

Master on Foundations of Data Science



Recommender Systems

Course Presentation



Associate Professor at the Department of Mathematics and Computer Science
from the University of Barcelona.

PhD in 2011 on Computer Vision and Machine Learning
Graduate on Computer Science in 2007

Santi Seguí Email: santi.segui@ub.edu

Who are you?

Why recommender systems?



“We are leaving the information age, and
entering into the recommendation age”

–Chris Anderson, from the Long Tail Book
- 2008

The Amazon.com logo, featuring the word "amazon.com" in white lowercase letters with a registered trademark symbol, and a yellow curved arrow underneath.

YOU MAY ALSO LIKE



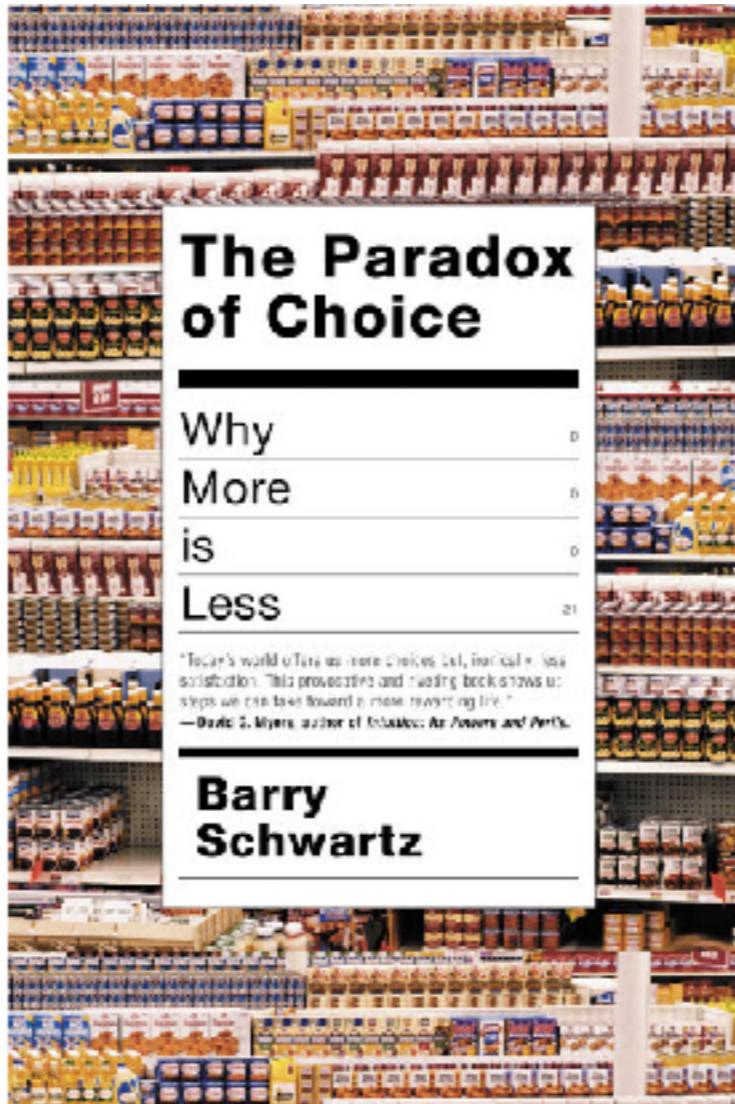
*“Judging by Amazon’s success, the recommendation system works. The company reported a 29% sales increase to \$12.83 billion during its second fiscal quarter, up from \$9.9 billion during the same time last year. A lot of that growth arguably has to do with **the way Amazon has integrated recommendations** into nearly every part of the purchasing process...”*

2012

“A lot of times, people doesn’t know what they want until you show it to them...”

