



Personalization

A Collaborative Filtering Approach

Here are some movie ratings...













	1	2	3	4
Alice	4	4		1
Bob		2	2	3
Carol	1	5	3	
Dennis	3		4	1
Emma	5	2	1	4
Frank	3	1		5

Predict Alice's Rating for Movie 3!

	1	2	3	4
Alice	4	4	?	1
Bob		2	2	3
Carol	1	5	3	
Dennis	3		4	1
Emma	5	2	1	4
Frank	3	1		5

The YouTube recommendations algorithm is way too reactive. I watched one Jordan Peterson video and this is my home page.

Recommended

 <p>11:47</p>	 <p>JORDAN PETERSON HYPER-INTELLECTUALS 5:12</p>	 <p>TV 3:32</p>	 <p>Access to a scammer's PC Ooooooh Yes 8:52</p>	 <p>PETERSON GENIUS 12:07</p>	 <p>vs. PETERSON MUKAMI 7:09</p>
<p>'His Ideas Are Idiotic' Jordan Peterson DESTROYS Justin</p> <p>Conservative Network 403K views · 1 month ago</p>	<p>Jordan Peterson: Advice for Hyper-Intellectual People</p> <p>PhilosophyInsights 855K views · 8 months ago</p>	<p>Jordan Peterson Dissects the Mind of a Mass Murderer</p> <p>Cheap Virtue 671K views · 1 year ago</p>	<p>Accessing a scammer's PC</p> <p>Jim Browning 2.1M views · 1 year ago</p>	<p>7 Times Jordan Peterson Went Unhinged Genius</p> <p>ScienceNET 458K views · 8 months ago</p>	<p>Jordan Peterson Destroys Islam in 15 Seconds</p> <p>Acts17Apologetics 504K views · 1 month ago</p>
 <p>12:59</p>	 <p>13:48</p>	 <p>9:55</p>	 <p>6:54</p>	 <p>10:49</p>	 <p>4:41</p>
<p>'All White men are R@CIST' Smart Man OWNS Race-</p> <p>50 Stars 159K views · 5 months ago</p>	<p>Jordan Peterson: Milo is a walking Contradiction and He</p> <p>Conservatism 277K views · 1 month ago</p>	<p>Jordan Peterson Destroys Gender Denying Ideologue</p> <p>AustralianRealist 733K views · 2 years ago</p>	<p>Jordan Peterson: My Encounter With Hells Angels</p> <p>Clash of Ideas 282K views · 6 months ago</p>	<p>Leftist Host SNAPS At Jordan Peterson, Instantly</p> <p>Conservative Network 993K views · 1 week ago</p>	<p>How to Easily Overcome Social Anxiety - Prof. Jordan</p> <p>Psyche Matters 908K views · 7 months ago</p>

SHOW MORE



Nakilis 2 years ago

I couldn't agree more. After watching one Tyler Perry interview on Jimmy Fallon, all of my recommendations are now Tyler Perry and Jimmy Fallon related.. And not all of the other content that I watch in ungodly amounts. But sure, Youtube still knows what they're doing.

↑ 1 ↓ Share Report Save



Poenaconda 2 years ago

I once watched ONE video from a creator I enjoy. The next day my ENTIRE recommended was their videos. I understand that YouTube thinks I will watch them but that is insane.

↑ 1 ↓ Share Report Save

有人只是去参加了一次北京婚博会，晚上回到家打开微博和微信，发现信息流广告全部变成了婚纱照、婚庆公司、婚礼礼服等。令他感到恐怖的是在此之前从未在手机进行过结婚相关的任何搜索。这一切发生改变的原因仅仅是因为他去了一次婚博会这个地方而已。

有人在知乎看到除甲醛的相关问题，只是百度了一下，结果连一个美食app都开始推荐除甲醛公司。在百度上打开某理财网站，不到半小时推销电话就打过来了。

有时候你在网上搜了一本小說，然后突然就会有假网站在百度上显示他们网站有这本小说可以下载，然后让百度把他推在首页，你打开链接一看其实里面没有，但是有其他东西的广告。

TECHNOLOGY

Google Knows You Better Than You Know Yourself

Predictive analysis combs through calendars and search histories—and gets in the way of routine self-deception.

