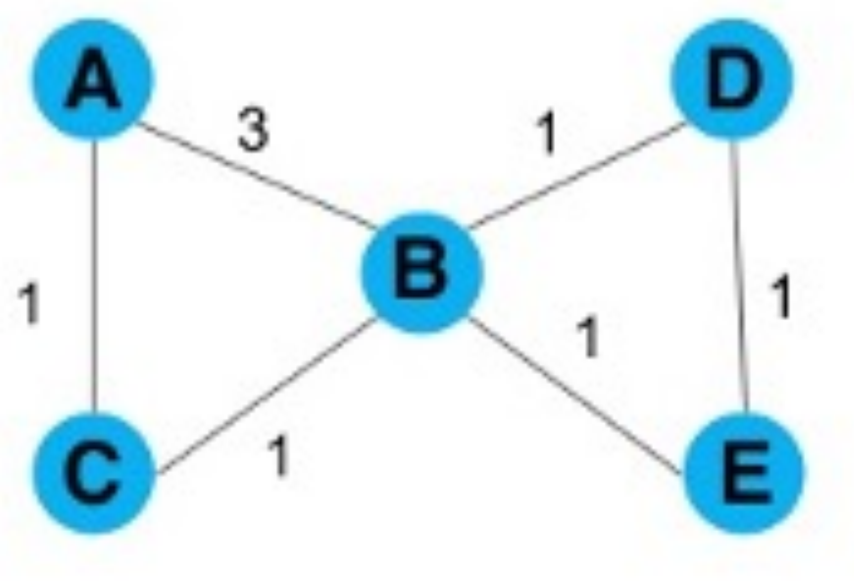
user-item bipartite graph:

- user-user graph
- item-item graph

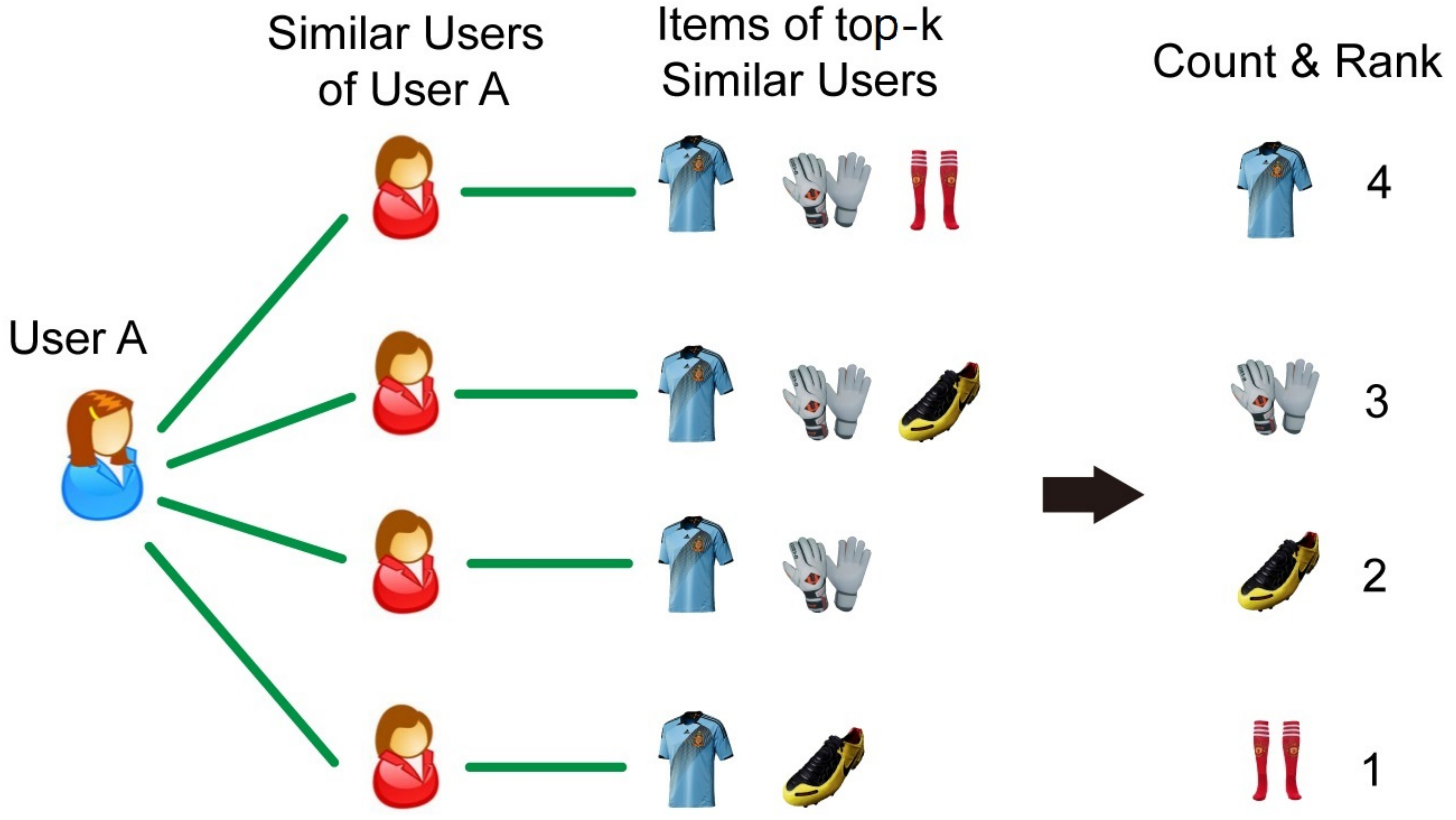


COLLABORATIVE FILTERING

User-based top-N recommendation algorithm (user-user)