–Steve Jobs, 1997

Information overload



(offering) more **choice** can sometimes mean fewer sales

**MOST OF THE
TIME WE DO NOT
WANT TO
CHOOSE**

“In 2015, user consumption will raise
to 74GB a day”

—UCSD Study, 2014

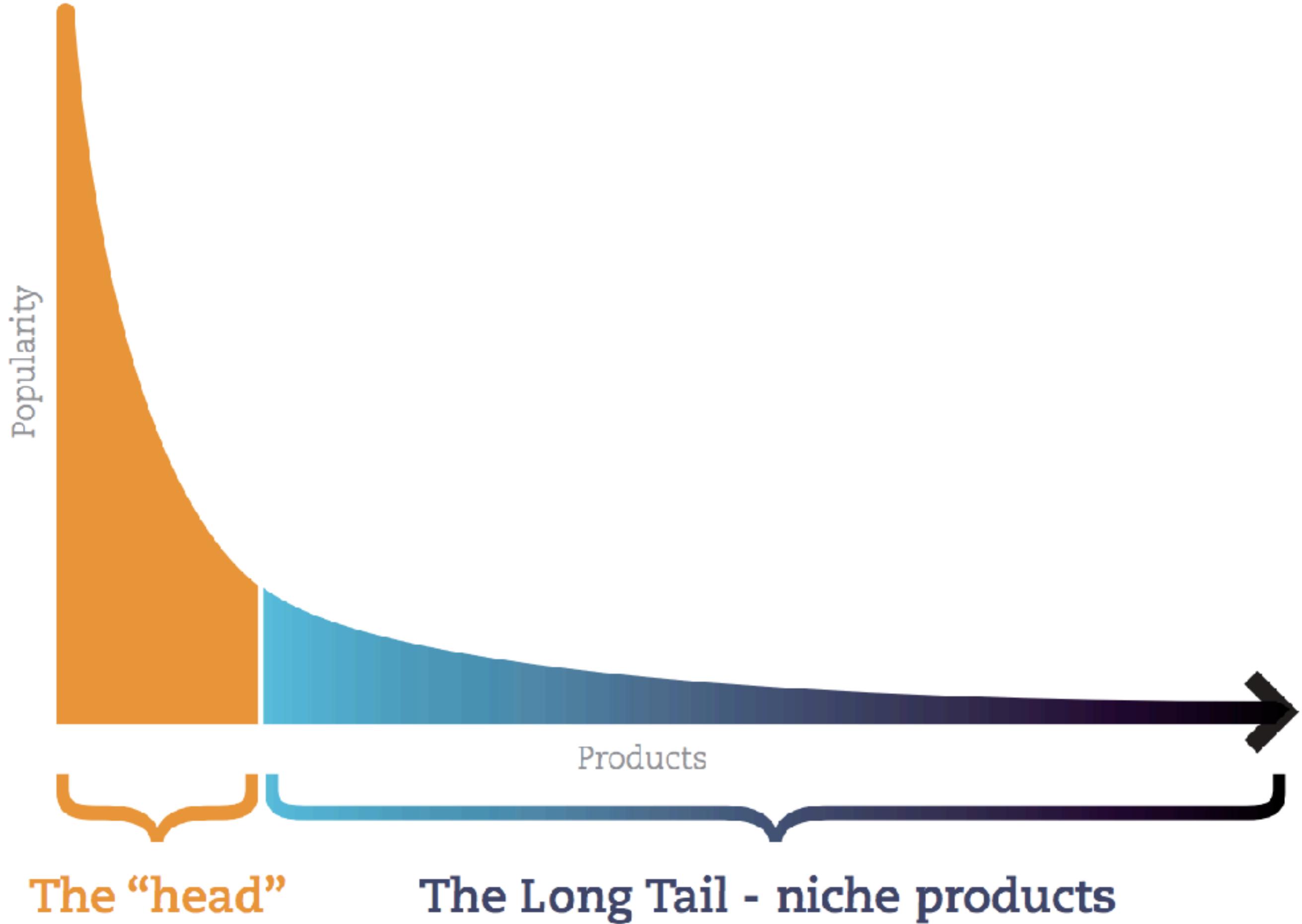
Today:

on facebook: More than 300 million photos get uploaded per day

on Instagram: Each day 95 million photos and videos are shared on Instagram

Why recommender systems?

