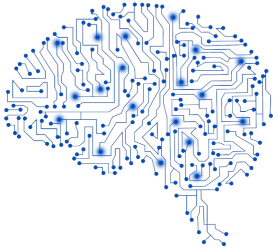
Recommendation System

Part I: Basic Concepts

Duc-Trong Le (PhD), VNU-UET

November 2022



Outline

- Real-life Examples
- Introduction to Recommendation Systems (RSs)
  - Item, User, Behavior
  - Role of RSs ? How do RSs work?
- Types of Recommendation Systems
- Learning Principles
  - Learning User-Item Associations
  - Learning Item-Item Associations
- Evaluation & Metrics
- Introduction to Cornac, a Recommendation Framework

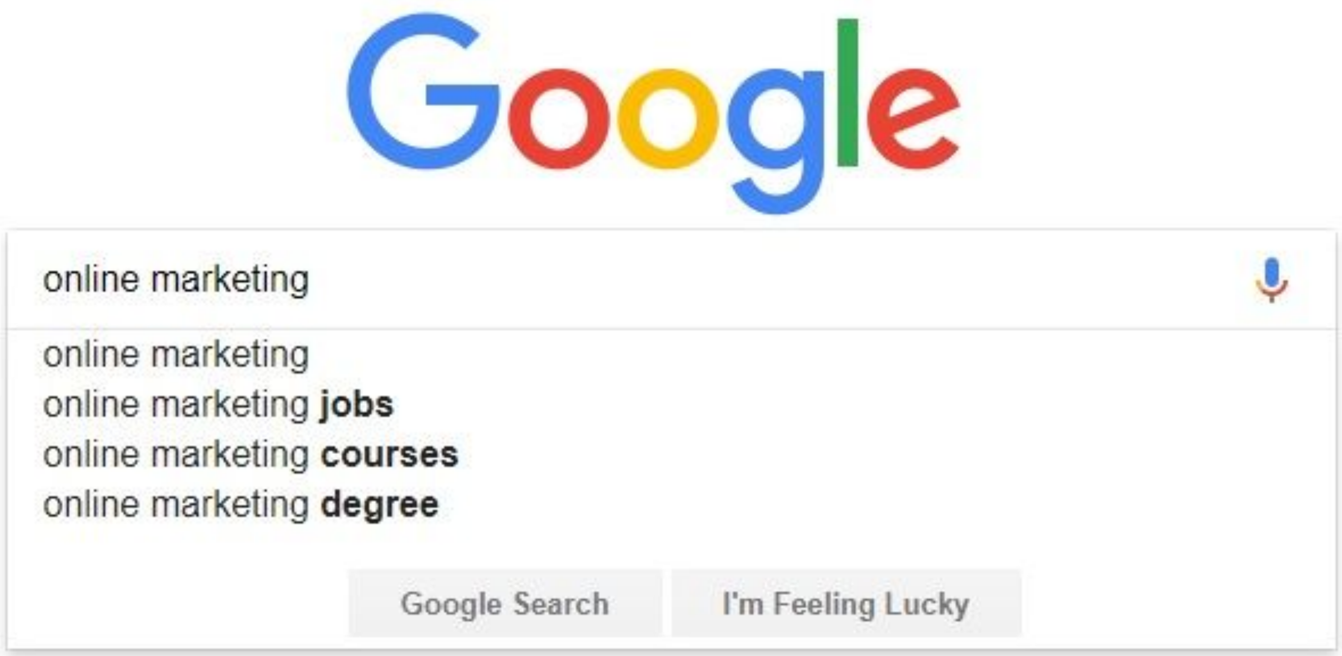
1.

Real-life Examples

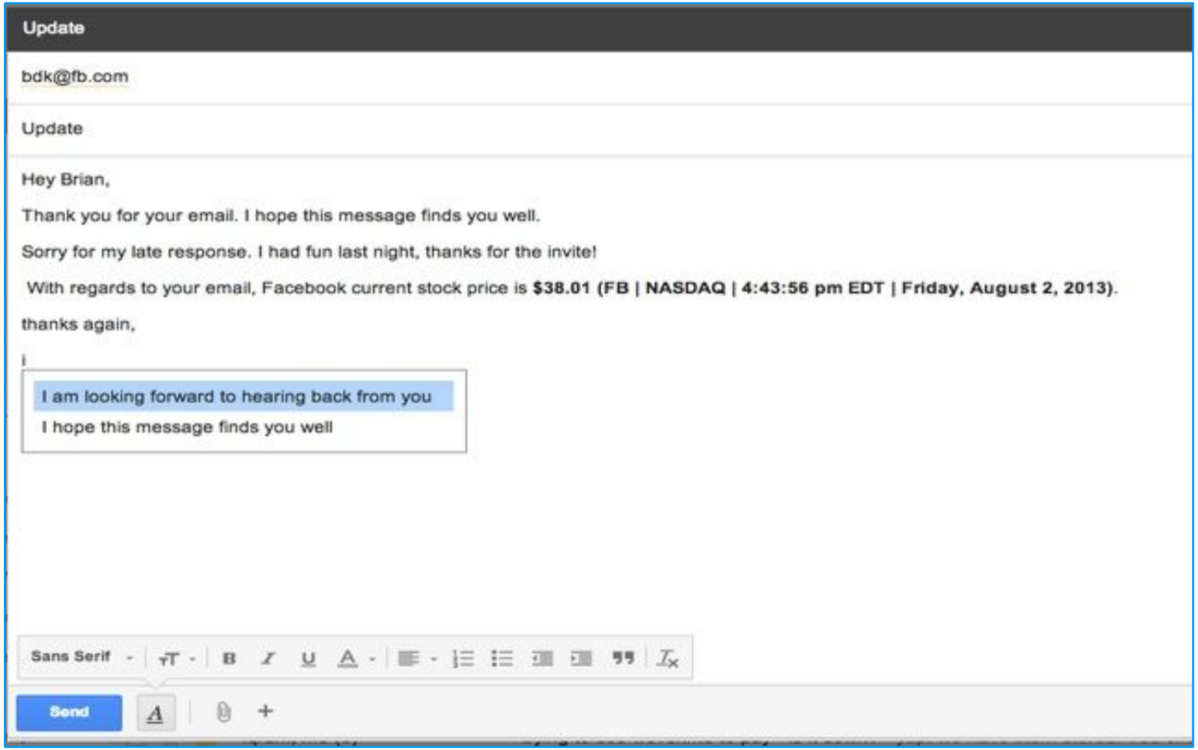
YouTube

Ad muted. Undo
We'll do our best to show you more relevant ads in the future.
Help us show you better ads by updating your ads preferences.
Google
3 Sick Bodybuilding Tips
by sixpackshortcuts
1,602,424 views
Ad
3:32
Ab To Hai Tum Sa Her Khushi Apni lata
by malikje1
45,661
FEATURED
4:00
Pink Interview on her husband and daughter
by MsSupergalaxy
Recommended for you
12:44
Abhimaan (1973): Loote Koi Man Ka Nagar
by tonguesonfire
5:56 / 9:55
Song: Tere Mere Milan Ki Yeh Raina Film: Abhimaan (1973) With Sinhala S...
madhukaw · 107 videos
Subscribe
1,738
35,944
41
4
Like
About
Share
Add to
Uploaded on Apr 20, 2010
No description available.
Buy "Tere Mere Milan Ki Yeh Raina" on

Google Search



Gmail

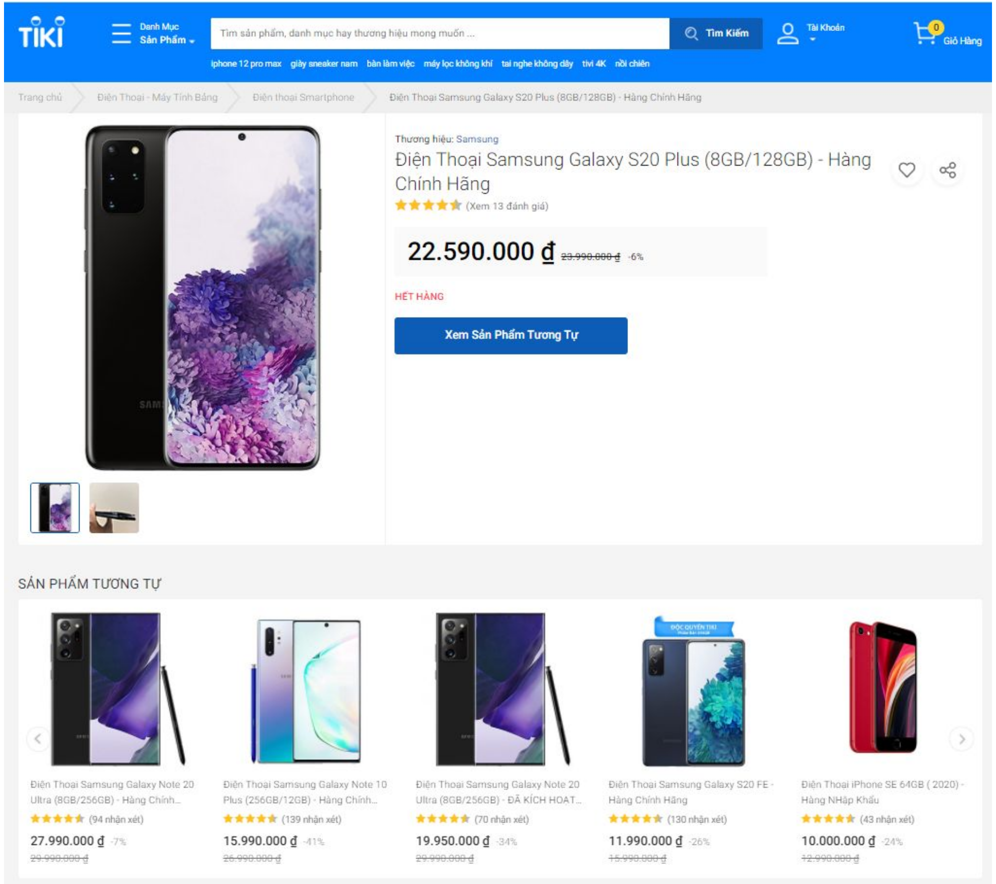


Update
bdk@fb.com
Update
Hey Brian,
Thank you for your email. I hope this message finds you well.
Sorry for my late response. I had fun last night, thanks for the invite!
With regards to your email, Facebook current stock price is $38.01 (FB | NASDAQ | 4:43:56 pm EDT | Friday, August 2, 2013).
thanks again,
I am looking forward to hearing back from you
I hope this message finds you well
Sans Serif
Send

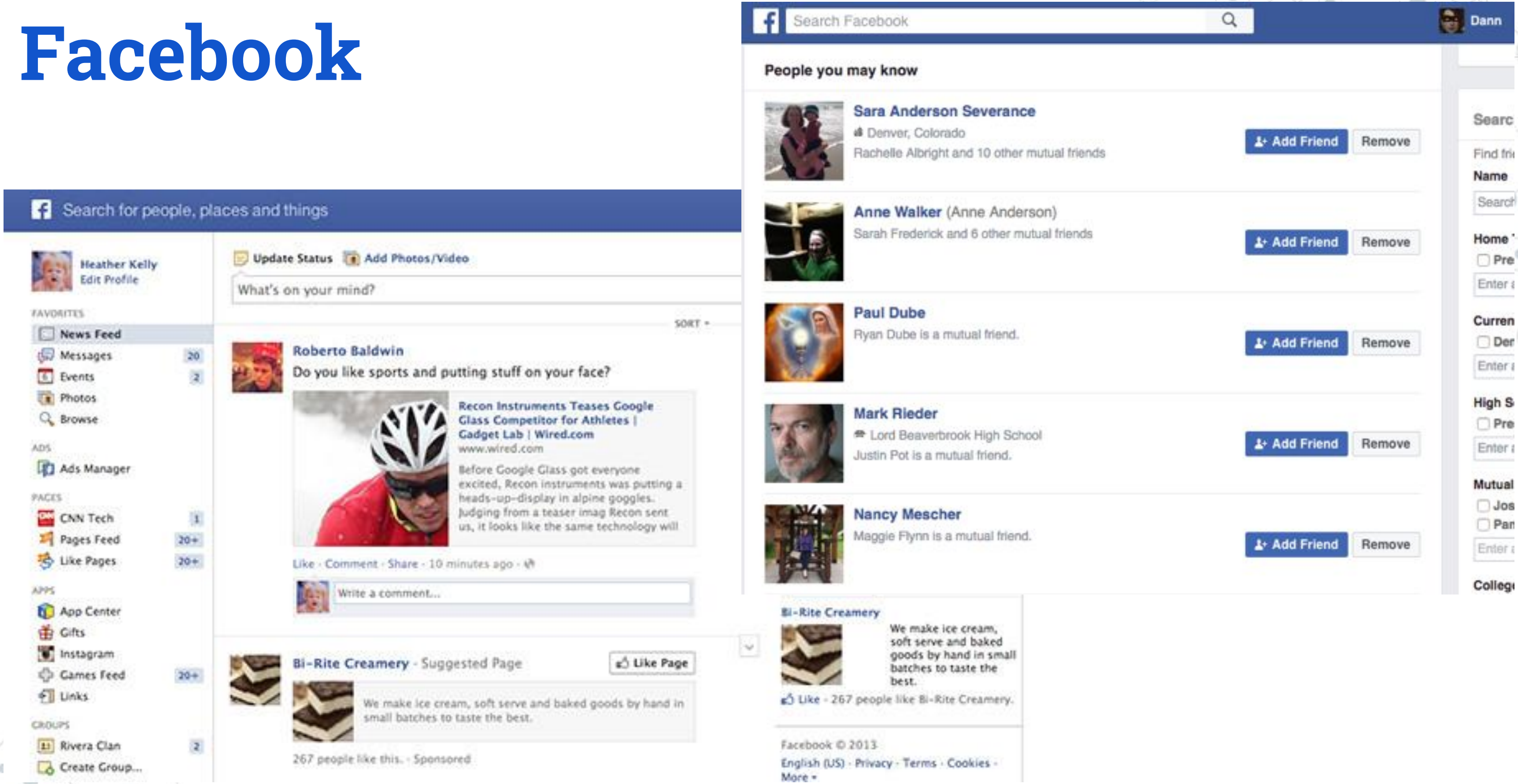
Amazon

Frequently Bought Together
Price for all three: $74.20
Add all three to Cart
Add all three to Wish List
Show availability and shipping details
This item: Beginning Ruby: From Novice to Professional (Expert's Voice in Open Source) by Peter Cooper Paperback $27.78
Learn to Program, Second Edition (The Facets of Ruby Series) by Chris Pine Paperback $16.94
Ruby on Rails Tutorial: Learn Web Development with Rails (2nd Edition) (Addison-Wesley Professional Ruby … by Michael Hartl Paperback $29.48
Customers Who Bought This Item Also Bought
Learn to Program, Second Edition (The Facets of…
Chris Pine
42
Paperback
$16.94 Prime
The Well-Grounded Rubyist
David A. Black
39
Paperback
$32.49 Prime
Ruby on Rails Tutorial: Learn Web Development…
Michael Hartl
70
Paperback
$29.48 Prime
The Ruby Programming Language
David Flanagan
74
Paperback
$26.35 Prime
The Well-Grounded Rubyist
David A. Black
19
#1 Best Seller in Ruby Programming Computer
Paperback
$29.67 Prime