JAMES CARMICHAEL AUGUST 19, 2014

Facebook Knows You Better than You Know Yourself



Erman Misirlisoy, PhD Oct 18, 2018 · 7 min read ★



The Internet Knows You Better Than You Know Yourself

When Amazon or eBay recommend us something we like but were not looking for, they effectively know us better than we know ourselves.

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did


Kashmir Hill Former Staff

Welcome to The Not-So Private Parts where technology & privacy collide



QUESTION

How did Target find out a girl
was pregnant?






TARGET HAS GOT YOU IN ITS AIM
COMPANY USES CONSUMER DATA TO ADVERTISE





QUESTION


What data do firms collect
from us?





DATA THAT FITMS COLLECT


Search history: Likewise, almost every search engine collects data on your search histories. Whenever you search on Google, it knows what you are looking for, the items you viewed on Google, the webpages you clicked...





DATA THAT FITMS COLLECT


Geolocation and device information: A GPS and Wi-Fi chip are installed in every smartphone. The dozens of apps on our phones, most of them free, aren't just serving up information and entertainment. They are collecting and selling your data to digital marketers who will then offer you personalized ads.





DATA THAT FITMS COLLECT


Purchase histories: Online sellers frequently collect data on your purchase histories. They know what you have purchased and what you have not purchased. As we will figure out later, the purchase history data is the most important data for online sellers.





DATA THAT FITMS COLLECT

IP addresses: You have to get an IP address to search the internet. You can find your IP address here: <https://whatismyipaddress.com/>. Then, even if you are not using a mobile device, firms can still know your geolocation (e.g., your country and even neighborhood).





DATA THAT FIRMS COLLECT

Firms also collect other types of data. Let's take Google as an example. Do you know what data that Google collects from you?





DATA THAT FITMS COLLECT

Number of email exchanges you've had in Gmail; number of files in Drive; number of photos Google stores for you.

Your location or searches or browsing history: Google Maps keeps track of everywhere you go and when, alongside the photos taken that day and travel times down to the minute.

Your Google Account: your photo and birthdate.







DATA THAT FITMS COLLECT

Even the traditional brick-and-mortar (offline) shops are also collecting your data.

Your payment method (Credit? Mobile pay? Cash?)

Loyalty program information (Are you using Yuu?)

Personal profile (If you ever registered there...)





DATA THAT FITMS COLLECT

With new technologies, brick-and-mortar stores can also get much more information than what they had before.

As described in the video, if you use the free Wi-Fi they provide you, they will be able to collect data from your smartphone!

Facial recognition and mobile payments help collect data from you (“刷臉支付”).




**BUYING FOOD
WITH FACIAL
RECOGNITION**





QUESTION

Suppose that you are an Internet company, and you have access to all this consumer data, what would you do?





How Firms Use Your Data

**Personalized
Pricing**

**Personalized
Recommendation**



Personalized Pricing

With personalized pricing, a seller offers each consumer an individualized price, and two persons can receive two different prices at the same time.

Note that personalized pricing is different from dynamic pricing. With dynamic pricing, the price is changing over time. For personalized pricing, the price is changing over consumers.

Example of dynamic pricing: Uber adjusts prices timely.





Price Discrimination

Broadly speaking, personalized pricing is a form of price discrimination. Let's review types of price discrimination (video [here](#)):

1st degree: The firm sells a product at the maximum price that every consumer is willing to pay.

2nd degree: price varies according to quantity demanded.

3rd degree: charging a different price to different consumer groups.






Personalized Pricing

Personalized pricing is close to first-degree price discrimination.

Firms can learn about your income (e.g., from your bank account), your geo-location (e.g., in the US or India), your neighborhood (a high-end one?), your device (iOS or Android), your purchase habits (bargain hunter?), your gender,...

Based on this information, firms can infer how much you are willing to pay for the product and offer you a personalized price.



Examples

Are you using a Mac or a PC?

THE WALL STREET JOURNAL.

English Edition | Print Edition | Video | Podcasts | [Latest Headlines](#)

[Home](#) [World](#) [U.S.](#) [Politics](#) [Economy](#) [Business](#) **Tech** [Markets](#) [Opinion](#) [Life & Arts](#) [Real Estate](#) [WSJ.M](#)

On Orbitz, Mac Users Steered to Pricier Hotels

On Orbitz, Mac users spend as much as 30% more a night on hotels that PC users do.