- Some of the various potential benefits of recommendation systems in business, and the companies that use them:
 - “**Improving with use**” (**retention**)
 - Ability to continuously calibrate to the preferences of the user. You’re much less likely to switch to a Netflix competitor when Netflix has such a wonderful sense of which movies and shows you might want to watch next
 - **Improving cart value**
 - Help the companies to decide what to offer (to have in their catalog)
 - **Improved engagement and delight**
 - Long-term or short term optimization. YouTube doesn’t optimize for short-term view length, as this might encourage pushy or flashy tactics that wouldn’t genuinely delight users. Instead, the service aims to encourage long-term use, because advertising views is the ROI that these systems serve at YouTube



Recommender Systems

Master on Foundations of Data Science



Goal of the course:

Learn the basics on recommender systems, and finally become an expert on this topic

What do we expect from the students?

Active participation in class
Some work at home

Course Agenda - 2020

Thursday	
13/02	Introduction to Recommender Systems
20/02	Non-Personalized Recommenders
27/03	Collaborative-Based Recommenders
05/03	Collaborative-Based Recommenders
12/03	Dimensionality Reduction for Recommender Systems
19/03	Content-Based Recommender Systems *
26/03	RecSys Challenge *
02/04	RecSys Challenge *
09/04	Easter BREAK
16/04	Evaluation/ Graph Based Models *
23/04	SANTI JORDI
30/05	Deep Learning Models
07/05	Context Based Models
14/05	Group Based Models / Knowledge Recommendations
21/05	Current Practices in Industry and Research
28/05	Exam

Course Evaluation

Assignment #1 : 25% - Classical Rec System

Assignment #2: 25% - Graph / DL Rec Systems

Presentation #3: 25%

Final Exam : 25%

A bit of History

- 1985 - First ideas of recommender systems were mentioned
- 1992 - Tapestry by Xerox Palo Alto
 - First system designed by collaborative filtering
- 1994 - GroupLens
 - First recommender system using rating data
- 1997 - MovieLens
 - First movie recommender system

Who uses Recommender systems?

Who uses Recommender systems?



m o v i e l e n s
helping you find the *right* movies

last.fm™
the social music revolution

Google™
News

YouTube

XBOX
LIVE



Telefónica
Investigación y Desarrollo

BBVA
DATA & ANALYTICS

Santander

Recommender systems
are everywhere

Prizes were based on improvement over Netflix's own algorithm, called Cinematch, or the previous year's score if a team has made improvement beyond a certain threshold. A trivial algorithm that predicts for each movie in the quiz set its average grade from the training data produces an RMSE of 1.0540. Cinematch uses "straightforward statistical linear models with a lot of data conditioning"





Watch Instantly

Browse DVDs

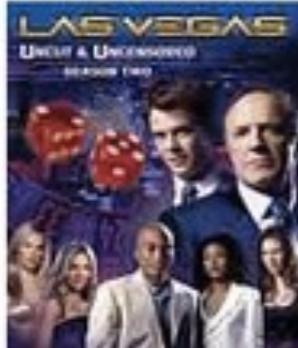
Your Queue

Movies You'll ❤️

Congratulations! Movies we think You will ❤️

Add movies to your Queue, or Rate ones you've seen for even better suggestions.

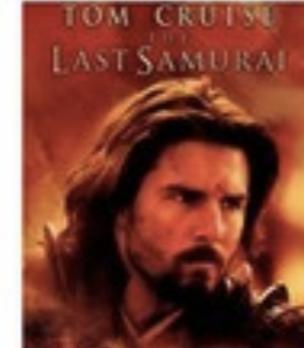
Spider-Man 3

[Add](#)Las Vegas: Season 2
(6-Disc Series)

300

[Add](#)

The Last Samurai



The Rundown

[Add](#)

Bad Boys II

[Add](#)

Star Wars: Episode III

Robot Chicken: Season 3
(2-Disc Series)