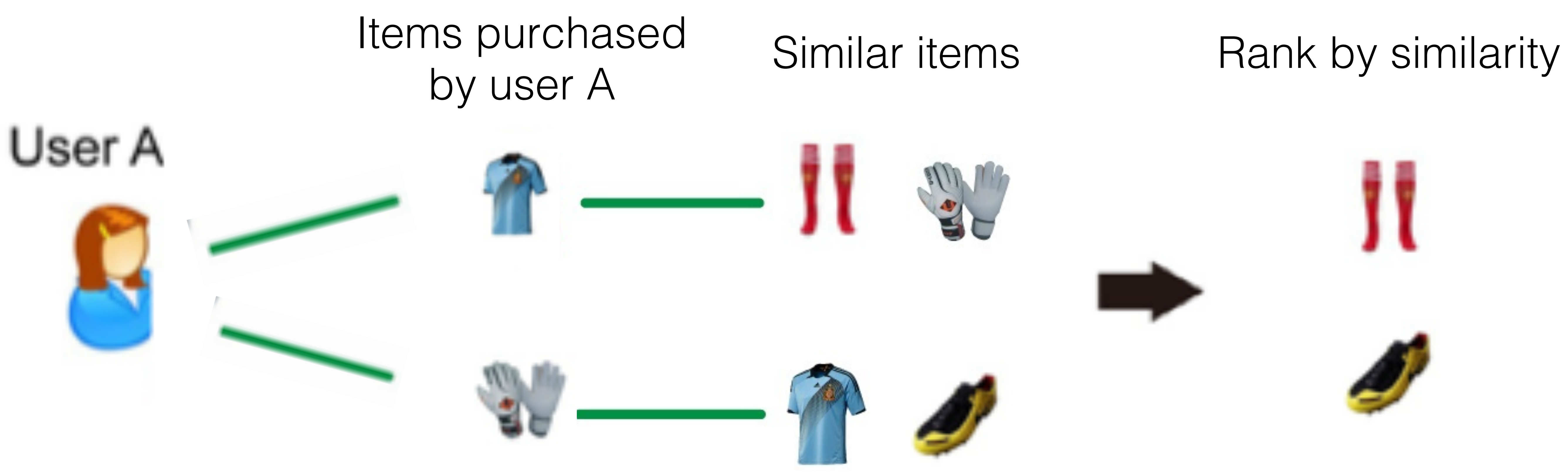


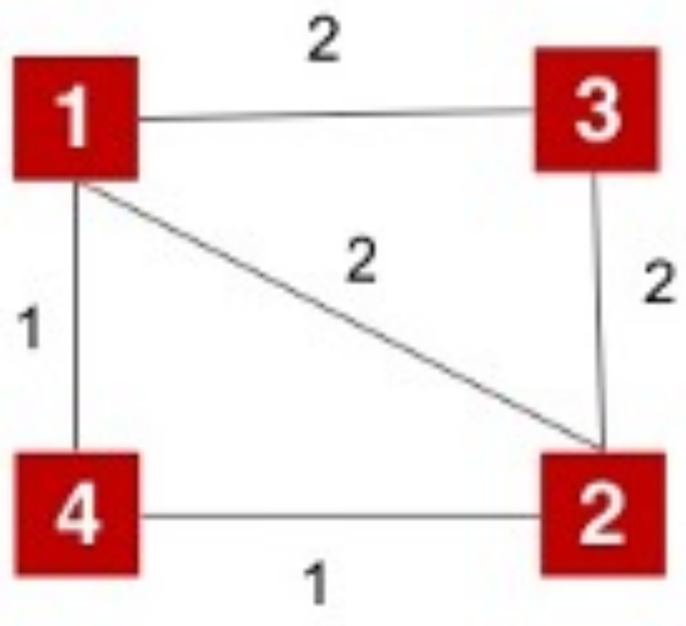
Users-Projection Graph



COLLABORATIVE FILTERING

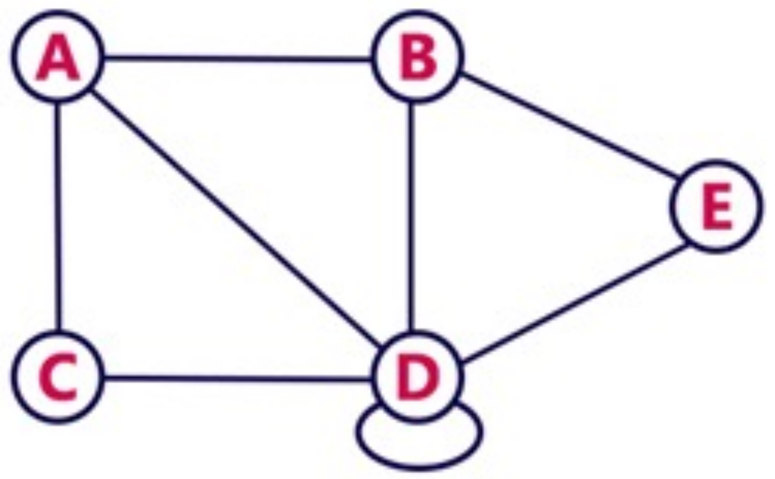
Item-based top-N recommendation algorithm (item-item)