Tiki


TIKI
Danh Mục
Sản Phẩm
Tìm sản phẩm, danh mục hay thương hiệu mong muốn ...
Tìm Kiếm
Tài Khoản
Giỏ Hàng
iphone 12 pro max giày sneaker nam bàn làm việc máy lọc không khí tai nghe không dây tivi 4K nồi chiên
Trang chủ Điện Thoại - Máy Tính Bảng Điện thoại Smartphone Điện Thoại Samsung Galaxy S20 Plus (8GB/128GB) - Hàng Chính Hãng
Thương hiệu: Samsung
Điện Thoại Samsung Galaxy S20 Plus (8GB/128GB) - Hàng Chính Hãng
(Xem 13 đánh giá)
22.590.000 ₫ 23.990.000 ₫ -6%
HẾT HÀNG
Xem Sản Phẩm Tương Tự
SẢN PHẨM TƯƠNG TỰ
Điện Thoại Samsung Galaxy Note 20 Ultra (8GB/256GB) - Hàng Chính...
(94 nhận xét)
27.990.000 ₫ -7%
29.990.000 ₫
Điện Thoại Samsung Galaxy Note 10 Plus (256GB/12GB) - Hàng Chính...
(139 nhận xét)
15.990.000 ₫ -41%
26.990.000 ₫
Điện Thoại Samsung Galaxy Note 20 Ultra (8GB/256GB) - ĐÃ KÍCH HOẠT...
(70 nhận xét)
19.950.000 ₫ -34%
29.990.000 ₫
Điện Thoại Samsung Galaxy S20 FE - Hàng Chính Hãng
(130 nhận xét)
11.990.000 ₫ -26%
15.990.000 ₫
Điện Thoại iPhone SE 64GB (2020) - Hàng NHập Khẩu
(43 nhận xét)
10.000.000 ₫ -24%
12.990.000 ₫

Facebook



Search for people, places and things
Heather Kelly
Edit Profile
FAVORITES
News Feed
Messages 20
Events 2
Photos
Browse
ADS
Ads Manager
PAGES
CNN Tech 1
Pages Feed 20+
Like Pages 20+
APPS
App Center
Gifts
Instagram
Games Feed 20+
Links
GROUPS
Rivera Clan 2
Create Group...
Update Status
Add Photos/Video
What's on your mind?
SORT
Roberto Baldwin
Do you like sports and putting stuff on your face?
Recon Instruments Teases Google Glass Competitor for Athletes | Gadget Lab | Wired.com
www.wired.com
Before Google Glass got everyone excited, Recon Instruments was putting a heads-up-display in alpine goggles. Judging from a teaser imag Recon sent us, it looks like the same technology will
Like · Comment · Share · 10 minutes ago
Write a comment...
Bi-Rite Creamery · Suggested Page
Like Page
We make ice cream, soft serve and baked goods by hand in small batches to taste the best.
267 people like this. · Sponsored
Search Facebook
Dann
People you may know
Sara Anderson Severance
Denver, Colorado
Rachelle Albright and 10 other mutual friends
Add Friend
Remove
Anne Walker (Anne Anderson)
Sarah Frederick and 6 other mutual friends
Add Friend
Remove
Paul Dube
Ryan Dube is a mutual friend.
Add Friend
Remove
Mark Rieder
Lord Beaverbrook High School
Justin Pot is a mutual friend.
Add Friend
Remove
Nancy Mescher
Maggie Flynn is a mutual friend.
Add Friend
Remove
Bi-Rite Creamery
We make ice cream, soft serve and baked goods by hand in small batches to taste the best.
Like · 267 people like Bi-Rite Creamery.
Facebook © 2013
English (US) · Privacy · Terms · Cookies ·
More

LinkedIn

Jobs you may be interested in

Data Scientist/Machine Learning Expert
Prudential Vietnam Assurance Private Ltd.
Vietnam
1 week ago · Easy Apply

Data Scientist [Ho Chi Minh]
OYO
Vietnam
5 days ago · Easy Apply

Software Engineer
ZALORA Group
Ho Chi Minh, Thành Phố Hồ Chí Minh, VN
4 days ago

Software Engineer (Java)
Kofax
Hanoi, VN
3 days ago

Professional Software Engineer
DXC Technology
Ho Chi Minh City, VN
2 days ago

AI Engineers
RGF Executive Search Vietnam
Vietnam
2 days ago · Easy Apply

Principal Consultant - Application Development
Oracle
Vietnam

Because you viewed

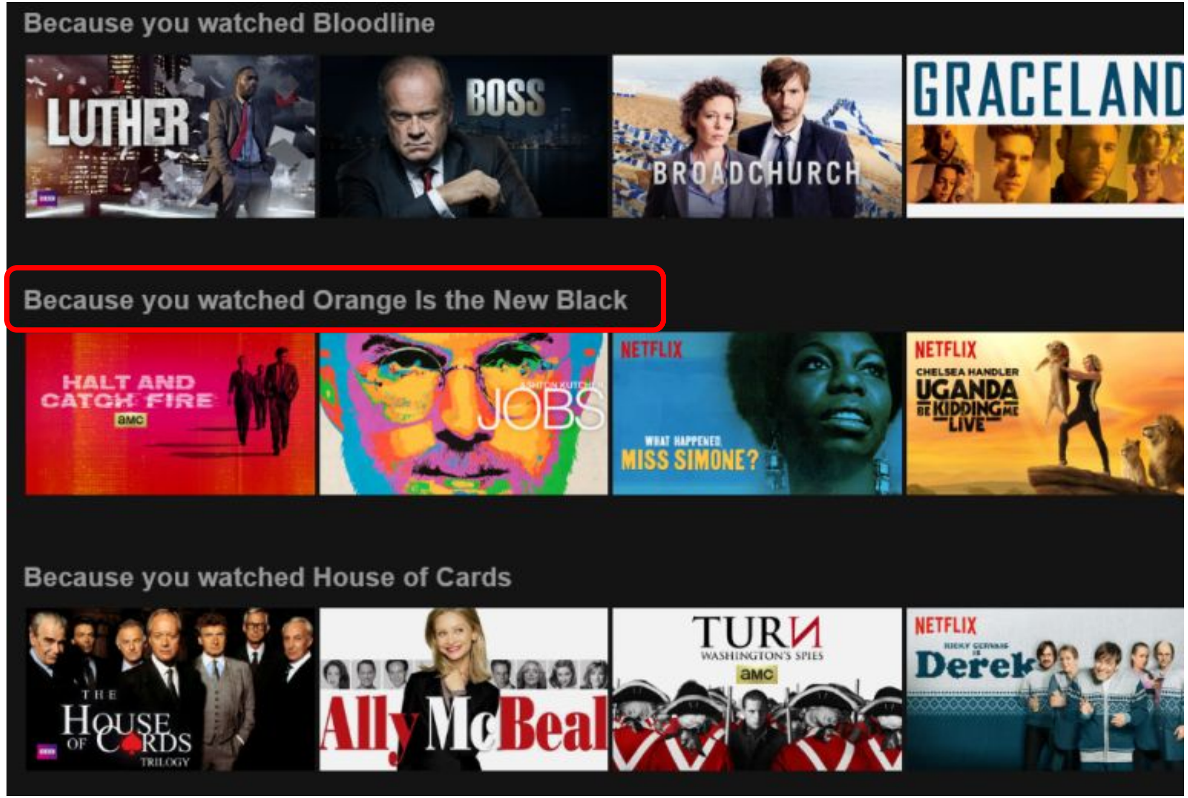
IBM Research Scientist - Health AI Postdoc FTH 24Months Melbourne at IBM

Assoc Research Scientist
PPD
Richmond, VA, US
Be an early applicant
3 weeks ago

Post-Doctoral Research Fellow
Murdoch University
Perth, Australia
3 weeks ago

Post-Doctoral Position in Biophotonics (modelling) (Ref...
Tyndall National Institute
County Cork, Ireland
Be an early applicant
1 month ago

Netflix



Because you watched Bloodline
LUTHER
BOSS
BROADCHURCH
GRACELAND
Because you watched Orange Is the New Black
HALT AND CATCH FIRE
amc
JOBS
NETFLIX
WHAT HAPPENED, MISS SIMONE?
NETFLIX
CHELSEA HANDLER
UGANDA BE KIDDING ME LIVE
Because you watched House of Cards
THE HOUSE OF CARDS TRILOGY
Ally McBeal
TURN
WASHINGTON'S SPIES
amc
NETFLIX
Derek

Medium

BASED ON YOUR READING HISTORY

Visualizing Data with Pairs Plots in Python

How to quickly create a powerful exploratory data analysis visualization

Will Koehrsen in Towards Data Science

Apr 7, 2018 · 8 min read

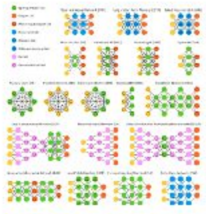
ARTIFICIAL INTELLIGENCE

Cheat Sheets for AI, Neural Networks, Machine Learning, Deep Learning & Big Data

The Most Complete List of Best AI Cheat Sheets

Stefan Kojouharov in Becoming Human: Artificial Intelligence Magazine

Jul 9, 2017 · 7 min read



ARTIFICIAL INTELLIGENCE

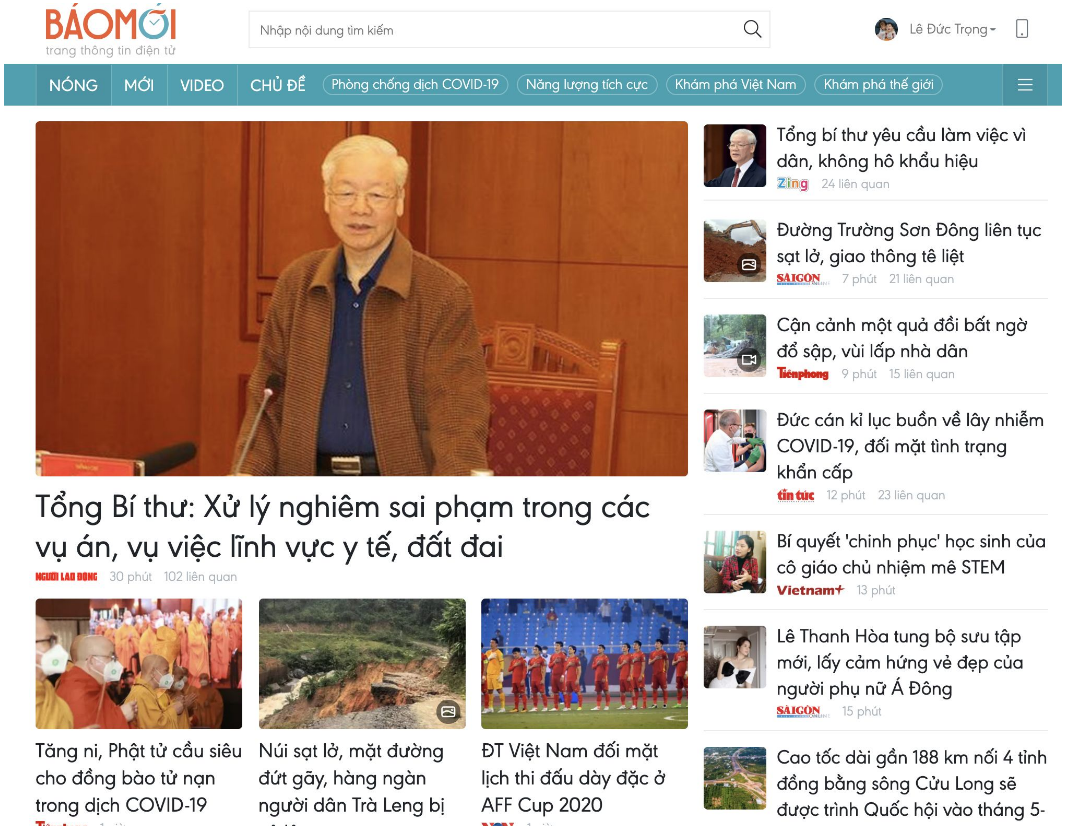
A Conceptual Explanation of Bayesian Hyperparameter Optimization for Machine Learning