IMDb Charts

Top Rated Movies

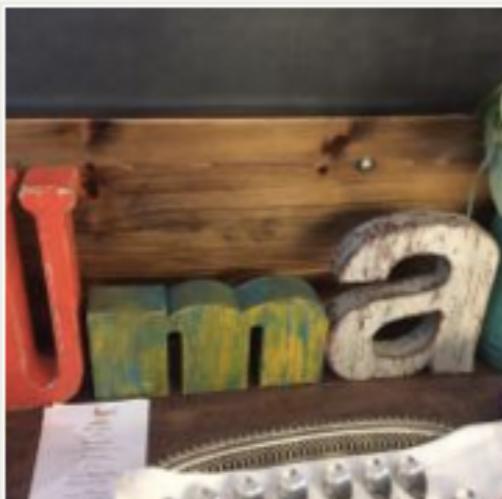
Top 250 as voted by IMDb Users

Showing 250 Titles

Sort by: Ranking



Rank & Title	IMDb Rating	Your Rating	
1. Cadena perpetua (1994)	★ 9,2	☆	
2. El padrí (1972)	★ 9,2	☆	
3. El padrí II (1974)	★ 9,0	☆	
4. El caballero oscuro (2008)	★ 8,9	☆	
5. Pulp Fiction (1994)	★ 8,9	☆	
6. La llista de Schindler (1993)	★ 8,9	☆	
7. 12 hombres sin piedad (1957)	★ 8,9	☆	
8. El senyor dels anells: El retorn del rei (2003)	★ 8,9	☆	



Uma

1 de 7.365 Restaurantes en Barcelona

●●●●● 227 opiniones

"Probarlo" 08/01/2016

"Creatividad en estado puro" 05/01/2016

Precio: 65 € - 75 € | Mapa | Fotos de los visitantes (345)

Cocina:

[Internacional](#)

[Mediterránea](#)

[Fusión](#)

[Española](#)



Tast-Ller

2 de 7.365 Restaurantes en Barcelona

●●●●● 324 opiniones

"COMPARTIENDOLO CON AMIGOS" 13/01/2016

"Excelente" 15/12/2015

Precio: 50 € - 60 € | Mapa | Fotos de los visitantes (312)

Cocina:

[Delicatessen](#)



Fulla d'Ostra

3 de 7.365 Restaurantes en Barcelona

●●●●● 166 opiniones

"Exquisito" 30/12/2015

"Excelente menú" 23/12/2015

Precio: 80 € - 100 € | Mapa | Fotos de los visitantes (75)

Cocina:

[Fusión](#)

[Mediterránea](#)

[Española](#)



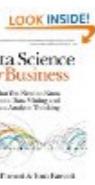
Informal

4 de 7.365 Restaurantes en Barcelona

●●●●● 215 opiniones



Frequently Bought Together



Price for both: \$158.15

Add both to Cart

Add both to Wish List

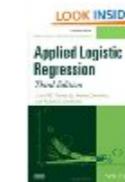
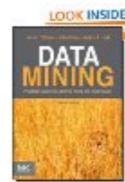
One of these items ships sooner than the other. [Show details](#)

This item: **Introduction to Data Mining** by Pang-Ning Tan Hardcover \$120.16

Data Science for Business: What you need to know about data mining and data-analytic thinking by Foster Provost Paperback \$37.99

Customers Who Bought This Item Also Bought

Page 1 of 15



[Data Science for Business: What you need...](#)
by Foster Provost
 102
#1 Best Seller in Data Mining
Paperback
\$37.99 

[Data Mining: Practical Machine Learning Tools...](#)
by Ian H. Witten
 52
Paperback
\$40.65 

[Data Mining: Concepts and Techniques, Third...](#)
by Jiawei Han
 28
Hardcover
\$60.22 

[Regression Analysis by Example](#)
by Samprit Chatterjee
 9
Hardcover
\$92.39 

[SAS Statistics by Example](#)
by Ron Cody
 10
Perfect Paperback
\$44.37 

[Applied Logistic Regression](#)
by David W. Hosmer Jr.
 9
Hardcover
\$62.33 

[An Introduction to Statistical Learning:...](#)
by Gareth James
 56
#1 Best Seller in Mathematical & Statistical...
Hardcover
\$72.79 

Frequently Bought Together



Price for both: \$158.15

Add both to Cart

Add both to Wish List

One of these items ships sooner than the other. [Show details](#)

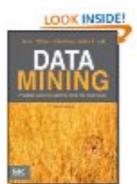
This item: Introduction to Data Mining by Pang-Ning Tan Hardcover \$120.16

Data Science for Business: What you need to know about data mining and data-analytic thinking by Foster Provost Paperback \$37.99