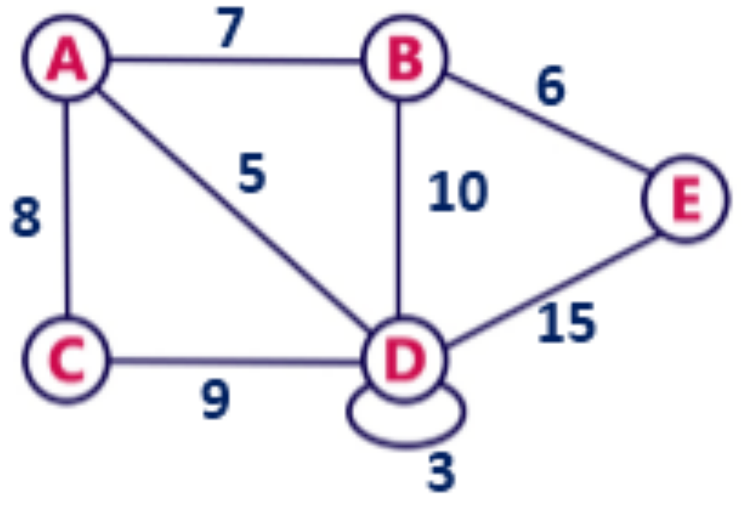


Items-Projection Graph

MATRIX GRAPH REPRESENTATION

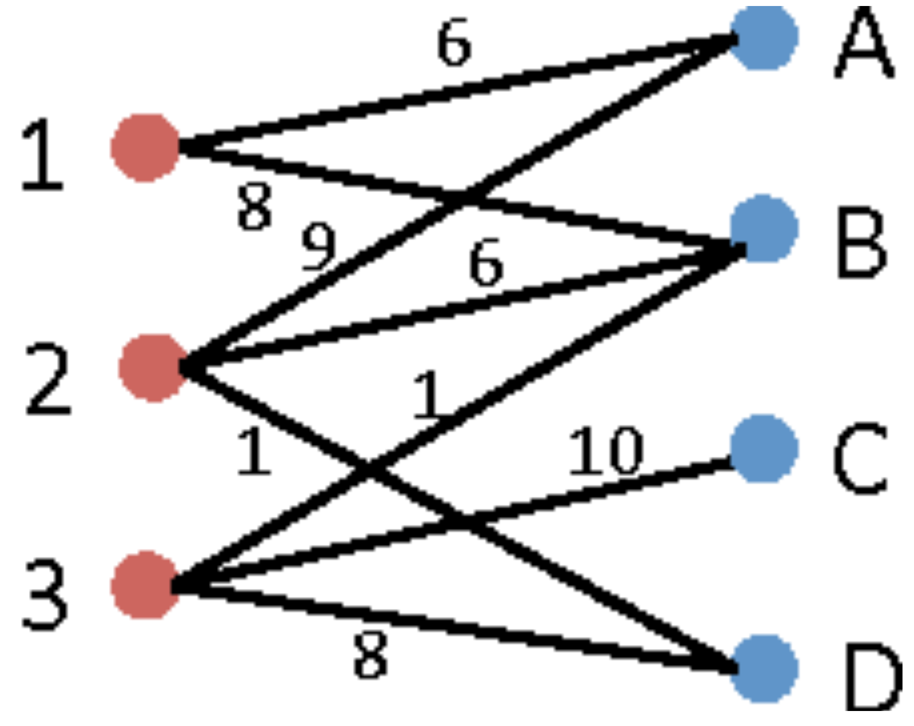


	A	B	C	D	E
A	0	1	1	1	0
B	1	0	0	1	1
C	1	0	0	1	0
D	1	1	1	1	1
E	0	1	0	1	0



	A	B	C	D	E
A	0	7	8	5	0
B	7	0	0	10	6
C	8	0	0	9	0
D	5	10	9	3	15
E	0	6	0	15	0

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 6 & 8 & 0 & 0 \\ 9 & 6 & 0 & 1 \\ 0 & 1 & 10 & 8 \end{bmatrix} \end{array}$$



PREDICTING USER RATINGS

INPUT

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

Introduction to Recommender Systems
Machine learning Paradigms
Social Network-based Recommender Systems
Learning Spark
Recommender Systems Handbook
Recommender Systems and the Social Web