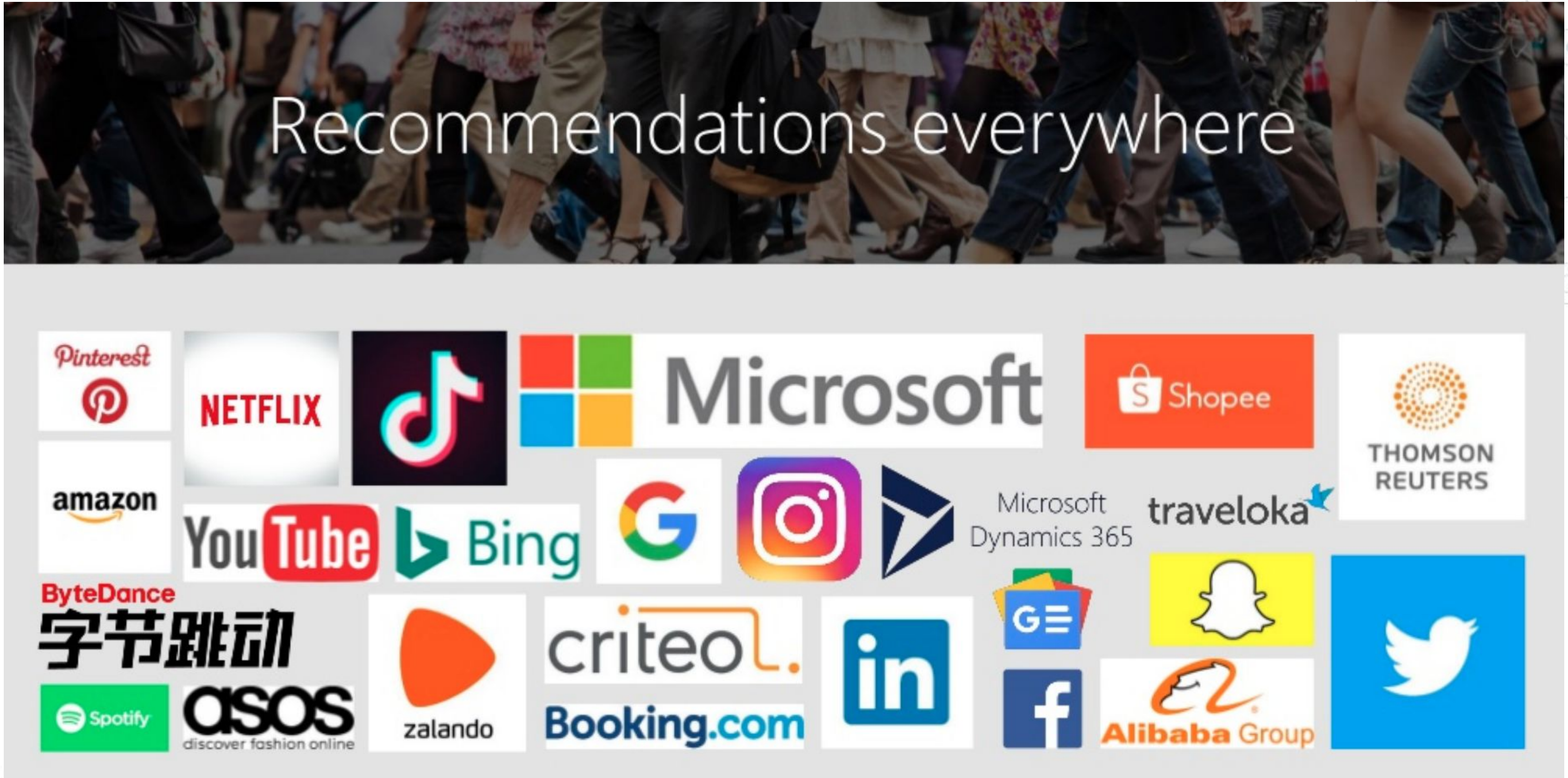
The concepts behind efficient hyperparameter tuning using Bayesian optimization

Will Koehrsen in Towards Data Science

Jun 24, 2018 · 14 min read

BaoMoi

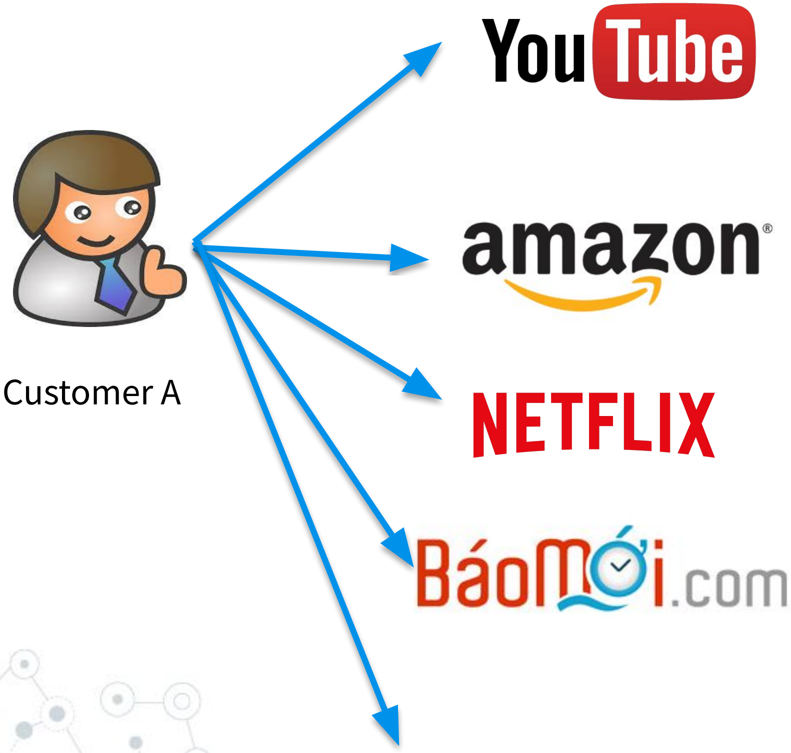

BÁOMỚI
trang thông tin điện tử
Nhập nội dung tìm kiếm
Lê Đức Trọng
NÓNG
MỚI
VIDEO
CHỦ ĐỀ
Phòng chống dịch COVID-19
Năng lượng tích cực
Khám phá Việt Nam
Khám phá thế giới
Tổng Bí thư: Xử lý nghiêm sai phạm trong các vụ án, vụ việc lĩnh vực y tế, đất đai
NGƯỜI LAO ĐỘNG 30 phút 102 liên quan
Tăng ni, Phật tử cầu siêu cho đồng bào tử nạn trong dịch COVID-19
Núi sạt lở, mặt đường đứt gãy, hàng ngàn người dân Trà Leng bị
ĐT Việt Nam đối mặt lịch thi đấu dày đặc ở AFF Cup 2020
Tổng bí thư yêu cầu làm việc vì dân, không hô khẩu hiệu
Zing 24 liên quan
Đường Trường Sơn Đông liên tục sạt lở, giao thông tê liệt
SÀIGÒN 7 phút 21 liên quan
Cận cảnh một quả đồi bất ngờ đổ sập, vùi lấp nhà dân
Tiền phong 9 phút 15 liên quan
Đức cán kỉ lục buồn về lây nhiễm COVID-19, đối mặt tình trạng khẩn cấp
tin tức 12 phút 23 liên quan
Bí quyết 'chinh phục' học sinh của cô giáo chủ nhiệm mê STEM
Vietnam+ 13 phút
Lê Thanh Hòa tung bộ sưu tập mới, lấy cảm hứng vẻ đẹp của người phụ nữ Á Đông
SÀIGÒN 15 phút
Cao tốc dài gần 188 km nối 4 tỉnh đồng bằng sông Cửu Long sẽ được trình Quốc hội vào tháng 5-


Recommendations everywhere
Pinterest
NETFLIX
Microsoft
Shopee
THOMSON REUTERS
amazon
YouTube
Bing
Microsoft Dynamics 365
traveloka
ByteDance
字节跳动
criteo
zalando
Spotify
asos
discover fashion online
Booking.com
Alibaba Group

2.

Introduction to Recommendation Systems

Item, User & Behavior

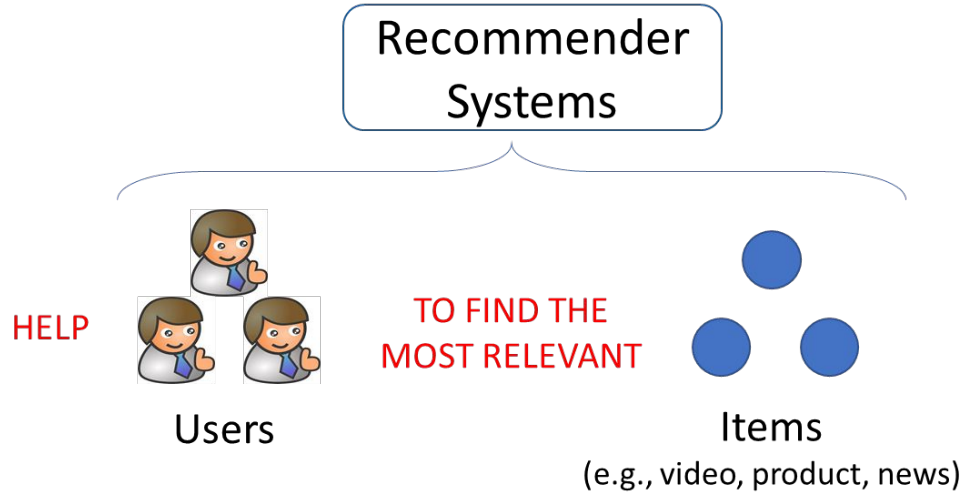


	Behavior	Item
YouTube	Click, Like/Dislike, Comment, Search	**Video**
amazon	Search, Click, Add-to-cart, Purchase, Review, Rate	**Product**
NETFLIX	Search, Click, Rate, Like	**Movie**
BáoMới.com	Search, Read, Duration, Comment	**News**
...	...	...

Why do we need Recommendation?

Millions of Items ...

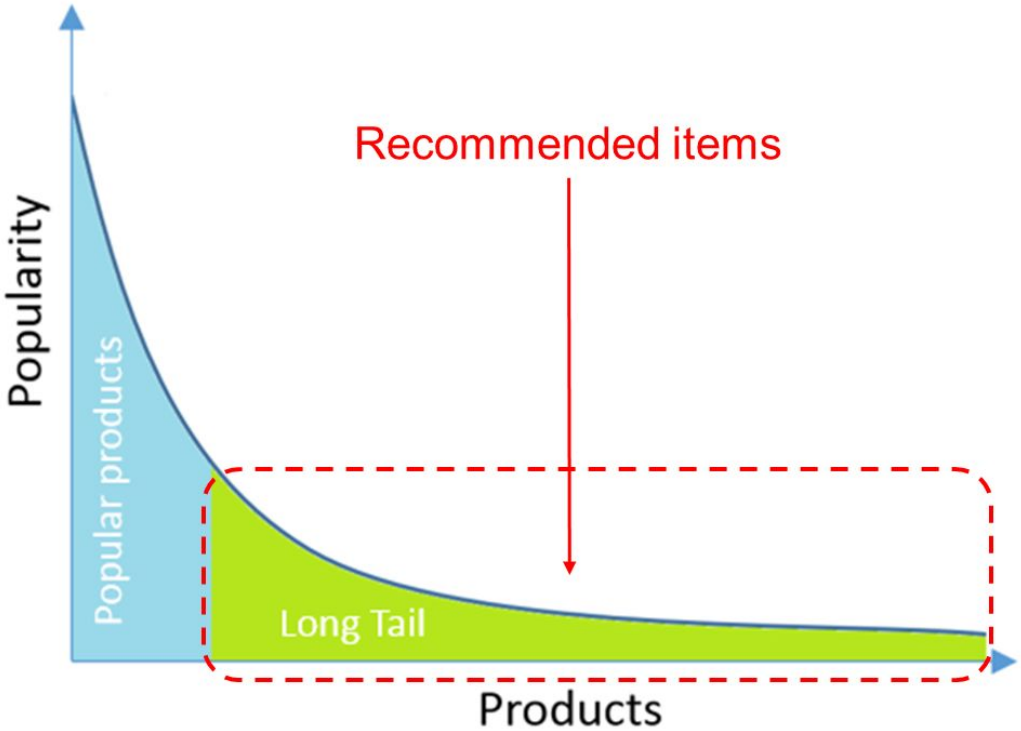
Roles of Recommendation Systems



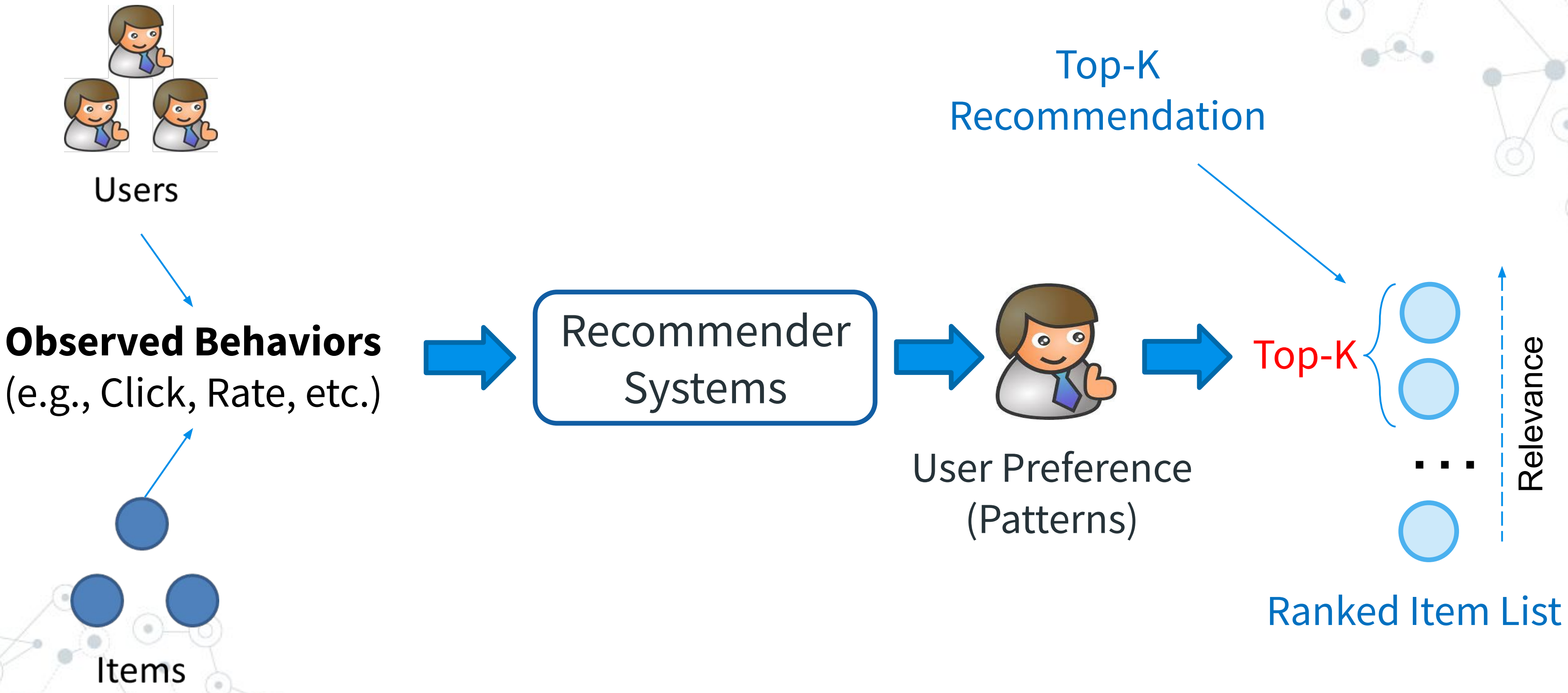
Advantages of RS:

Enterprise Perspective	Customer Perspective
• Reduce search costs • Increase customer satisfaction • Educate customers about product domains. • Optimize sales and profit	• Easy to find what he/she might want • Having an assistant in websites/systems

Long-tail Recommendation



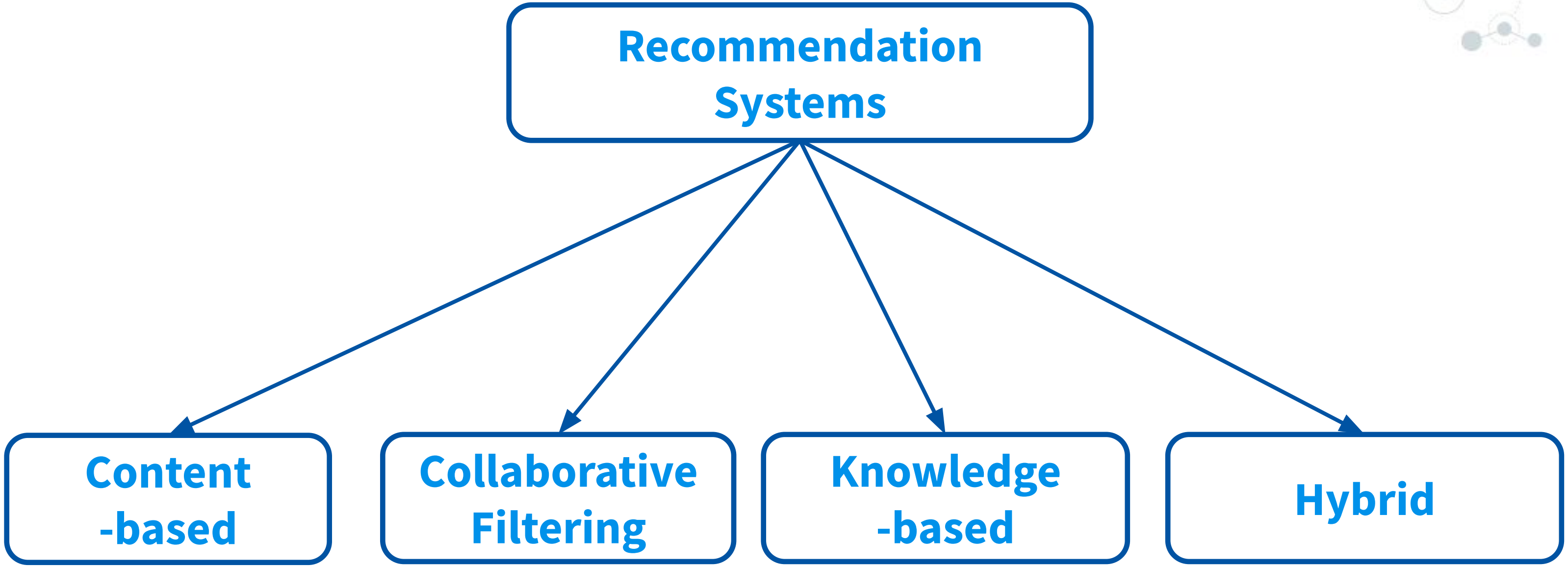
How does RS work?



3.

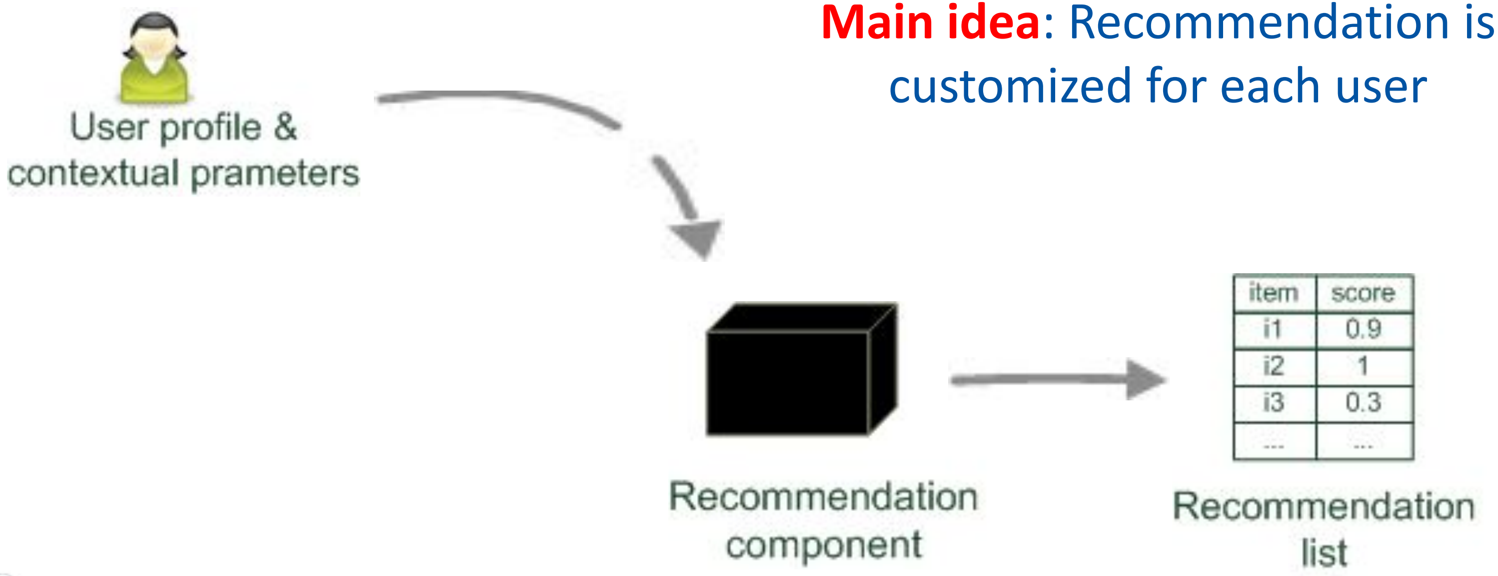
Types of Recommendation Systems

Types of Recommendation Systems



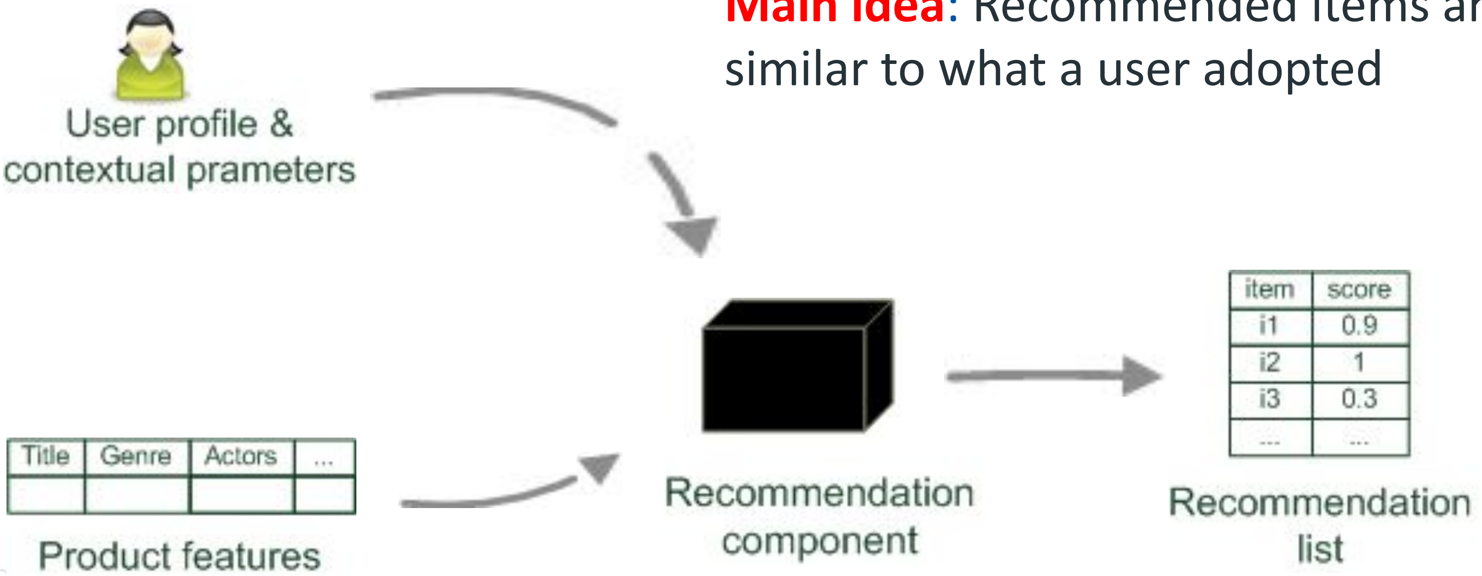
Personalized Recommendation

Main idea: Recommendation is customized for each user

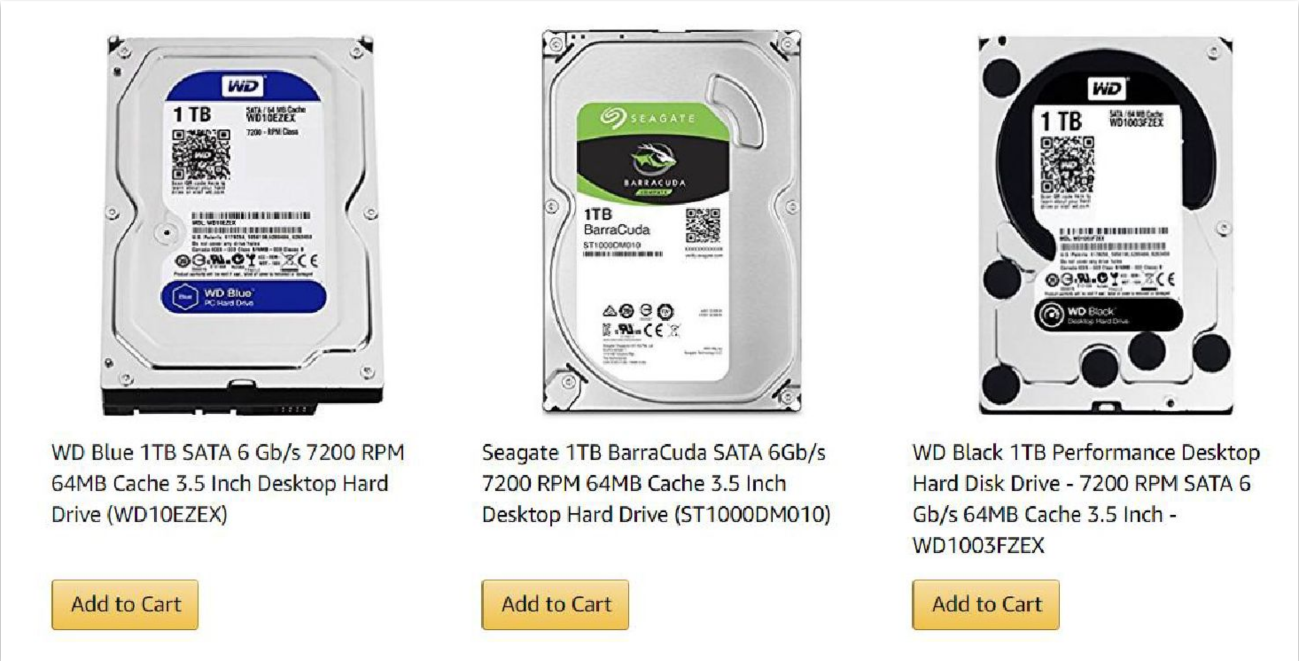


Content-based Recommendation (1)

Main idea: Recommended items are similar to what a user adopted

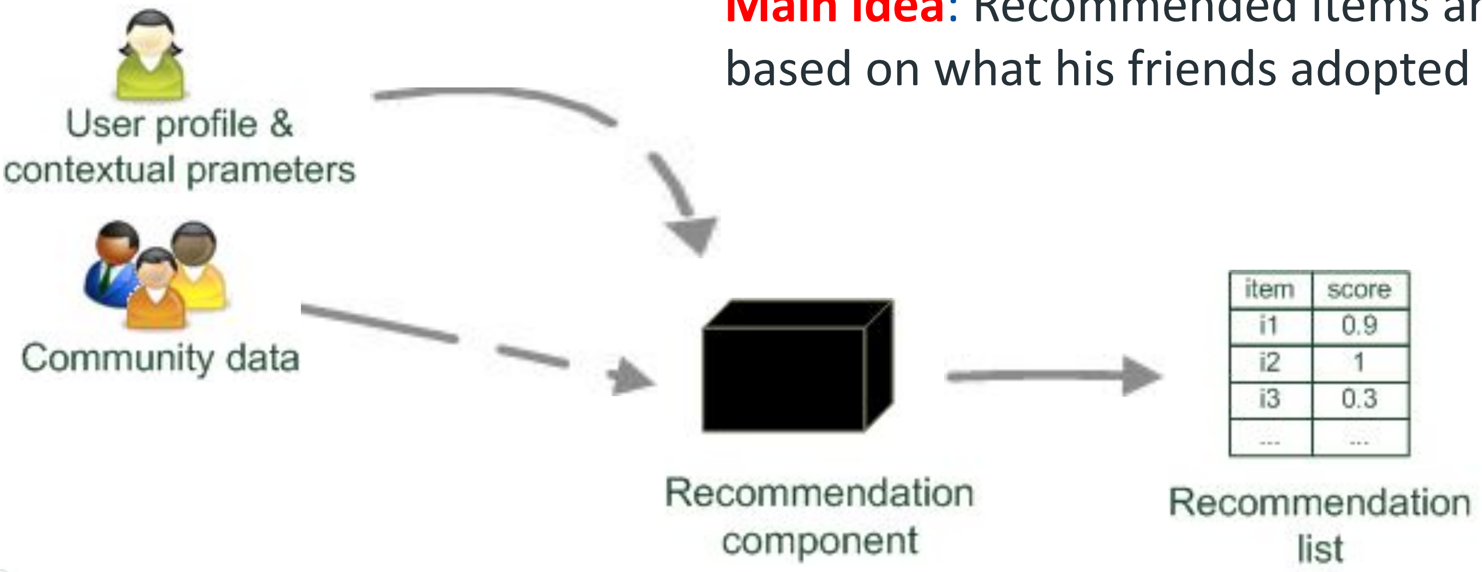


Content-based Recommendation (2)



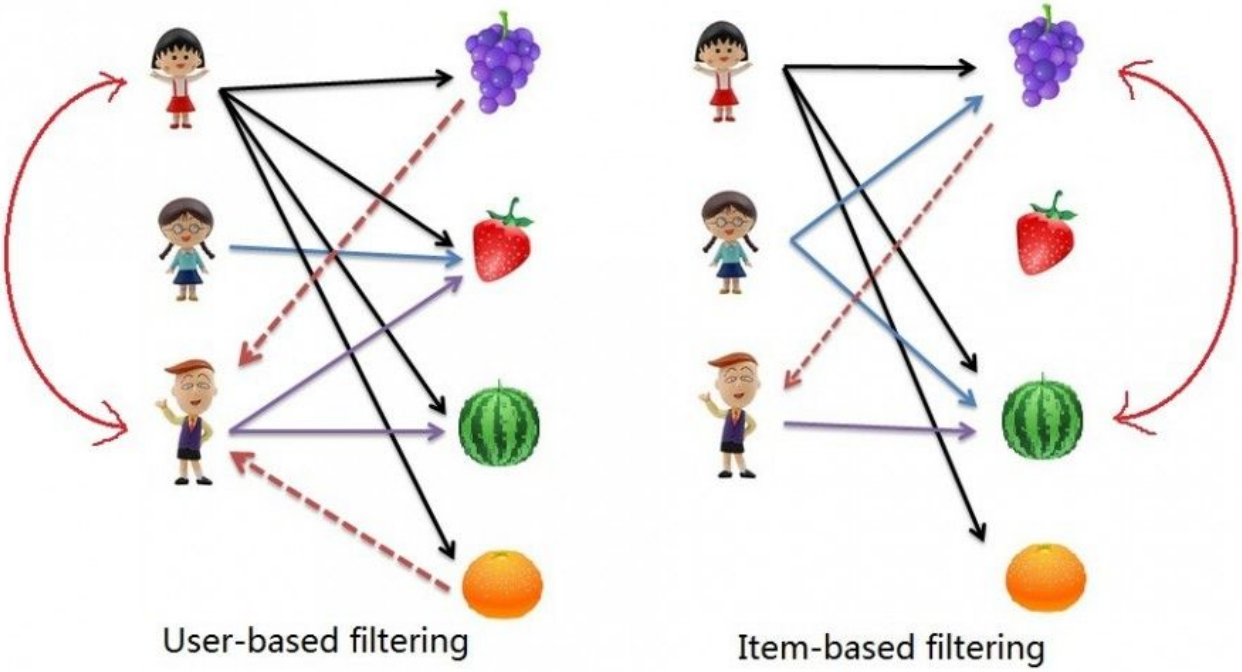
Recommendations for 1TB HDD on Amazon

Collaborative Filtering (1)



Main idea: Recommended items are based on what his friends adopted

Collaborative Filtering (2)



Netflix Prize

Netflix Prize

COMPLETED

Home | Rules | Leaderboard | Update | Download

Leaderboard

Showing Test Score. Click here to show quiz score

Display top 20 leaders.

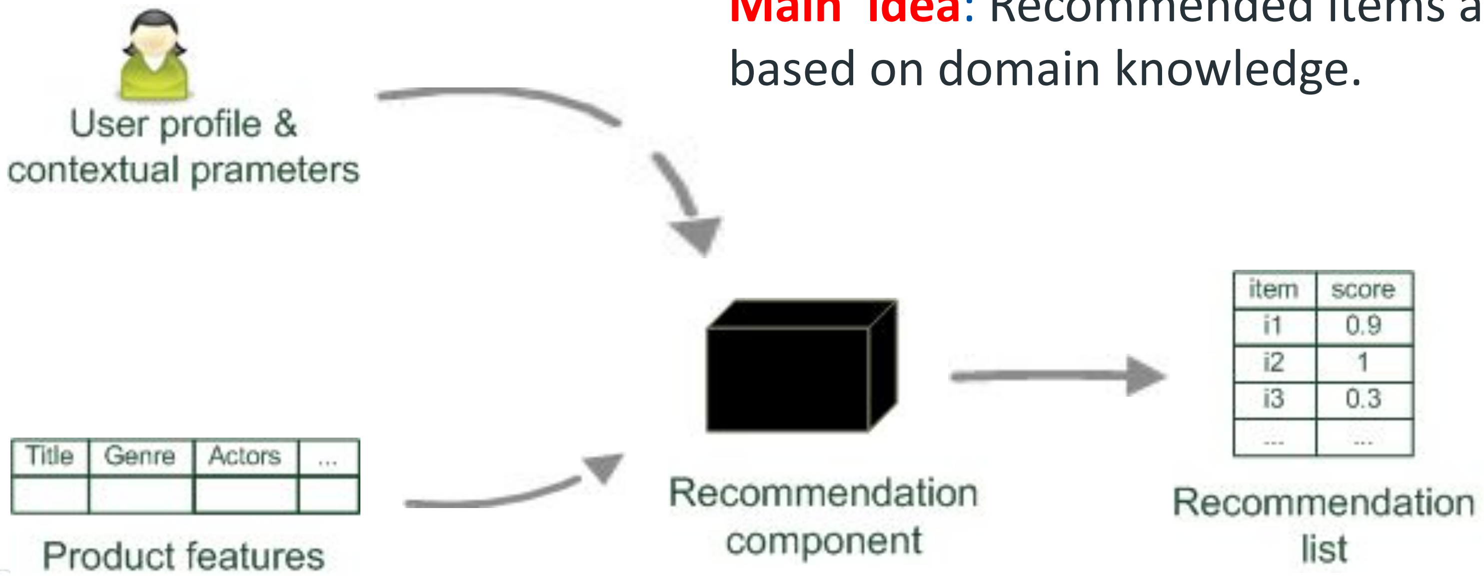
Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos				
1	BellKor's Pragmatic Chaos	0.8567	10.06	
2	The Ensemble	0.8567	10.06	
3	Grand Prize Team	0.8582	9.90	
4	Opera Solutions and Vandelay United	0.8588	9.84	
5	Vandelay Industries !	0.8591	9.81	
6	PragmaticTheory	0.8594	9.77	
7	BellKor in BigChaos	0.8601	9.70	
8	Dace	0.8612	9.59	

	Movie 1	Movie 2	Movie 3
Ted	4	5	5
Carol		5	5
Bob		5	?

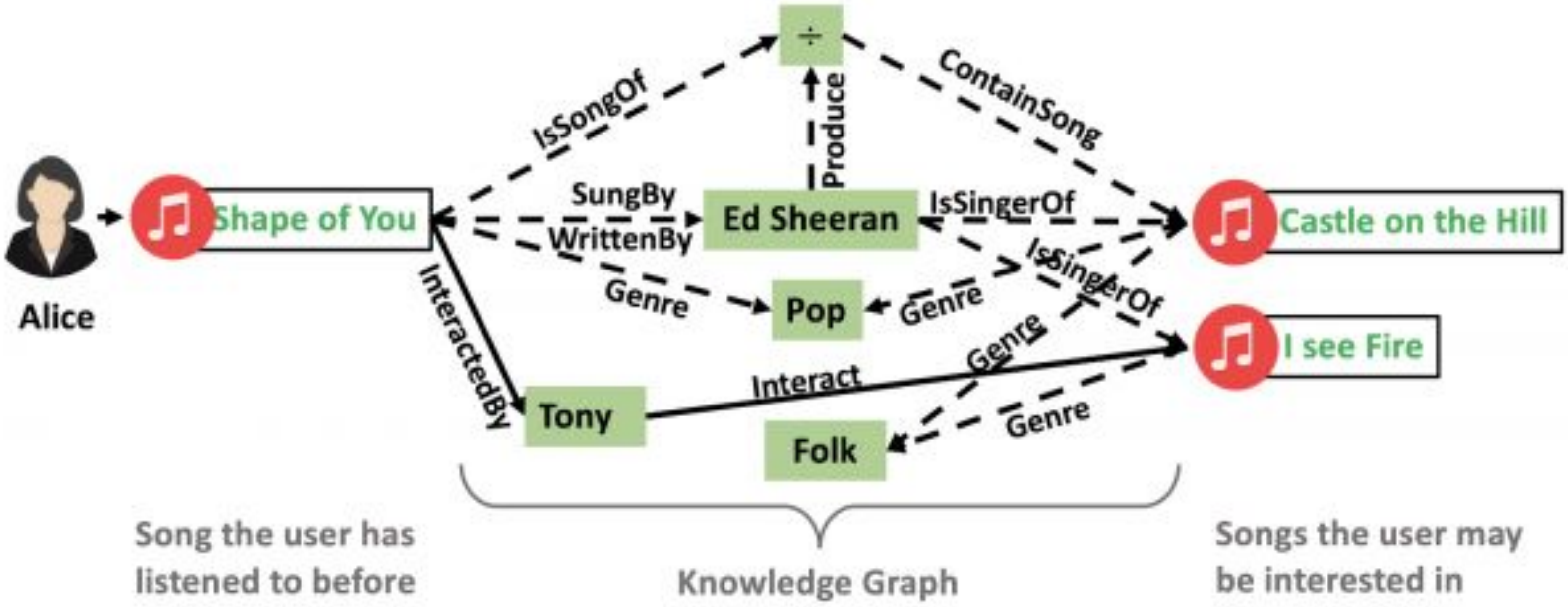
Netflix 1M $ challenge

Knowledge-based Recommendation (1)

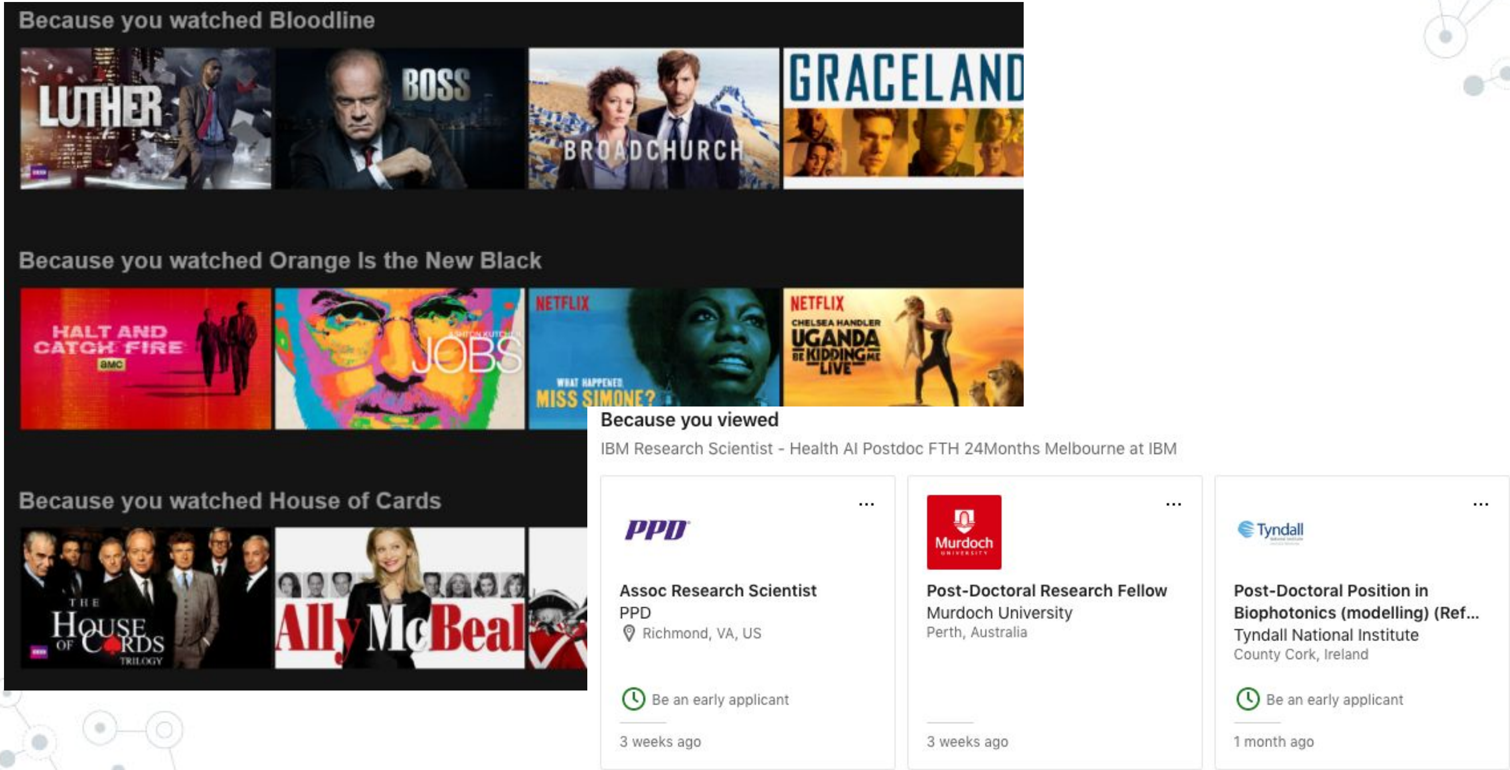
Main idea: Recommended items are based on domain knowledge.



Knowledge-based Recommendation (2)

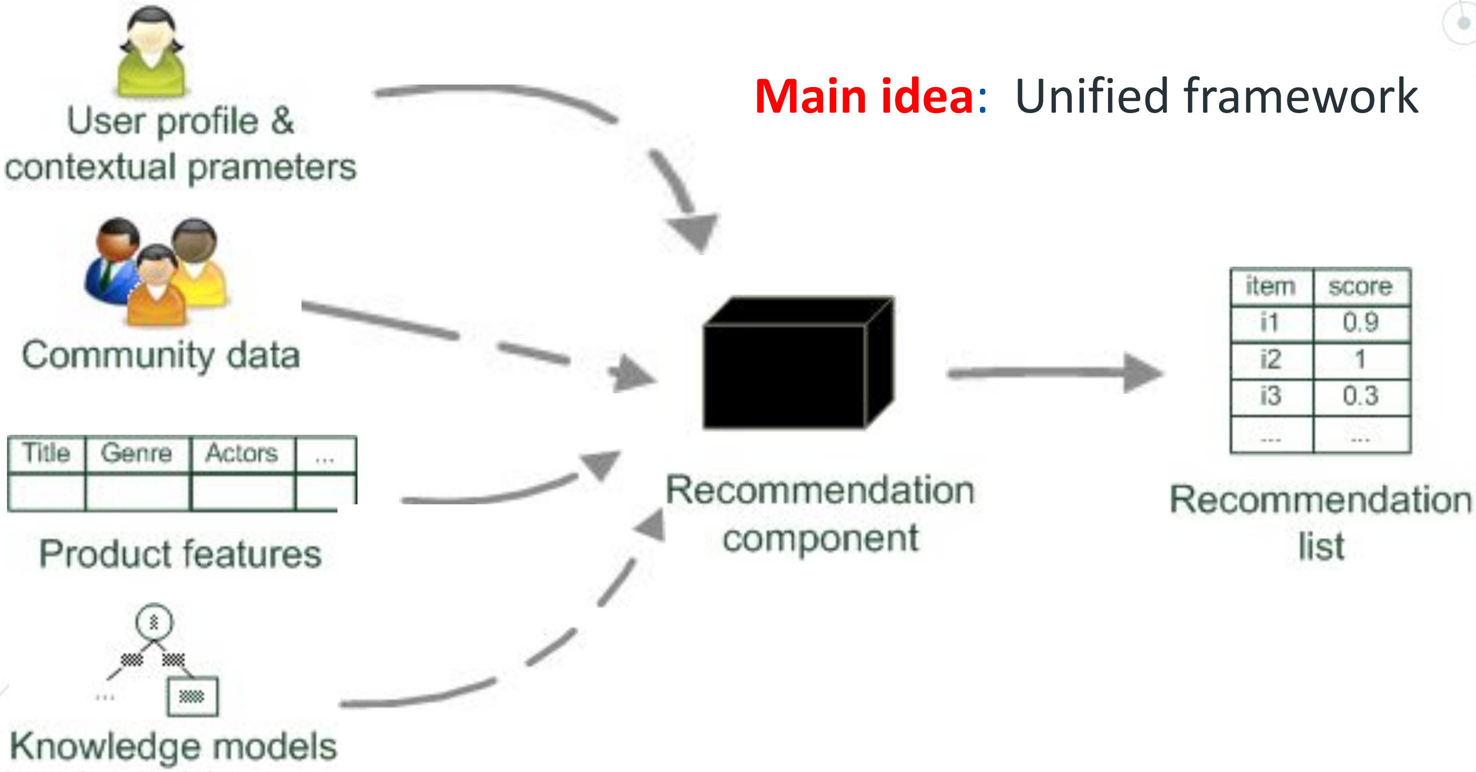


Knowledge-based Recommendation (3)

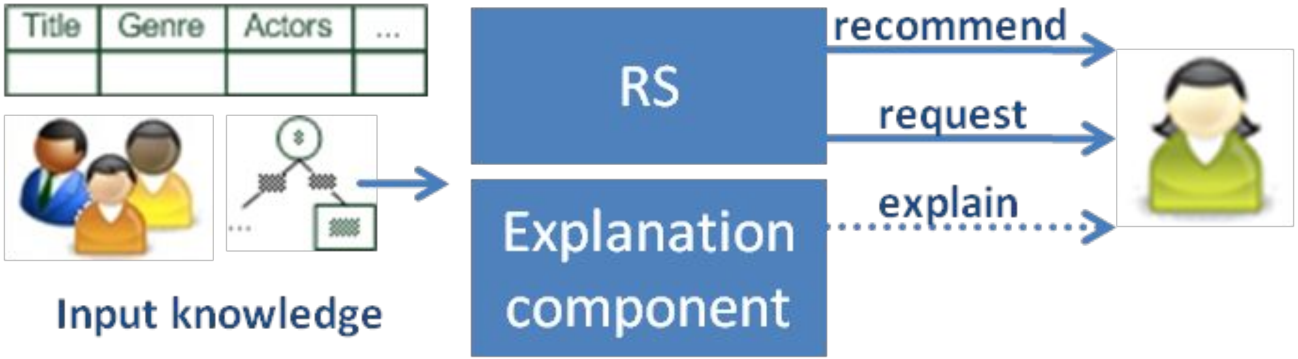


Hybrid Recommendation

Main idea: Unified framework



Explainable Recommendation

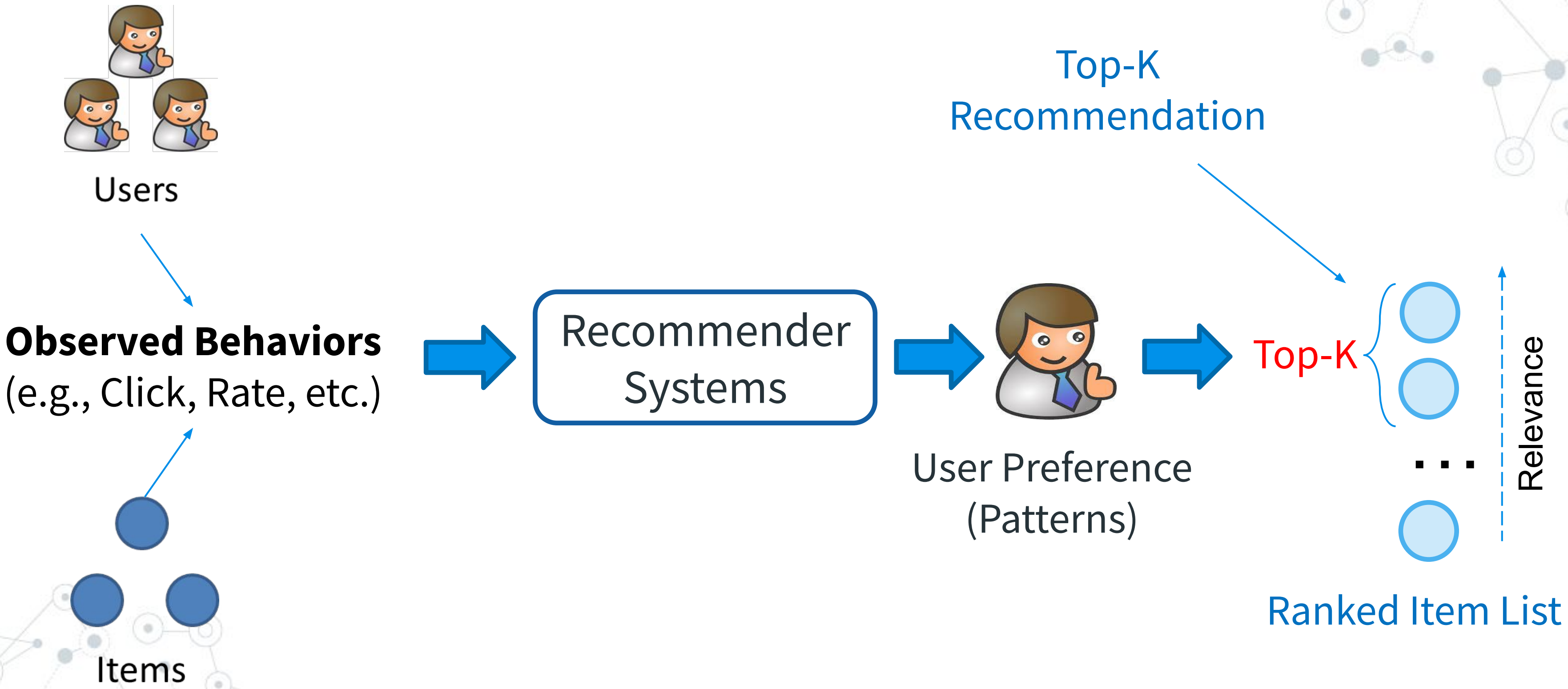


Main idea: Recommendation with explanation

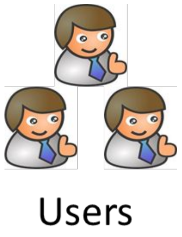
4.

Learning Principles

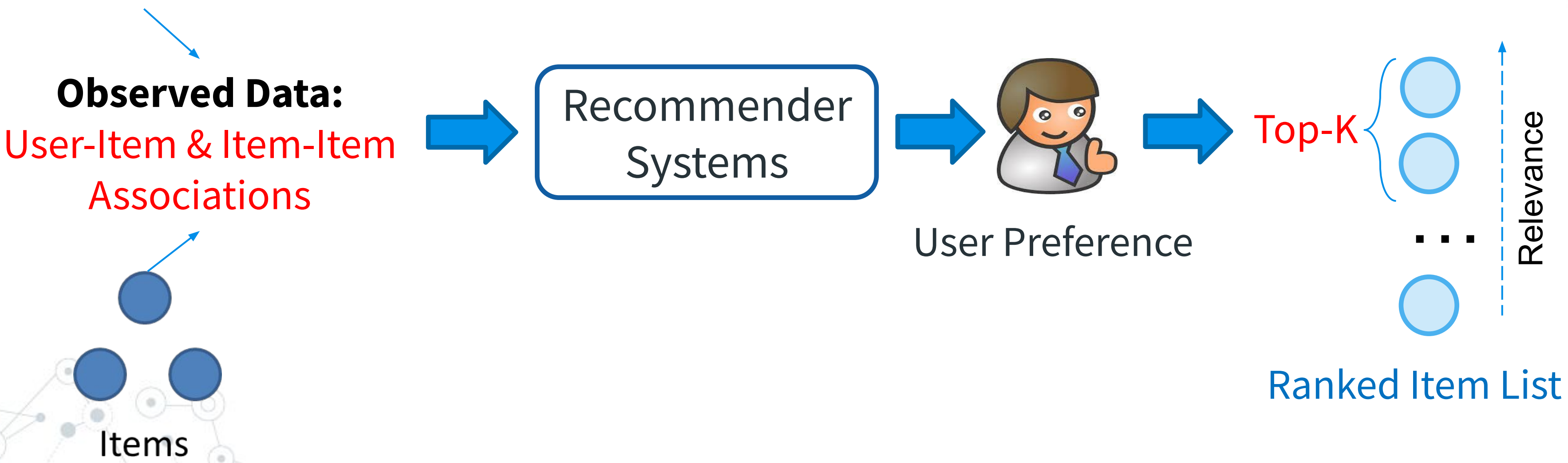
How does RS work?



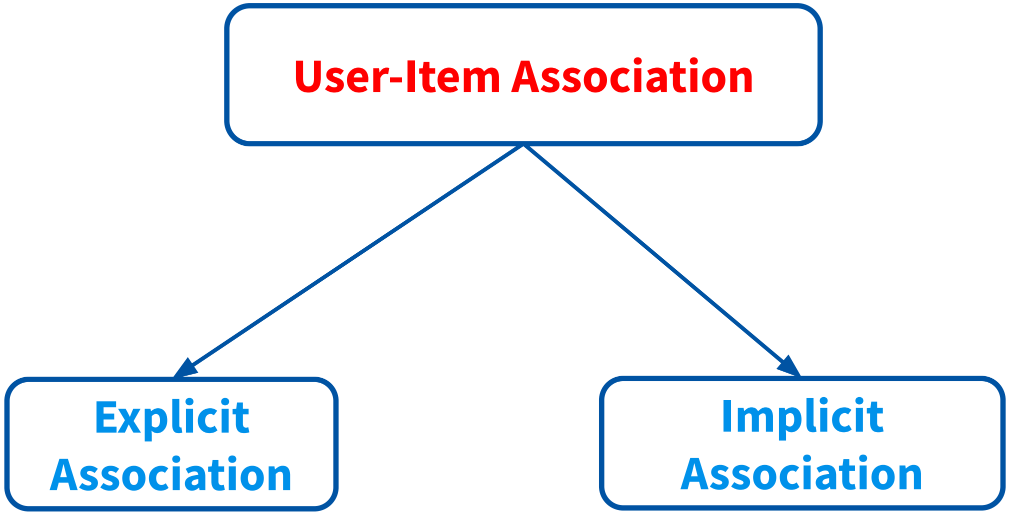
Top-K Recommendation



<u>Main idea</u>: Learning **user preferences** via modeling **associations** to generate **top-K recommendations**.



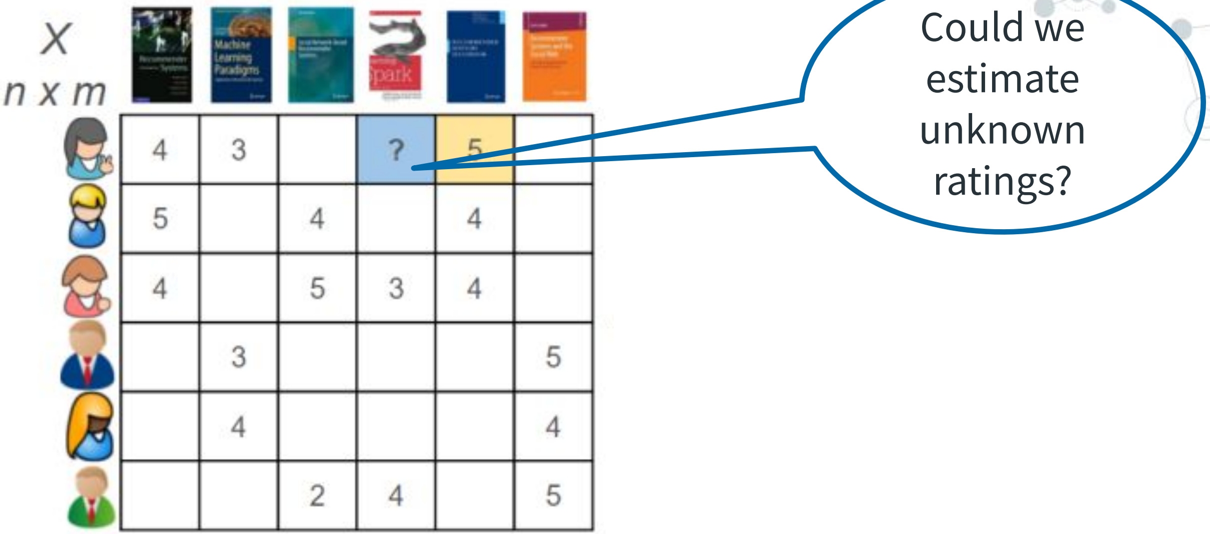
User-Item Association



e.g., rating, score, stars

e.g., click, view, purchase

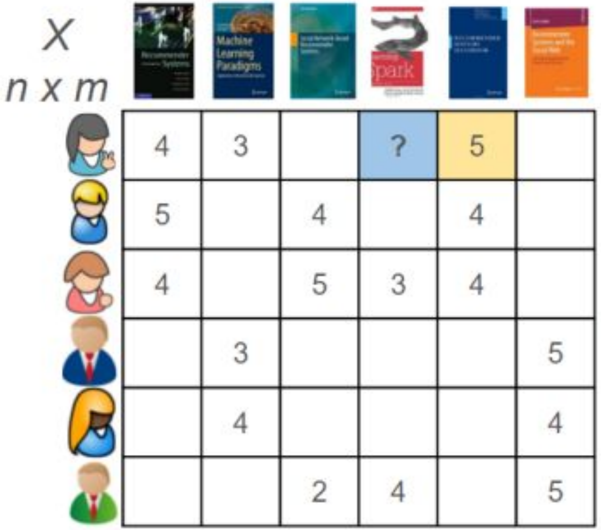
Explicit Association-based Recommendation



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

Main idea: Recommended items are high rating ones.

Rating Prediction

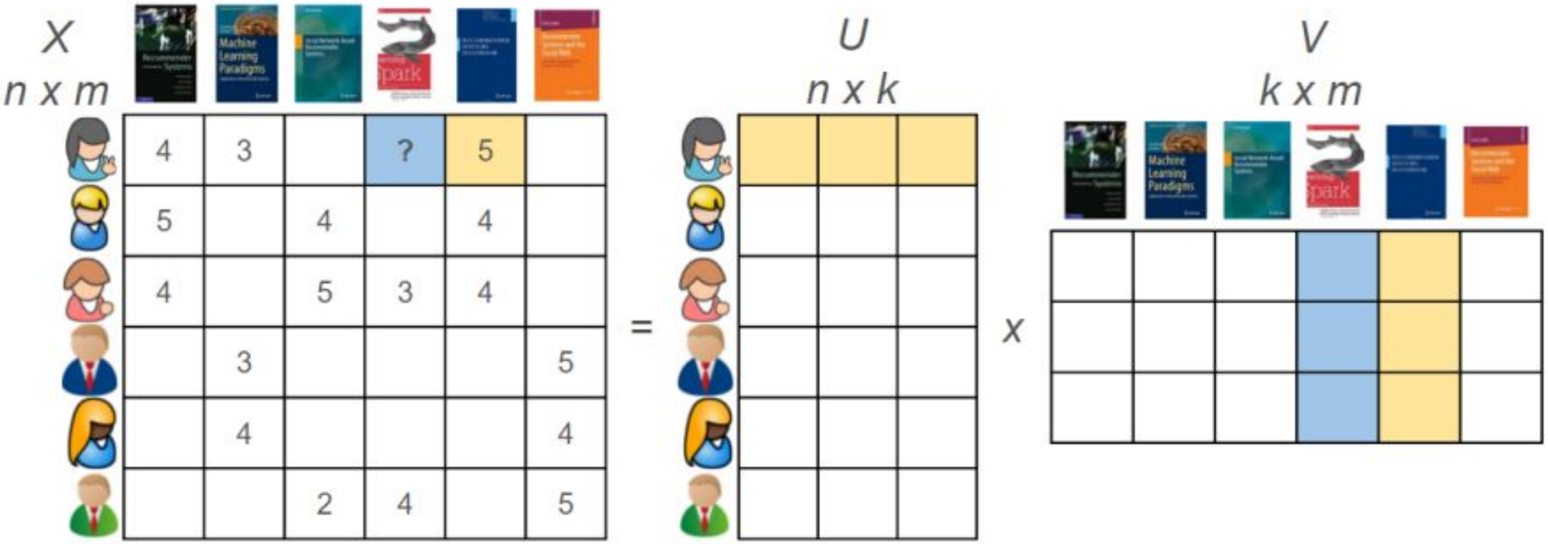


4	3		?	5	
5		4		4	
4		5	3	4	
	3				5
	4				4
		2	4		5

$$\hat{y}_{i,u} = \frac{\sum_{u_j \in \mathcal{N}(u,i)} \bar{y}_{i,u_j} \mathrm{sim}(u, u_j)}{\sum_{u_j \in \mathcal{N}(u,i)} |\mathrm{sim}(u, u_j)|}$$

Main idea: Rating is computed via similar users.

Rating Prediction with Matrix Factorization

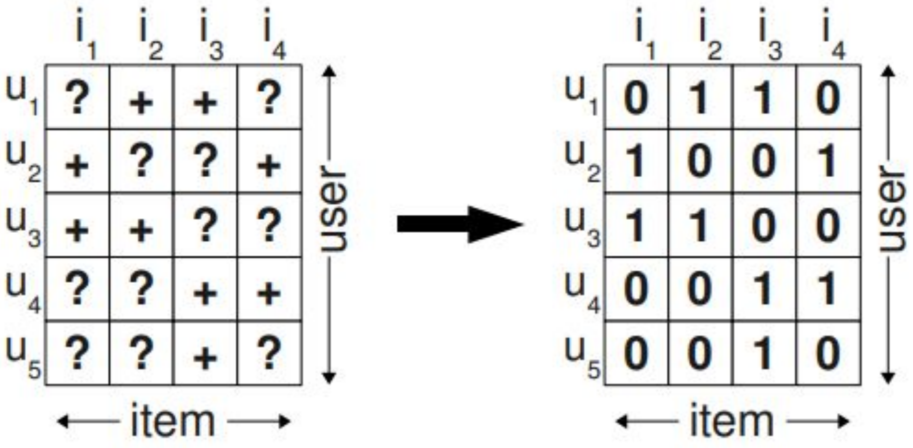


User
latent representation

Item
latent representation

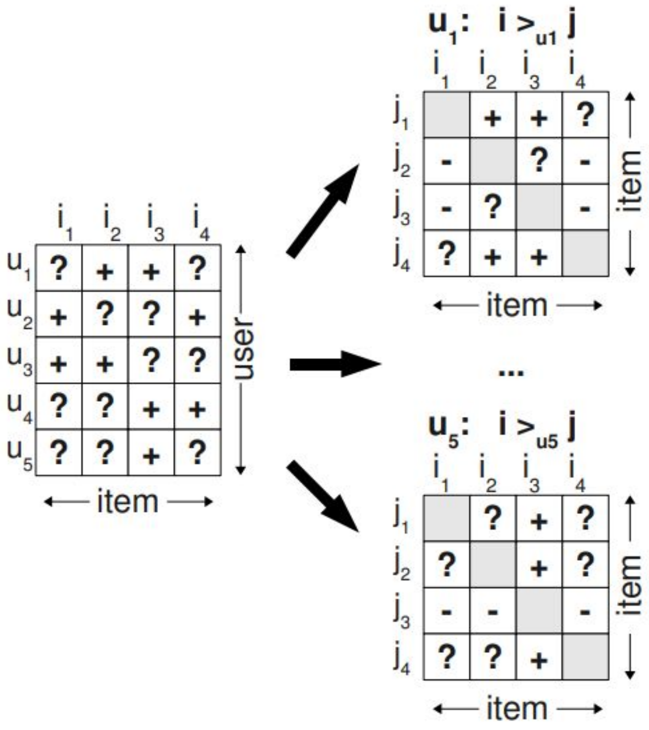
Learning Objective: Minimize $\left(X_{ij} - U_i \times V_j^T\right)^2 + \lambda(||U||_2 + ||V||_2)$

Implicit Association-based Recommendation



Main idea: Recommended items are popular (Counting)

Bayesian Personalized Ranking (BPR)



$$\prod_{u \in U} p(>_u | \Theta) = \prod_{(u,i,j) \in D_S} p(i >_u j | \Theta)$$

$$p(i >_u j | \Theta) := \sigma(\hat{x}_{uij}(\Theta))$$

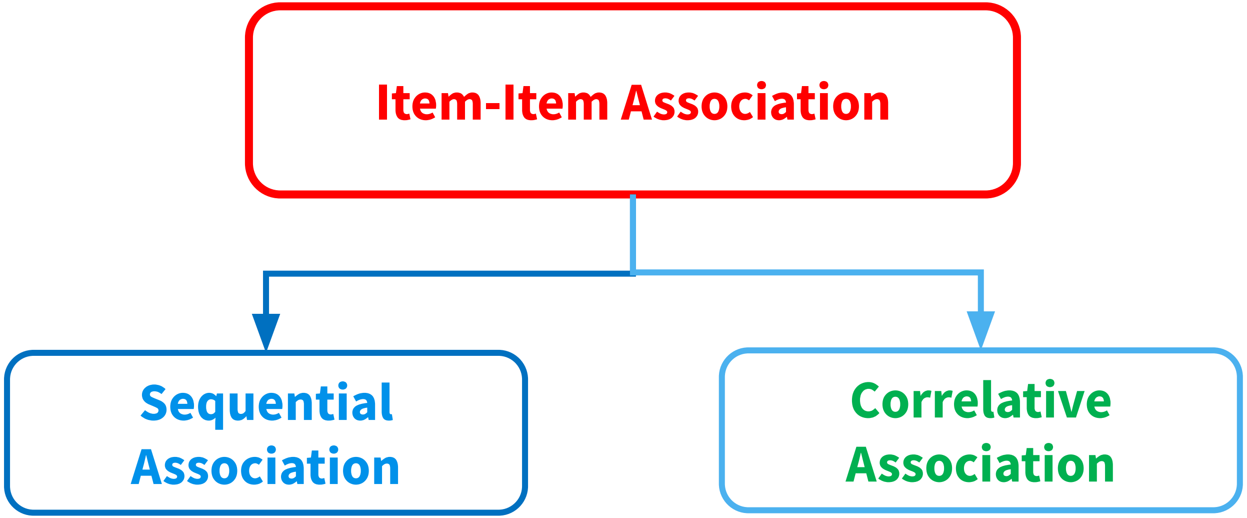
$$\hat{x}_{uij} := \hat{x}_{ui} - \hat{x}_{uj}$$

Main idea: Learn the relative ranking for each user

Ref: Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. BPR: Bayesian Personalized Ranking from Implicit Feedback.

Item-Item Association

Hypothesis: The adoption of a user on an item might be influenced by his past adoptions on other items.



Correlative Association

Data: Basket – Items are adopted at the same time.

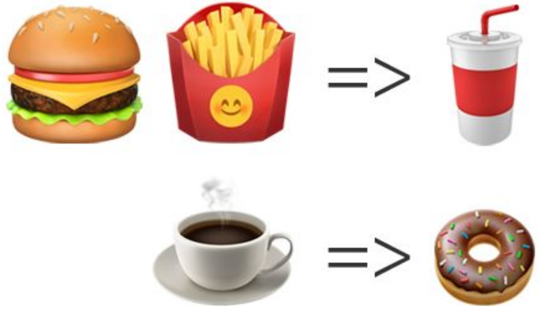
Hypothesis: Items within the same basket are correlated, refers to as *correlative dependencies*.

Association Rule-based Recommendation

instacart

Orders	Products	Users	Timeframe
+3M	+50K	+200K	1 Year

https://www.instacart.com/



Source:
https://www.quora.com/How-is-association-rule-compared-with-collaborative-filtering-in-recommender-systems

An Example of Amazon RS

Frequently Bought Together

Price for all three: $74.20

Add all three to Cart | Add all three to Wish List

Show availability and shipping details

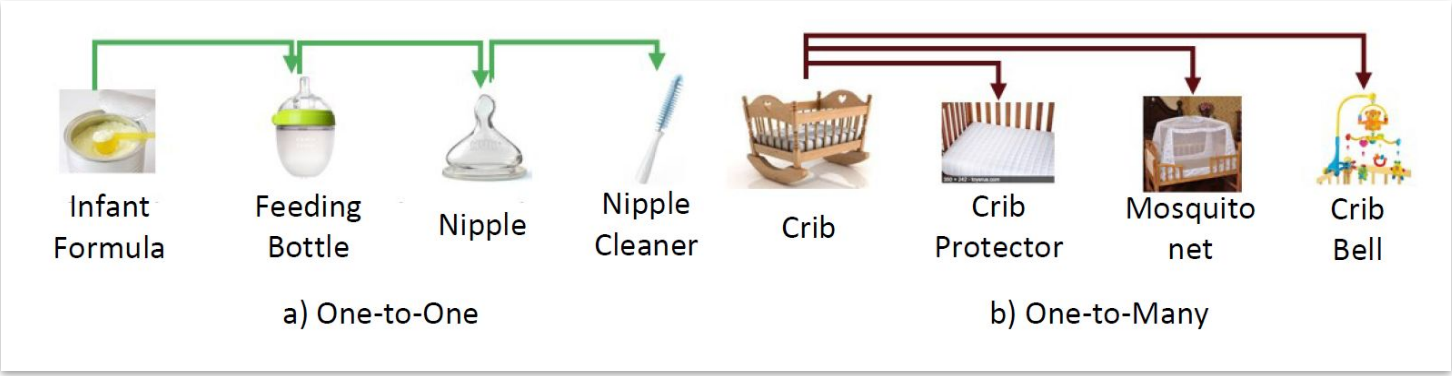
- ☑ **This item:** Beginning Ruby: From Novice to Professional (Expert's Voice in Open Source) by Peter Cooper Paperback $27.78
- ☑ Learn to Program, Second Edition (The Facets of Ruby Series) by Chris Pine Paperback $16.94
- ☑ Ruby on Rails Tutorial: Learn Web Development with Rails (2nd Edition) (Addison-Wesley Professional Ruby ... by Michael Hartl Paperback $29.48

Customers Who Bought This Item Also Bought

Learn to Program, Second Edition (The Facets of...	The Well-Grounded Rubyist	Ruby on Rails Tutorial: Learn Web Development...	The Ruby Programming Language	The Well-Grounded Rubyist
Chris Pine	› David A. Black	› Michael Hartl	› David Flanagan	› David A. Black
42	39	70	74	19
				#1 Best Seller in Ruby Programming Computer
Paperback	Paperback	Paperback	Paperback	Paperback
$16.94 Prime	$32.49 Prime	$29.48 Prime	$26.35 Prime	$29.67 Prime

Sequential Association

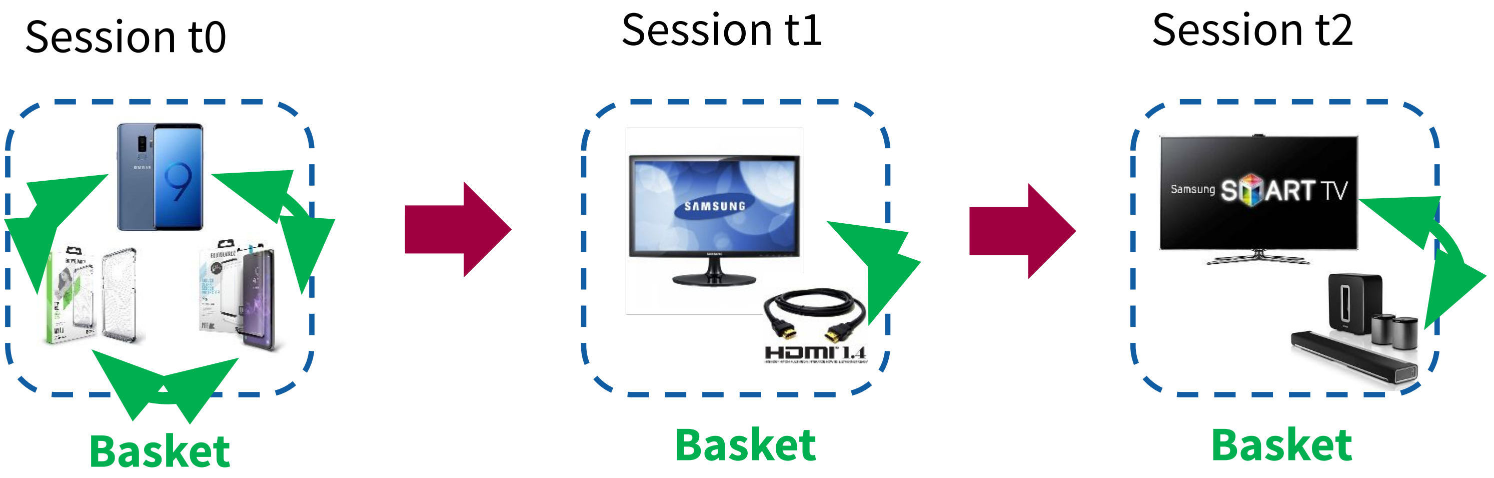
Data: Sequence – Items are adopted sequentially by time.



Hypothesis: The next selection (item) of a user is affected by his preceding adoptions

One more thing …

The notion of Basket Sequence



Correlative associations among items in a basket

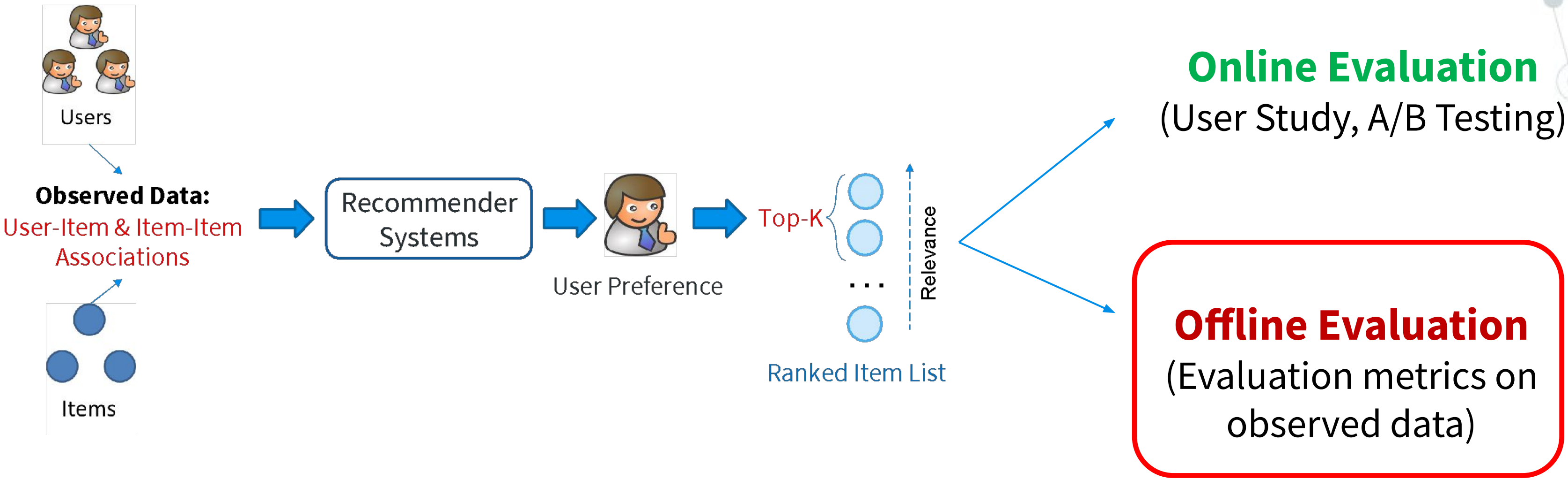
Sequential associations across baskets in a sequence

5.

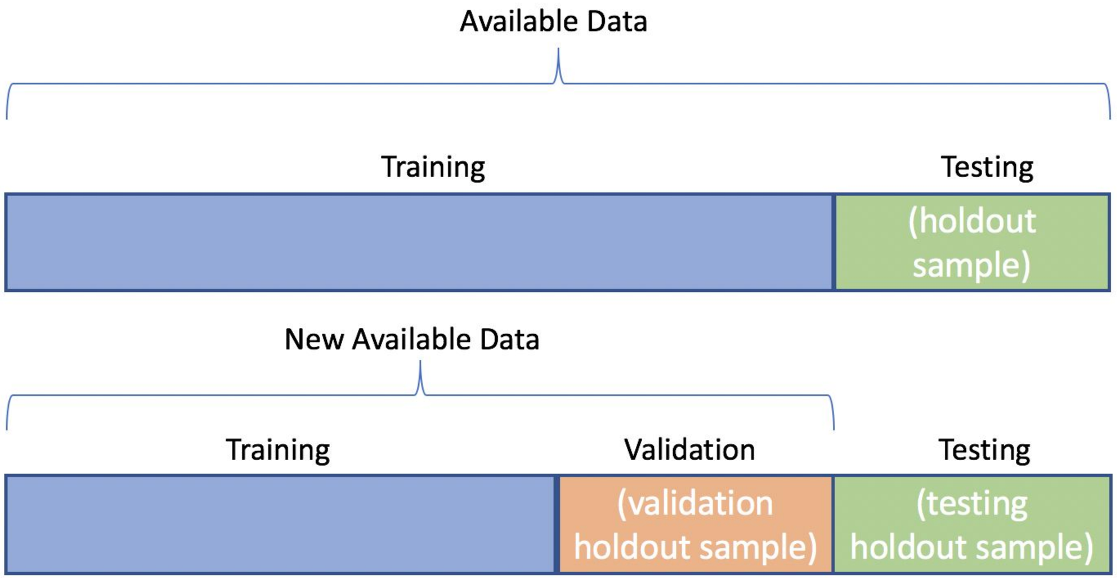
Evaluation and Metrics

“***Recommender systems** are systems that **help users discover items they may like**.*”

Evaluation Strategy



Offline Evaluation



Note: If the available data consists of item/basket sequence, the train/val/test should be split from non-overlapping consecutive periods

Error-based Metrics

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{t=1}^{T}(\hat{y}_t - y_t)^2}{T}}.$$

Recommendation via Rating Prediction

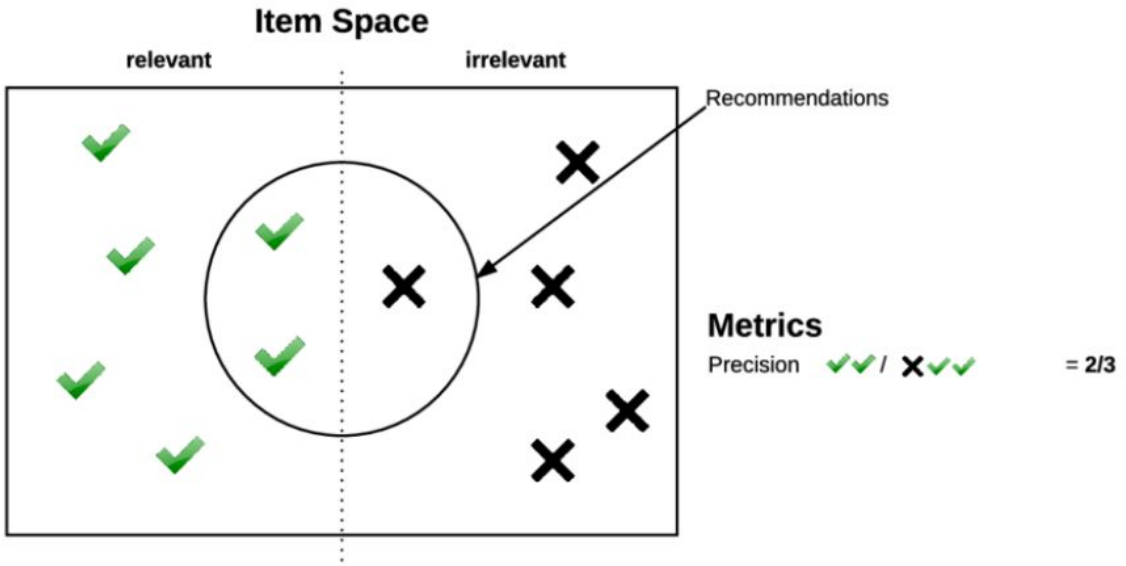
User	Actual	Pedicted	Absolute Error	Squared Error
	1: Toy Story (1995)	1: Toy Story (1995)	1: Toy Story (1995)	1: Toy Story (1995)
755	2	4	2	4
5277	1	2	1	1
1577		3	3	9
4388	2	2	0	0
1202		2	2	4
3823	3	5	2	4
5448		4	4	16
5347	2	3	1	1
4117	4	5	1	1
2765	4	4	0	0

MAE	MSE	RMSE
1.6	4	2.0

Source: https://towardsdatascience.com/an-exhaustive-list-of-methods-to-evaluate-recommender-systems-a70c05e121de

Classification-based Metrics (1)

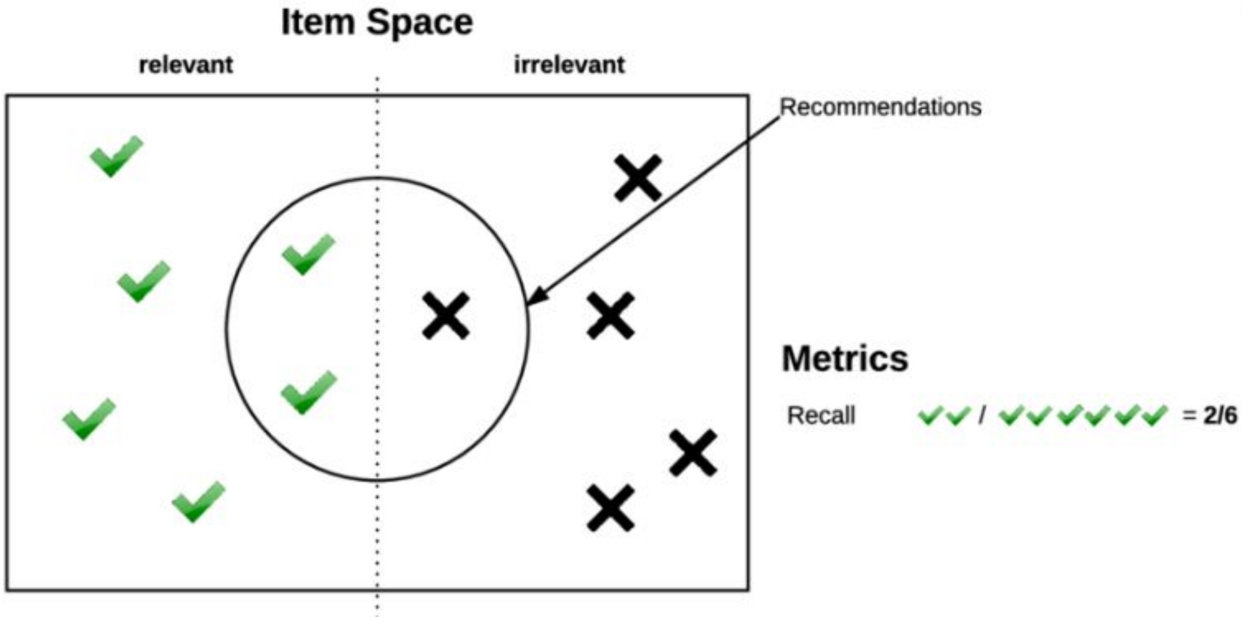
Precision@K for Top-K Recommendation



Source: https://towardsdatascience.com/an-exhaustive-list-of-methods-to-evaluate-recommender-systems-a70c05e121de

Classification-based Metrics (2)

Recall@K for Top-K Recommendation



Source: https://towardsdatascience.com/an-exhaustive-list-of-methods-to-evaluate-recommender-systems-a70c05e121de

Ranking-based Metrics (1)

Normalized Discounted Cummulative Gain (nDCG@K) for Top-K Recommendation

$$\mathrm{nDCG_p} = \frac{DCG_p}{IDCG_p}, \qquad \mathrm{DCG_p} = \sum_{i=1}^{p} \frac{rel_i}{\log_2(i+1)} \qquad \mathrm{CG_p} = \sum_{i=1}^{p} rel_i$$

Items Ranking	Relevancy Score
Movie 1	1
Movie 3	2
Movie 2	2
Movie 5	0
Movie 4	1

Perfect Ranking	Relevancy Score
Movie 3	2
Movie 2	2
Movie 1	1
Movie 4	1
Movie 5	0

CG =	6
DCG =	12.1

CG (p) =	6
DCG (p) =	13.9

nDCG = DCG/DCG (P) =	0.87

Source: https://towardsdatascience.com/an-exhaustive-list-of-methods-to-evaluate-recommender-systems-a70c05e121de

Ranking-based Metrics (2)

Mean Reciprocal Rank (MRR) for Top-K Recommendation

$$MRR = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{rank_i}$$

Items Ranking	Relavent Items	Reciprocol Ranking
Movie 1	No	0
Movie 3	Yes	1/2
Movie 2	Yes	1/3
Movie 5	No	0
Movie 4	Yes	1/5

MRR =	1/2+1/3+1/5 =	1.03

Source: https://towardsdatascience.com/an-exhaustive-list-of-methods-to-evaluate-recommender-systems-a70c05e121de

6.

Introduction to Cornac
A Recommendation Framework

Cornac

Cornac is a comparative framework for multimodal recommender systems, which focuses on making it **convenient** to work with models leveraging **auxiliary data** (e.g., item descriptive text and image, social network, etc.)

Flow of an Experiment in Cornac

Website: https://cornac.preferred.ai/
Tutorials: Github Repo

Summary

- Real-life Examples
- Introduction to Recommendation Systems (RSs)
  - Item, User, Behavior
  - Role of RSs? How do RSs work?
- Types of Recommendation Systems
- Learning Principles
  - Learning User-Item Associations
  - Learning Item-Item Associations
- Evaluation & Metrics
- Introduction to Cornac, a Recommendation Framework

Thanks!

trongld@vnu.edu.vn