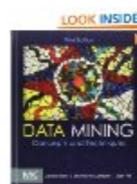
Customers Who Bought This Item Also Bought



Data Science for Business: What you need...
by Foster Provost
 102
#1 Best Seller in Data Mining
Paperback \$37.99 



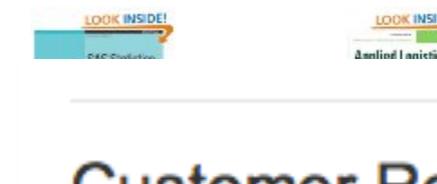
Data Mining: Practical Machine Learning Tools...
by Ian H. Witten
 52
Paperback \$40.65 



Data Mining: Concepts and Techniques, Third...
by Jiawei Han
 28
Hardcover \$60.22 



Regression Analysis by Example
by Samprit Chatterjee
 9
Hardcover \$92.39 

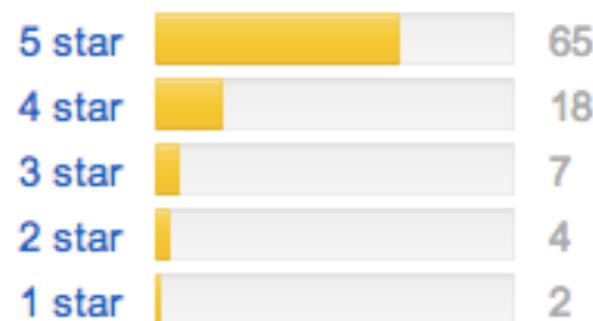


Introduction to Data Mining:...
James
 56
Seller in Mathematical & Statistical...


Customer Reviews

 (96)

4.5 out of 5 stars



[See all 96 customer reviews](#) ▶

Page 1 of 15



Introduction to Data Mining

\$120.16 FREE Shipping. | Temporarily out of stock. Order now and we'll deliver when available. We'll e-mail you w

What Other Items Do Customers Buy After Viewing This Item?



[Data Science for Business: What you need to know about data mining and data-analytic thinking](#) Paperback

› Foster Provost

102

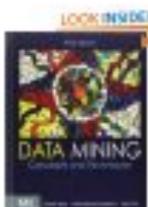
\$37.99



[Introduction to Data Mining](#) Paperback

Pang-ning Tan

4



[Data Mining: Concepts and Techniques, Third Edition \(The Morgan Kaufmann Series in Data Management Sy](#)

› Jiawei Han

28

\$60.22



[Data Mining: Practical Machine Learning Tools and Techniques, Third Edition \(The Morgan Kaufmann Series i](#)

› Ian H. Witten

52

\$40.65

People You May Know



Pixel Park

6 mutual friends

+1 Add Friend



Nolan Bushnell

8 mutual friends

+1 Add Friend



Juan Carlos Anorga

16 mutual friends

+1 Add Friend



Jiwoong Lee (Fragile)

31 mutual friends

+1 Add Friend



Ana Milat

+1 Add Friend



Kate Roberts

+1 Add Friend

See All

We created Pandora to put the [Music Genome Project](#) directly in your hands

**It's a new kind of radio –
stations that play only music you like**

Enter artist, genre or composer to create a station

Already have a Pandora account? [Sign in](#)

Company

[About](#)
[Advertising](#)
[Submit Your Music](#)
[Blog](#)
[Jobs](#)
[Investor](#)
[Press](#)
[Privacy](#)
[Your CA Privacy Rights](#)
[Terms](#)
[Ad Preferences](#)

Music Genre Stations

[Today's Hits](#)
[Today's Country](#)
[Today's Hip Hop And Pop Hits](#)
[Today's R&B And Hip Hop Hits](#)
[Contemporary Gospel](#)
[Classic Rock](#)
[Classical](#)
[Browse Genre Stations](#)

Pandora Everywhere

[Pandora On Your Phone](#)
[Pandora In Your Car](#)
[Pandora In Your Home](#)
[Pandora For Your Business](#)

© 2005-2014 Pandora Media, Inc. All Rights Reserved.

Pandora and the Music Genome Project are registered trademarks of Pandora Media, Inc.