USER-BASED COLLABORATIVE FILTERING

1.00	0.75	0.63	0.22	0.30	0.00
0.75	1.00	0.91	0.00	0.00	0.16
0.63	0.91	1.00	0.00	0.00	0.40
0.22	0.00	0.00	1.00	0.97	0.64
0.30	0.00	0.00	0.97	1.00	0.53
0.00	0.16	0.40	0.64	0.53	1.00

(0.7 x) + (0.6 x) =

already rated by user

(0.7 x 4 + 0.6 x 5) / (0.7 + 0.6) = 4.5

(0.6 x 3) / 0.6 = 3.0

already rated by user

ITEM-BASED COLLABORATIVE FILTERING

1.00	0.27	0.79	0.32	0.98	0.00
0.27	1.00	0.00	0.00	0.34	0.65
0.79	0.00	1.00	0.69	0.71	0.18
0.32	0.00	0.69	1.00	0.32	0.49
0.98	0.34	0.71	0.32	1.00	0.00
0.00	0.65	0.18	0.49	0.00	1.00

(4 x) + (3 x) + (5 x) =

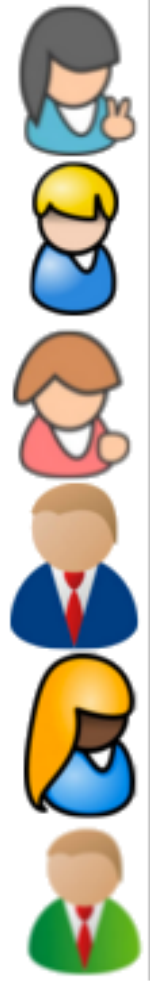
already rated by user

(0.8 x 4 + 0.7 x 5) / (0.8 + 0.7) = 4.5

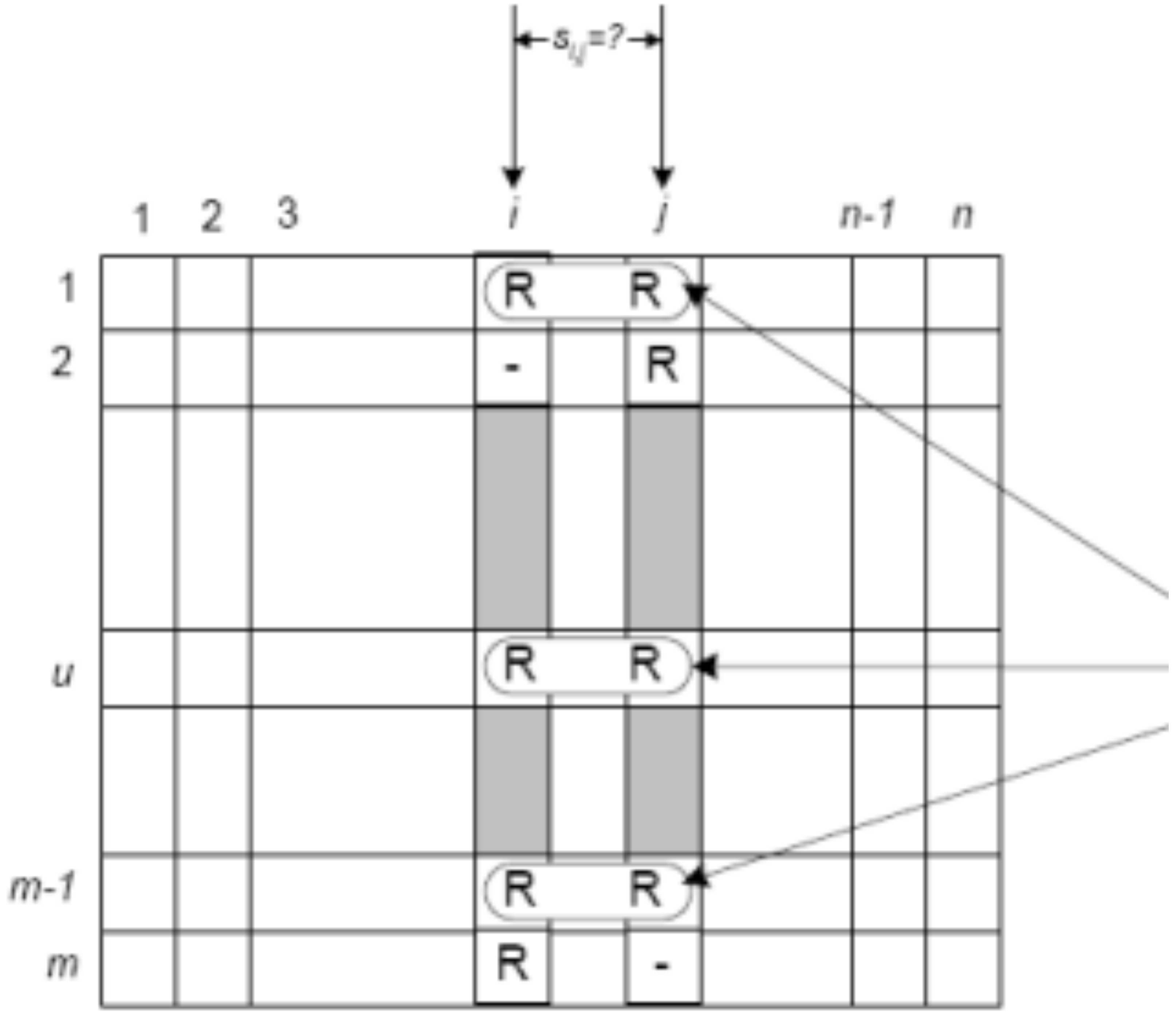
already rated by user

(0.7 x 3) / 0.7 = 3.0

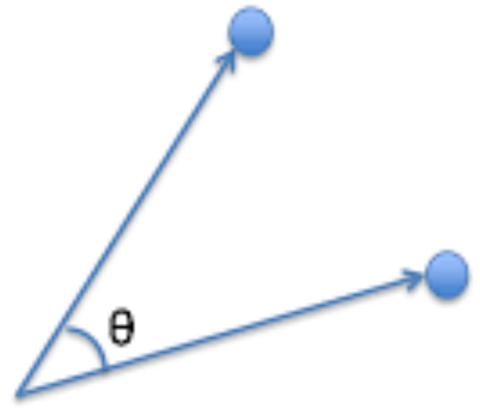
COSINE SIMILARITY



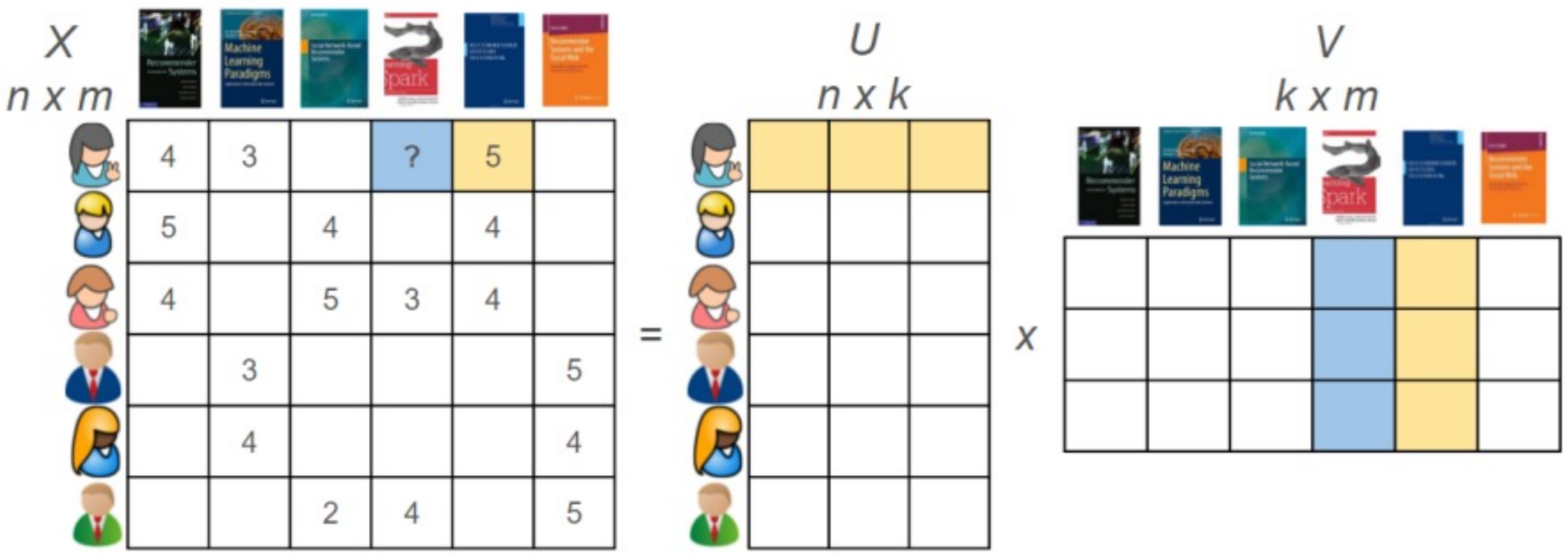
4	3			5	
5		4		4	
4		5	3	4	
	3				5
	4				4
		2	4		5