Examples

Websites Vary Prices, Deals Based on Users' Information

Getting Different Deals Online
A Journal examination found online retailers adjusted prices by a shopper's location, among other factors

Staples.com
SnapSafe Titan safe
HIGHER PRICE \$1,199.99
DISCOUNT PRICE \$1,099.99
DIFFERENCE 9.1%

Homedepot.com
A 250-foot spool of electrical wiring
Six price-groups, including:
\$70.80 in Ashtabula, Ohio
\$72.45 in Erie, Pa.
\$77.87 in Monticello, NY

RosettaStone.com
RosettaStone
A 20% DISCOUNT
...for buying multiple levels of German lessons, when test-shopping from the U.S. or Canada. But not from the U.K. or Argentina.

Photos: (l to r) SnapSafe Home Depot; Rosetta Stone Source: WSJ testing The Wall Street Journal

MOST POPULAR NEWS

1. What You Can and Can't Do if You've Been Vaccinated: Travel, Risk Factors, What You Need to Know
2. Europe Confronts Covid Rebound as Vaccine Hopes Recede
3. Biden's \$1,400 Stimulus Checks Hit Bank Accounts Starting Today
4. Schumer and Gillibrand Call for Cuomo to Resign

The US retailer *Office Depots* uses customers' browsing history and location data to vary prices

Examples

These Brands Have Some of the Best Abandoned Cart Email Strategies

Aug 28, 2019 5:03:58 PM


When you abandon an item from your online shopping cart, e-tailers may issue you a discount to lure you to make a purchase.



Behavior-Based Pricing

The more common approach is pricing with consumers' purchase history, a practice known as “**behavior-based pricing**”.


The idea is very simple: The price you receive depends on whether or not you have purchased the products before. In other words, we offer new and existing consumers different prices.





QUESTION

Suppose that a firm uses “behavior-based pricing”, how should the firm charge its prices? Should the firm offer new consumers a higher or lower price?



Examples

Amazon's old customers 'pay more'



Some Amazon customers are refusing to accept some DVD prices

In 2000, behavior-based pricing first appeared to the public. You can find the link to this phenomenon [here](#).

Examples

This is also evidence that airlines offer higher price to frequent travelers.

TRAVEL

Airfare Expert: Do cookies really raise airfares?

Rick Seaney, special for USA TODAY

Published 5:00 a.m. ET Apr. 30, 2013

Examples

In China, this is a very vivid description of this kind of behavior, i.e., “杀熟” --- “**killing existing consumers.**”

同样的订单，同一家外卖平台、同一家商户、同一处送餐地址、同一个时间段，会员却比非会员支出更多——近日，有网民几次测试发现，在注册成为美团会员后，相比非会员，外卖满减优惠力度不仅有所降低，配送费也不减反增。此事再次引发舆论对互联网平台“杀熟”现象的强烈关注。

After becoming a member of Meituan, an online food delivery platform in China, you will have to a higher price and receive lower price discount.



QUESTION

In most cases, firms offer high prices to existing consumers and lower prices to new consumers. But why?






Behavior-Based Pricing

The rationale is as follows. Consumers who really like the product will make the purchase early. So, compared to new consumers, existing consumers are likely to be fans of the product and are willing to pay a higher price for it.

Following the logic, the firm can take advantage of this and charges existing consumers a higher price, i.e., punishing existing consumers (杀熟).





Is It Legal?

While consumers often object to personalized pricing, it is legal in most countries.

In 1996, a consumer living in Manhattan sued Victoria's Secret for distributing different versions of catalogs with identical items but different prices. However, the New York Court dismissed the claim by noting that it was an accepted business practice to reward repeat consumers or to draw in new consumers with special savings.





Is It Legal?

Any form of price discrimination is legal in the United States, as long as the basis of discrimination is not race, religion, national origin, gender, and the like.

Recently, China banned behavior-based pricing in the traveling and hospitality industry. According to a 2020 regulation by the Ministry of Culture and Tourism, online traveling website is not allowed to offer consumers discriminated prices (see [news](#) here).