Jobs You May Be Interested In

Powered by
LinkedIn



THOMSON REUTERS

Senior Data Analyst Job

Thomson Reuters - Bangalore, KA



Data Scientist/ Senior Data Scientist

HeadHonchos.com - Bangalore - IN



Adobe

Hiring Computer Scientist (Java) for...

Adobe - Noida



People You May Know



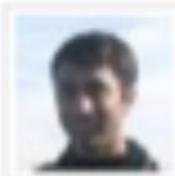
Rajeev Gupta, MBA Candidate '13 at Kelley School of Business

Connect



Alex Vauthey, Sr. Director of Engineering at LinkedIn

Connect



Bikram Bhatia, Graduate Research Assistant at University of Illinois at

Connect

[See more »](#)

What to Watch

My Subscriptions

Music

Recommended



**Steve Jobs at the
Antennagate Press...**
by EverySteveJobsVideo
18,990 views • 9 months ago



**Last Week Tonight with John
Oliver: Online Harassment...**
by LastWeekTonight
2,895,113 views • 1 week ago



**BASTILLE feat. Ella - No
Angels (HD)**
by dancenationedm
5,836,404 views • 2 years ago



Bill Maher with Charlie Rose
by DarkLordDabc
88,923 views • 6 months ago



**READING MEAN COMMENTS
w/ JAMES CORDEN // Grac...**
by Grace Helbig
733,926 views • 3 months ago



**The Attack That Could Disrupt
The Whole Internet -...**
by Computerphile
345,806 views • 1 year ago



**START! Walking at Home
American Heart Associa...**
by Leslie Sansone's Walk at Home
1,863,839 views • 7 months ago



**The 5 Best Shia Labeouf "DO
IT" Motivation Videos**
by The VGC
1,390,971 views • 4 weeks ago



**Elon Musk and Bill Gates
discuss AI, entrepreneurship...**
by Every Elon Musk Video
105,539 views • 2 months ago



**Hadfield demonstrates
Microbial Air Sampling on t...**
by Canadian Space Agency
281,611 views • 2 years ago

Show more





Send Save Now Discard Draft autosaved at 6:01 PM (1 minute ago)

To: "Dopey" <hikingfan@gmail.com>, "Grumpy" <farnsworthpt@gmail.com>, "Doc" <surfingfan@gmail.com>

Also include: [Happy](#), [Sneezy](#), [Sleepy](#)

[Add Cc](#) | [Add Bcc](#)

Subject: New in Labs: Suggest more recipients

[Attach a file](#) [Add event invitation](#)

Market Basket



Market Basket Example



- Where should detergents be placed in the Store to maximize their sales?
- Are window cleaning products purchased when detergents and orange juice are bought together?
- Is soda typically purchased with bananas? Does the brand of soda make a difference?
- How are the demographics of the neighborhood affecting what customers are buying?