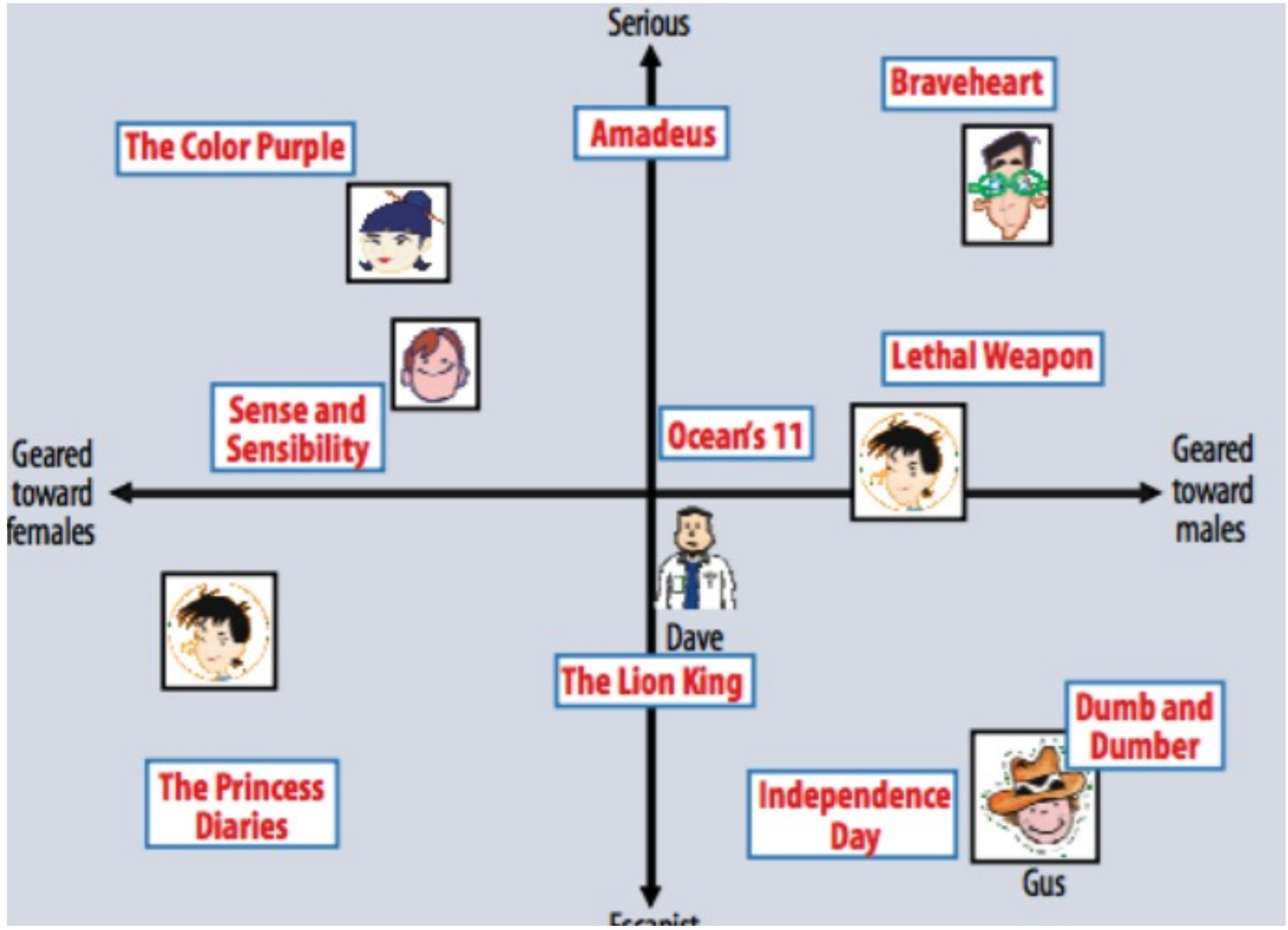
$$sim(A,B) = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$$