In the EU, there is a recent GDPR regulation on big data.






EU's GDPR regulation



QUESTION

As an individual consumer, do you like personalized pricing? What should you do when you know firms are using personalized pricing?





How Firms Use Your Data

**Personalized
Pricing**

**Personalized
Recommendation**

Recommendation is everywhere

amazon.com

Recommended for You

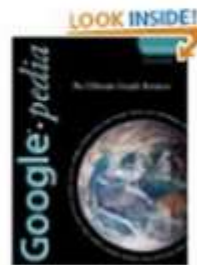
Amazon.com has new recommendations for you based on [items](#) you purchased or told us you own.



[Google Apps Deciphered: Compute in the Cloud to Streamline Your Desktop](#)



[Google Apps Administrator Guide: A Private-Label Web Workspace](#)



[Googlepedia: The Ultimate Google Resource \(3rd Edition\)](#)

Recommendation is everywhere

us house of representatives

myFINANCE
Did you know?
Two Savings Accounts
That Pay 10x What
Your Bank Pays.
Save More

Up next

(Part II) A Day in the Life of
Arizona Rancher: Fences, II
Center for Immigration Studie
43K views

13:29

CMJ
10:06

Obama promises to pass Trump
You promised Mexico would
CNN
2.7M views
New

People Are Fleeing President
Trump's America To This T
NBC News
308K views
7:23

Scrambling onto trucks for
better life
Sky News
2.4M views
3:58

Polar Bear vs Walrus colon
BBC Planet Earth | BBC Stu
BBC Studios
Recommended for you

WSJ

Arizona Border Ranchers Torn in Support for Trump's Wall

172,275 views

685 249 SHARE SAVE

Wall Street Journal
Published on Mar 16, 2017

SUBSCRIBE 1.2M

Despite enthusiastic backing for President Donald Trump and pleas for a stronger border, Arizona ranchers are conflicted in their support for Trump's promise to build a wall along the border with Mexico. Photo/Video: Jake Nicol/The Wall Street Journal

SHOW MORE

Recommendation is everywhere





The Importance of Recommendation

Netflix: 2/3 of the movies watched are recommended.

Google News: recommendations generate 38% more click-throughs.

Amazon: 35% sales from recommendations.


ChoiceStream: 28% of the people would buy more music if they found what they liked.





QUESTION

Have you ever thought about how these platforms and APPs make recommendations to you? Any idea?





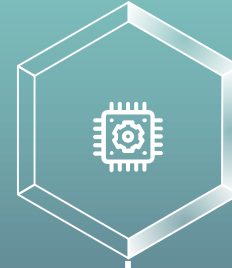
Items

movies, songs,
products, etc. (often
many thousands)



Users

watchers, listeners,
purchasers, etc.
(often many
millions)



Feedback

5-star ratings, not-
clicking 'next', purchases,
etc.

Collaborative Filtering

Everyday examples of collaborative filtering:

Bestseller lists, Top 40 music lists, The “recent returns” shelf at the library, “Read any good books lately?”

The intuition behind collaborative filtering: **personal tastes are correlated**

If Alice and Bob both like X and Alice likes Y, then Bob is more likely to like Y – especially (perhaps) if Bob knows Alice

Question



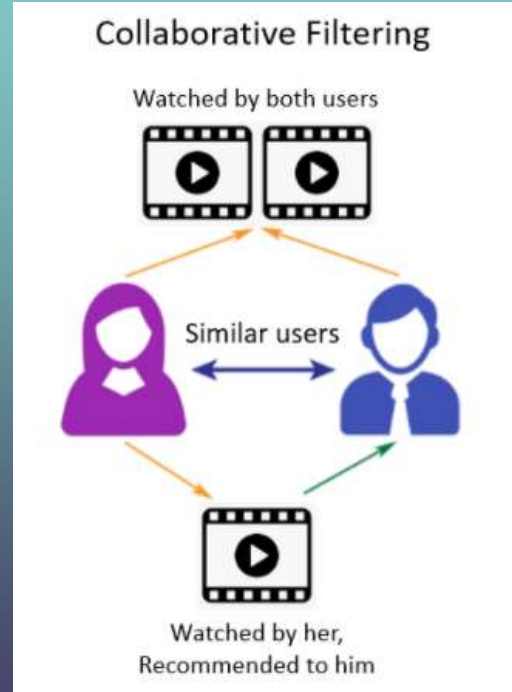
The image shows a grid of movie posters and user avatars. The posters are: *Reservoir Dogs*, *Batman Returns*, *Mickey Blue Eyes*, *Rainbow*, *The Untouchables*, and *Boyz n the City*. The avatars are: a man in a suit, a woman with blonde hair, a man in a suit, a man with glasses, a man with a target on his chest, a man in a blue shirt, and a man in a green shirt. The grid contains the following ratings:

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

Suppose that you want to recommend a movie to this person. Which movie would you recommend? Why?

Collaborative Filtering

In collaborative filtering, we make recommendation to one user based on the preference of similar users.





**User-Based
Collaborative
Filtering**

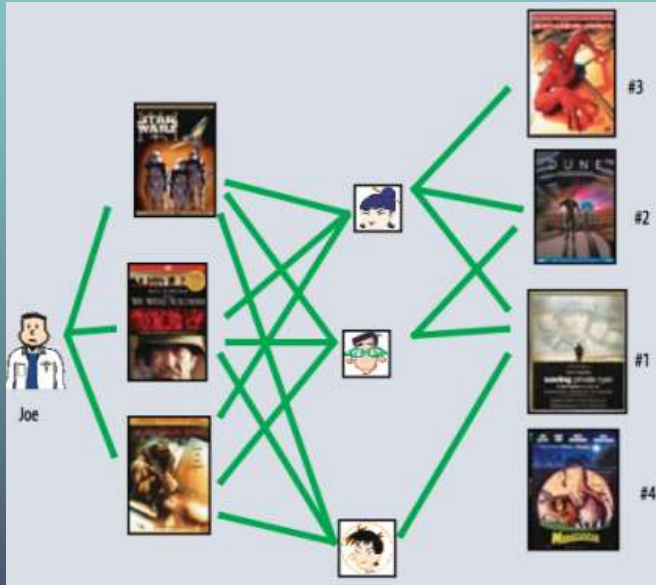


**Item-Based
Collaborative
Filtering**



**Model-Based
Collaborative
Filtering**

Neighborhood Method



In the figure, assume that a green line indicates the movie was watched.

Algorithm:

1. **Find neighbors** based on similarity of movie preferences
2. **Recommend** movies that those neighbors watched

User-Based Collaborative Filtering



	Ramona and Beezus	Date Night	Mickey Blue Eyes	The Rainbow	The Sandlot	The Hot Chick
User 1	2			4	5	
User 2	5		4			1
User 3			5		2	
User 4		1		5		4
User 5			4		2	
User 6	4	5		1		

Each user has reviewed some items, but not every item.

We want to know their preferences for the unrated items.


User-Based Collaborative Filtering



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

Suppose that you want to understand this specific user's preferences.

User-Based Collaborative Filtering

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

Identify items that have been rated by this user.

User-Based Collaborative Filtering

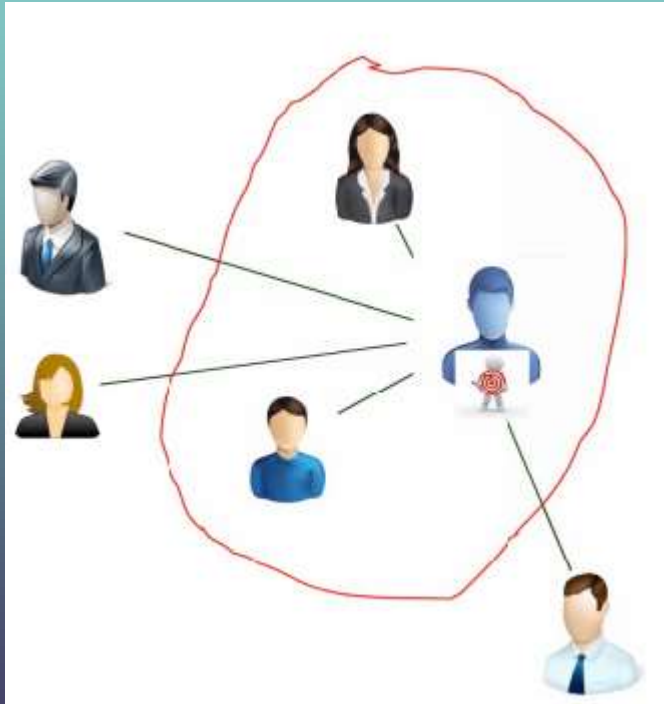


	Toy Story	Date Night	Mickey Blue Eyes	The Rainbow Connection	The Matrix	The Bourne Supremacy
User 1	2			4	5	
User 2	5		4			1
User 3			5		2	
User 4		1		5		4
User 5			4			2
User 6	4	5		1		

Identify items that have been rated by this user.

Identify other users that have rated the same items.

User-Based Collaborative Filtering



Compute how similar each neighbor is to the target user (similarity function). This is usually done by calculating the correlation between their ratings.

In case, select k most similar neighbors.

Make predictions based the similar neighbors' preferences.



**User-Based
Collaborative
Filtering**



**Item-Based
Collaborative
Filtering**




**Model-Based
Collaborative
Filtering**



Item-Based Collaborative Filtering

The idea is very similar to user-based collaborative filtering.

1. Identify set of users who rated the target item.
 2. Identify which other items (neighbors) were rated by the users set.
 3. Compute similarity between each neighbor & target item.
 4. In case, select k most similar neighbors.
 5. Predict ratings for the target item.
- 



**User-Based
Collaborative
Filtering**



**Item-Based
Collaborative
Filtering**



**Model-Based
Collaborative
Filtering**



Matrix Factorization Method

Here, we assume that each individual and each movie has some “latent factors”. For movies, these factors can measure dimensions such as **comedy versus drama, amount of action, or orientation to children; depth of character development or quirkiness, ...**

Each user has his or her preference for the factors and each movie has its value on each of these factors.



Matrix Factorization Method

Let us consider a very simple example. Suppose that there are two factors, amount of action (X) and seriousness (Y). A user also has her preference for action β_X and preference for seriousness β_Y .

When a movie has a large X , it means the movie has more actions, and when a movie has a large Y , it means the movie is more serious. Similarly, if β_X is large, it means the user prefers more actions in the movie.

Matrix Factorization Method

Then, if we know X, Y, β_X, β_Y , we can predict the user's preference for the movie, which is given by

$$\text{Preference score} = \beta_X X + \beta_Y Y$$

And we should recommend movies with the highest preference score.