Image source: deepclimate.org

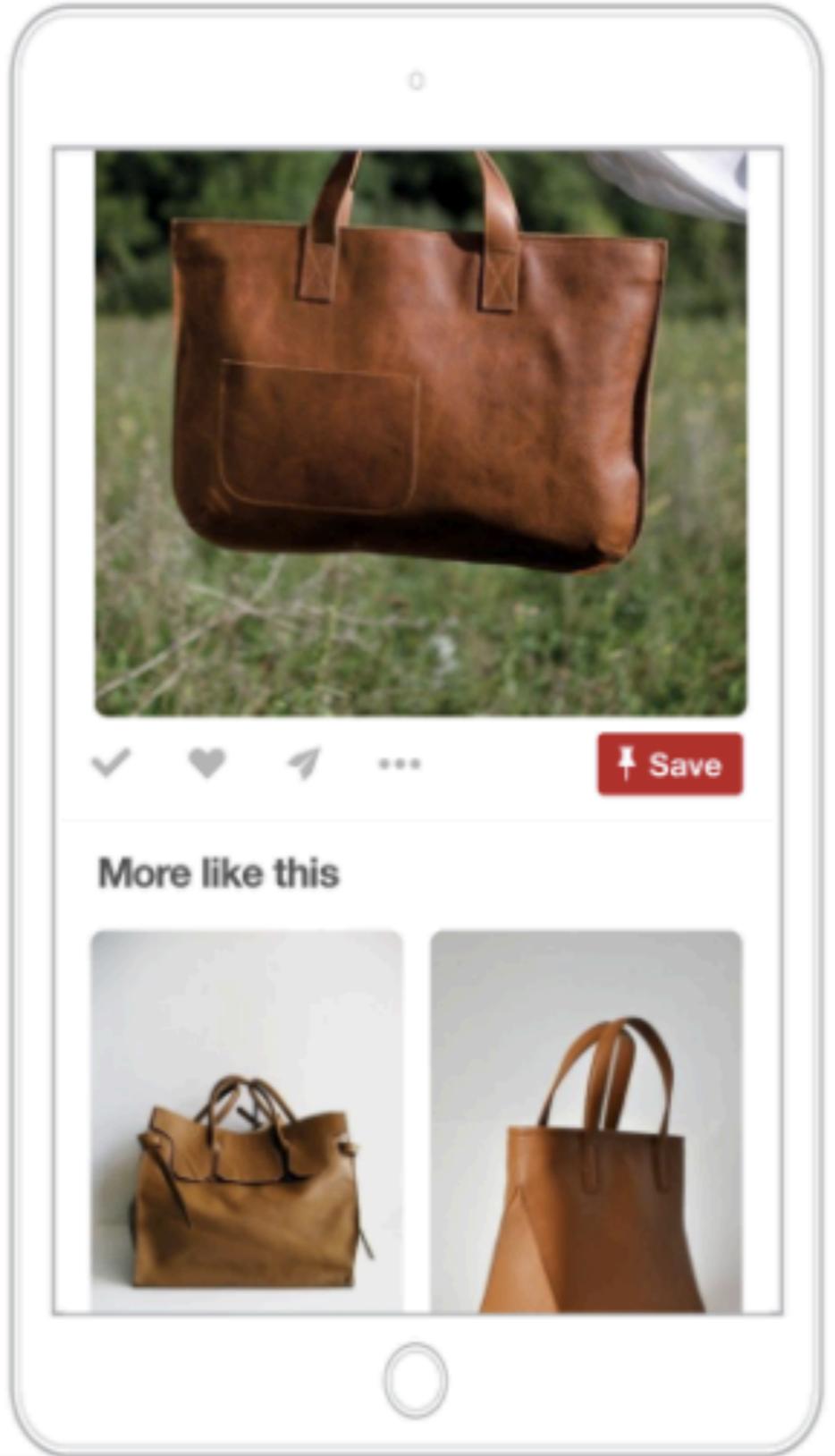


Figure 1. Related Pins recommendations for similar bags.





Zainab Mohiuddin, Blogger and Creator at Tinderella Worldwide (2017-present)

Answered Sep 22, 2017 · Author has 89 answers and 1m answer views

Hey Tinderos and Tinderellas! I'm a 26 year old woman who's been on tinder for 4 years and it's been **pretty awesome** for me. So awesome in fact, that I **research and blog about it**.

How the tinder algorithm works

Tinder isn't like Instagram, where you just need to upload some bullshit images of your breakfast and buy some bots. Nope, the rules for your visibility are more **complex**.

- Tinder calculates how attractive you are by using an **ELO score**. What the hell is an ELO score? It's this rating system thing. Here, read **THIS**.
- So it doesn't tell you how likely you are to be right-swiped, but **ranks you in terms of likability**. This is probably linked to it's **smart pictures** feature.
- The ELO score is made up of the % of people that like you as a whole, the % of matches you get and what the ELO scores of those people are.
- Your ELO score **determines who will be shown** your profile; so if you're a 7, you won't be shown to 4's but neither will you be shown to a 9. Harsh but true.
- You **can change your ELO score** by modifying your profile.
- You will **NEVER EVER** know your ELO score (because that's just too harsh).

What are the recommender systems for?

What are the recommender systems for?