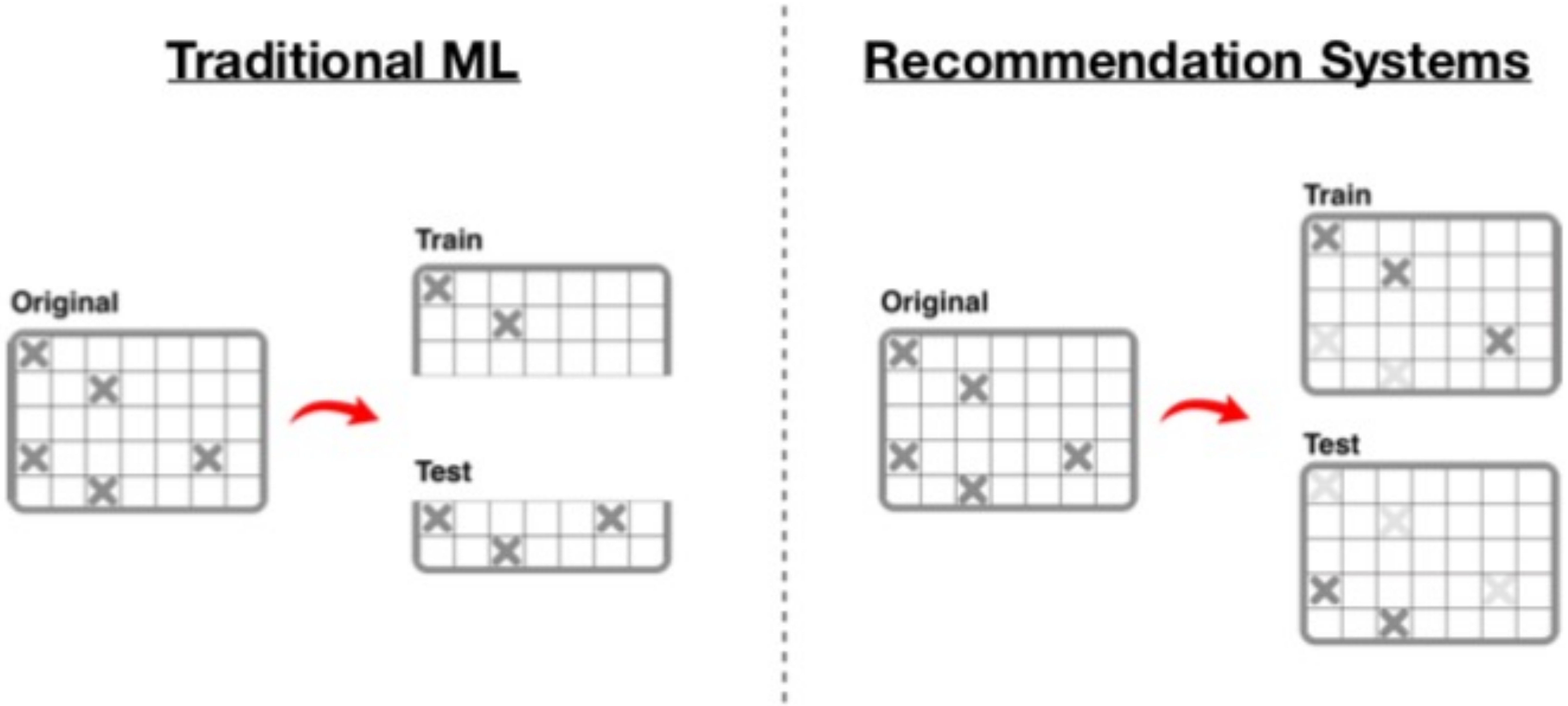
RECOMMENDATION BY MATRIX FACTORIZATION



$$A = U\,S\,V^{T}$$



EVALUATING RECOMMENDER SYSTEM QUALITY



Train-test splitting

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{N}(y - \hat{y})^2}{N}}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

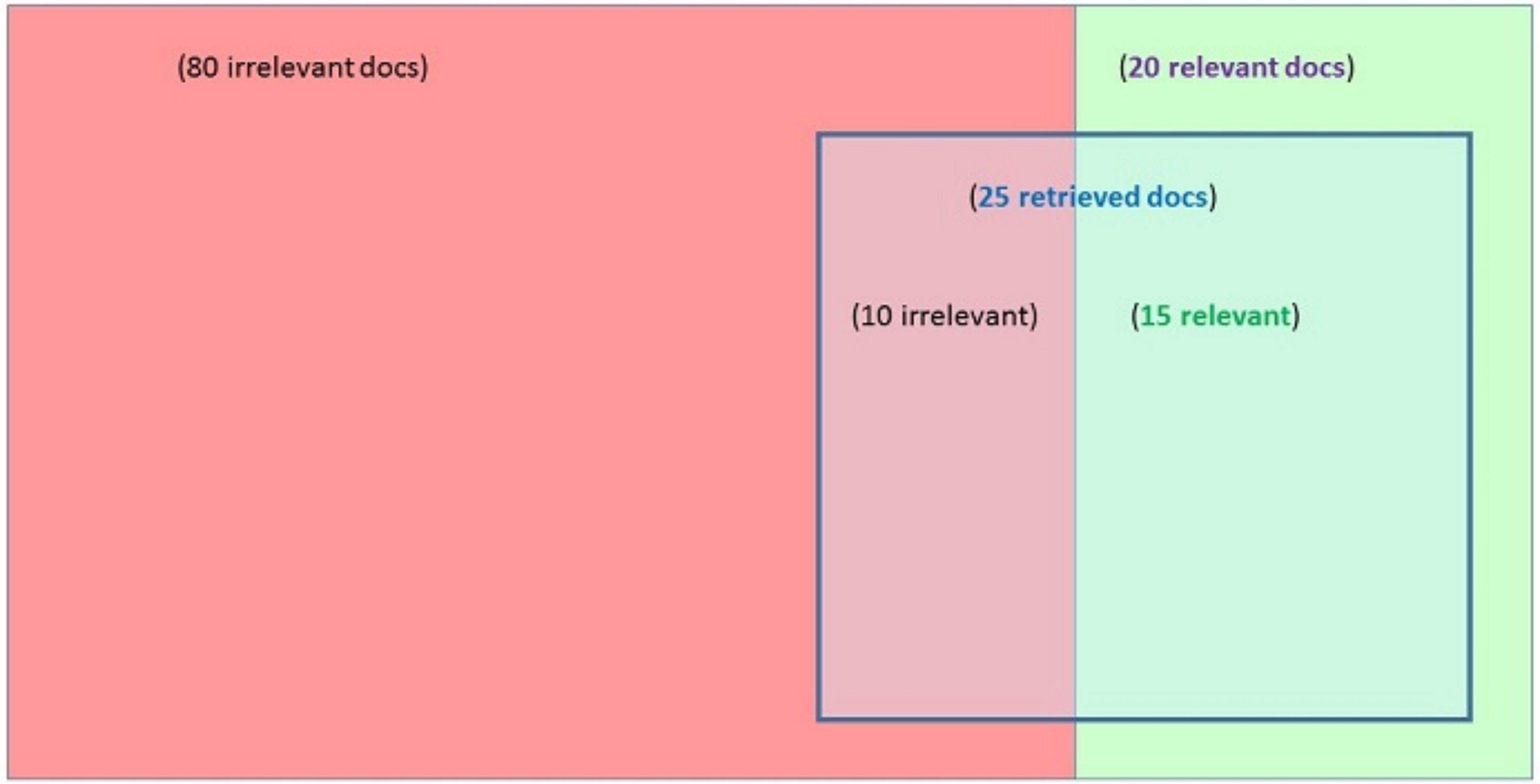
Quality metrics

PRECISION AND RECALL

Information retrieval metrics

$$\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}$$

$$\text{recall} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{relevant documents}\}|}$$



Precision = 15 / 25 = 0.60 Recall = 15 / 20 = 0.75

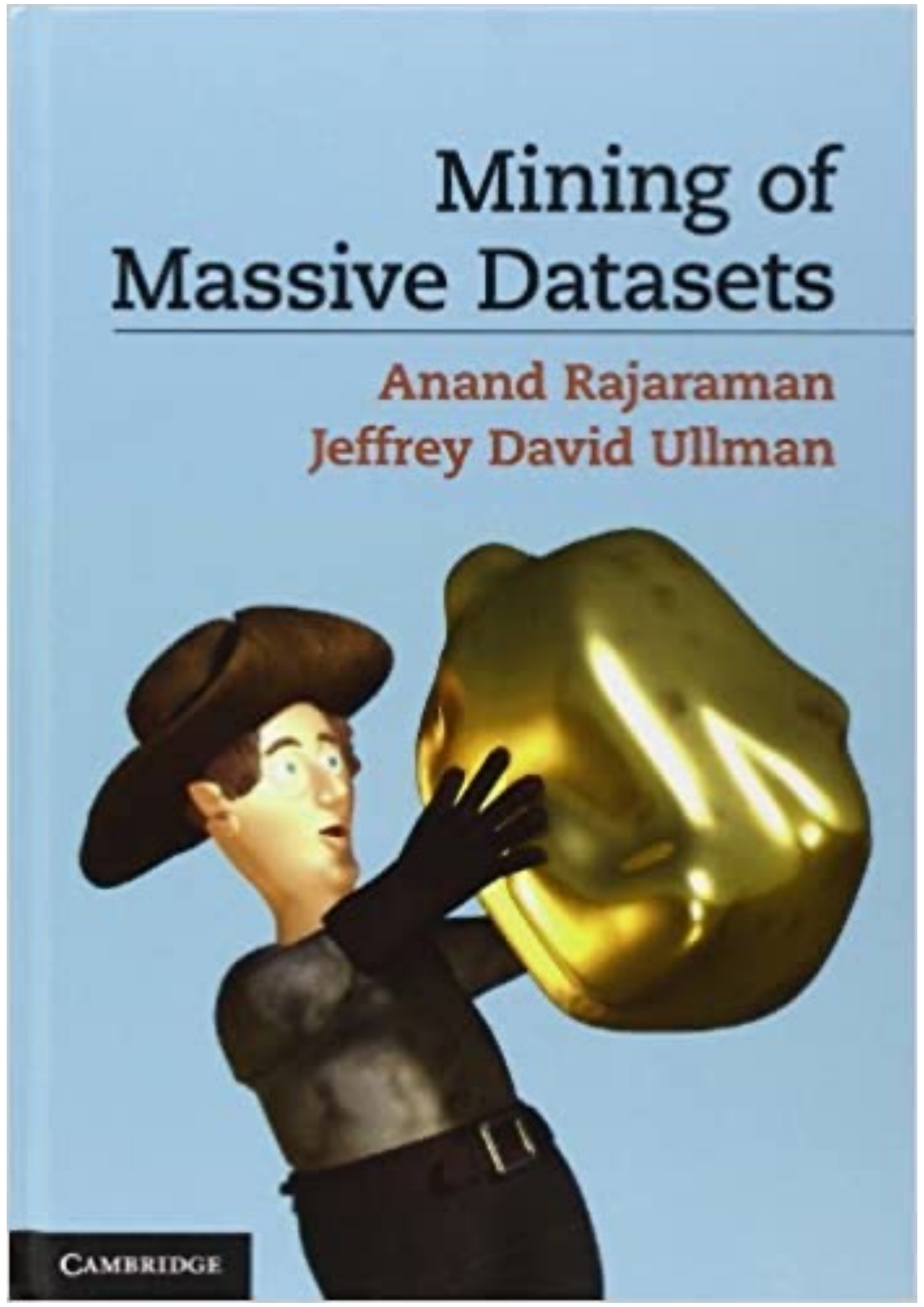
Precision/recall metrics do not use irrelevant documents (TN)!

RECOMMENDER SYSTEMS

Evaluating quality of recommendations

ASSOCIATION RULES VS RECOMMENDER SYSTEM

ONE MORE BOOK..



NATIONAL RESEARCH
UNIVERSITY