Rating Matrix as a product of its factors



1

1

F1

0

1

F2

F1

F2



3

2

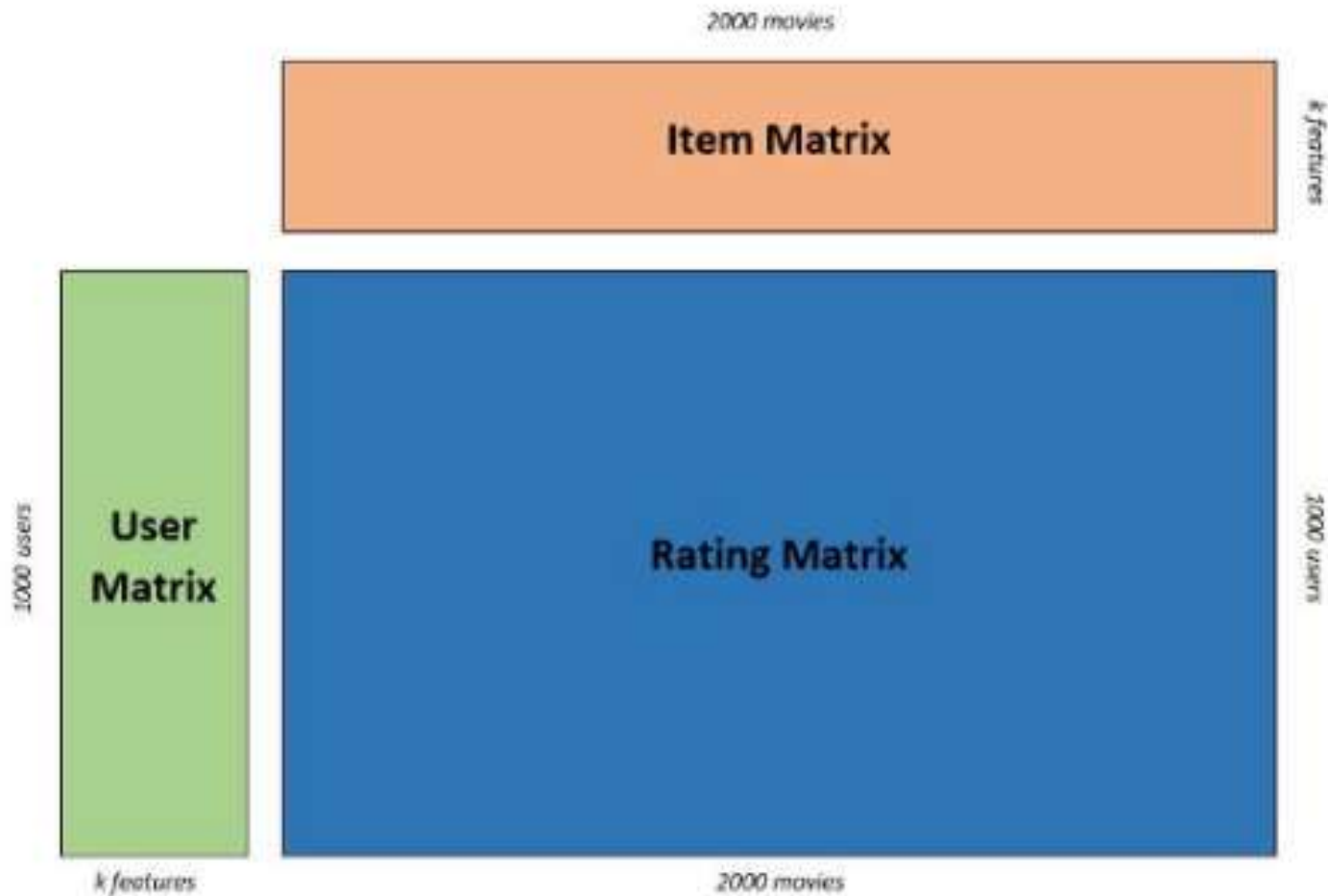
3

5

DOT PRODUCT

$(3 \times 1) + (2 \times 0)$

$(3 \times 1) + (2 \times 1)$



Matrix Factorization Method

	Doctor Strange	Star Trek: Beyond	Zootopia
Alice	1		5
Bob	3	4	
Charlie	3	5	2

Usually, each user has only watched or rated a few movies.

So, the entire rating matrix has a lot of missing values.

We want to fill these missing values.




Matrix Factorization Method

Based on the data that we already have (i.e., existing ratings from users), we can decompose the rating matrix into the user matrix and the movie matrix.

How to decompose? One approach is to minimize the sum of squares of errors like we do in linear regression.

Then, we can multiple these two matrices to predict a user's preference for an unwatched movie.






Matrix Factorization Method

In sum, based on what users have already watched, we can infer the user's preference for various movie attributes.

In addition, based on the ratings from the users who have watched the movie, we can infer the movie's attributes.

Finally, based on the user's preferences and the movie's attributes, we can predict a user's preference for this movie.



Matrix Factorization Method

We can compare it with linear regression:

In linear regression, we infer the value of α and β , and then we can use the regression formula $Y = \alpha + \beta X$ to make predictions.

In matrix factorization, we infer both β_X and β_Y for each user, and X and Y for each movie, and use the formula $\beta_X X + \beta_Y Y$ to predict the user's preference for the movie.

Predict Alice's Rating for Movie 3!

	1	2	3	4
Alice	4	4	?	1
Bob		2	2	3
Carol	1	5	3	
Dennis	3		4	1
Emma	5	2	1	4
Frank	3	1		5

Predict Alice's Rating for Movie 3!

	1	2	3	4
Alice	4	4	≈ 2.82	1
Bob	≈ 2.72	2	2	3
Carol	1	5	3	≈ 1.97
Dennis	3	≈ 5.70	4	1
Emma	5	2	1	4
Frank	3	1	≈ 0.54	5



Summary Video