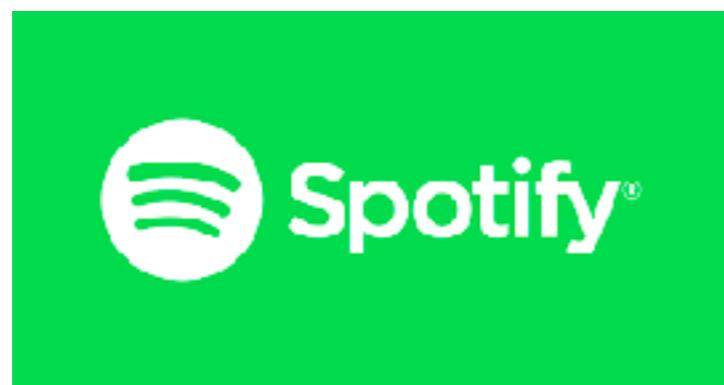
System designed to recommend **to me**

something **I** may like!

What are the recommender systems for?

Recommender systems are tools designed for interacting with **large and complex** information spaces and **providing information or items** that are likely to be of **interest to the user, in an automated fashion.**

**What they recommend?
Which data do they use?**



NETFLIX

[← Back to Help Home](#)

How Netflix's Recommendations System Works

Our business is a subscription service model that offers personalized recommendations, to help you find shows and movies of interest to you. To do this we have created a proprietary, complex recommendations system. This article provides a high level description of our recommendations system in plain language.

The basics

Whenever you access the Netflix service, our recommendations system strives to help you find a show or movie to enjoy with minimal effort. We estimate the likelihood that you will watch a particular title in our catalog based on a number of factors including:

- your interactions with our service (such as your viewing history and how you rated other titles),
- other members with similar tastes and preferences on our service, and
- information about the titles, such as their genre, categories, actors, release year, etc.

In addition to knowing what you have watched on Netflix, to best personalize the recommendations we also look at things like:

- the time of day you watch,
- the devices you are watching Netflix on, and
- how long you watch.

All of these pieces of data are used as inputs that we process in our algorithms. (An algorithm is a process or set of rules followed in a problem solving operation.) The recommendations system does not include demographic information (such as age or gender) as part of the decision making process.

If you're not seeing something you want to watch, you can always search the entire catalog available in your country. We try to make searching as easy and quick as possible. When you enter a search query, the top results we return are based on the actions of other members who have entered the same or similar queries.

Rows, rankings and title representation

In addition to choosing which titles to include in the rows on your Netflix homepage, our system also ranks each title within the row, and then ranks the rows themselves, using algorithms and complex systems to provide a personalized experience. To put this another way, when you look at your Netflix homepage, our systems have ranked titles in a way that is designed to present the best possible ordering of titles that you may enjoy.

In each row there are three layers of personalization:

- the choice of row (e.g. Continue Watching, Trending Now, Award-Winning Comedies, etc.)
- which titles appear in the row, and
- the ranking of those titles.

The most strongly recommended rows go to the top. The most strongly recommended titles start on the left of each row and go right -- unless you have selected Arabic or Hebrew as your language in our systems, in which case these will go right to left.

How we improve our recommendations system

We take feedback from every visit to the Netflix service and continually re-train our algorithms with those signals to improve the accuracy of their prediction of what you're most likely to watch. Our data, algorithms, and computation systems continue to feed into each other to produce fresh recommendations to provide you with a product that brings you joy.



By [Ashok Chandrashekhar](#), [Fernando Amat](#), [Justin Basilico](#) and [Tony Jebara](#)

For many years, the main goal of the Netflix personalized recommendation system has been to get the right titles in front each of our members at the right time. With a catalog spanning thousands of titles and a diverse member base spanning over a hundred million accounts, recommending the titles that are just right for each member is crucial. But the job of recommendation does not end there. Why should you care about any particular title we recommend? What can we say about a new and unfamiliar title that will pique your interest? How do we convince you that a title is worth watching? Answering these questions is critical in helping our



A Netflix homepage without artwork. This is how historically our recommendation algorithms viewed a page.



Artwork for Stranger Things that each receive over 5% of impressions from our personalization algorithm. Different images cover a breadth of themes in the show to go beyond what any single image portrays.

one user have seen:



because of this,
we want to recommend

PULP FICTION

one user have seen:



because of this,
we want to recommend



We want to recommend **PULP FICTION**

one user have seen:



another user have seen:

