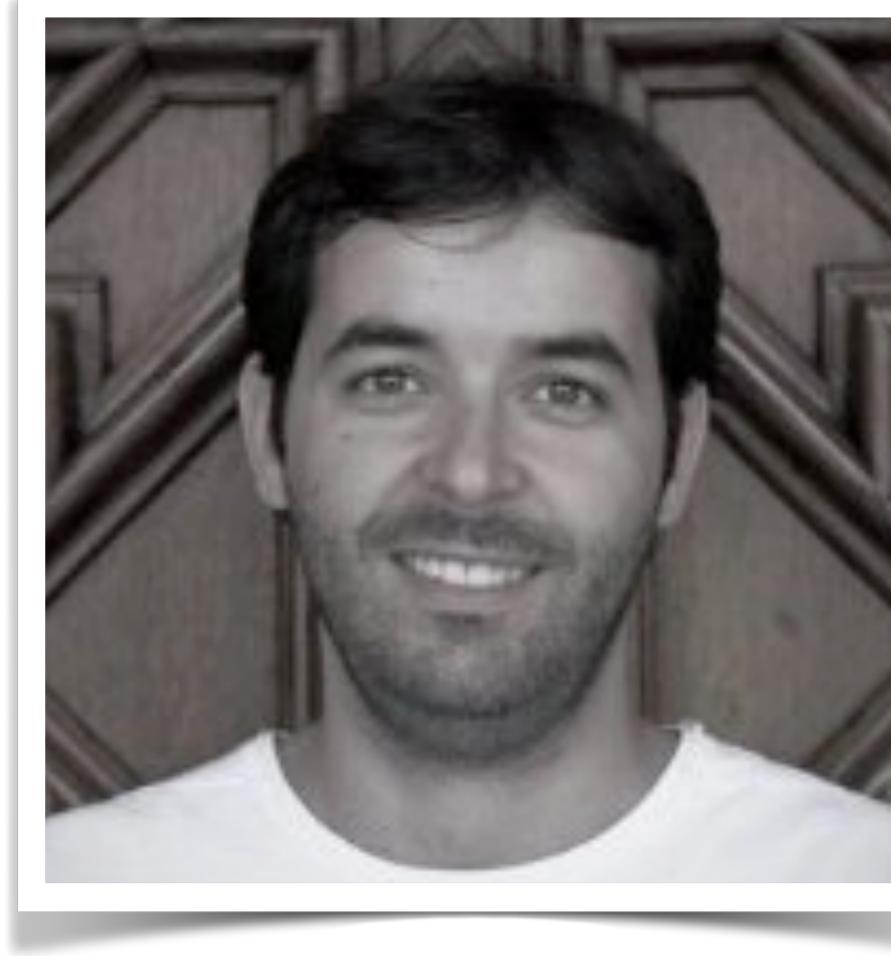


Master on Foundations of Data Science



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?

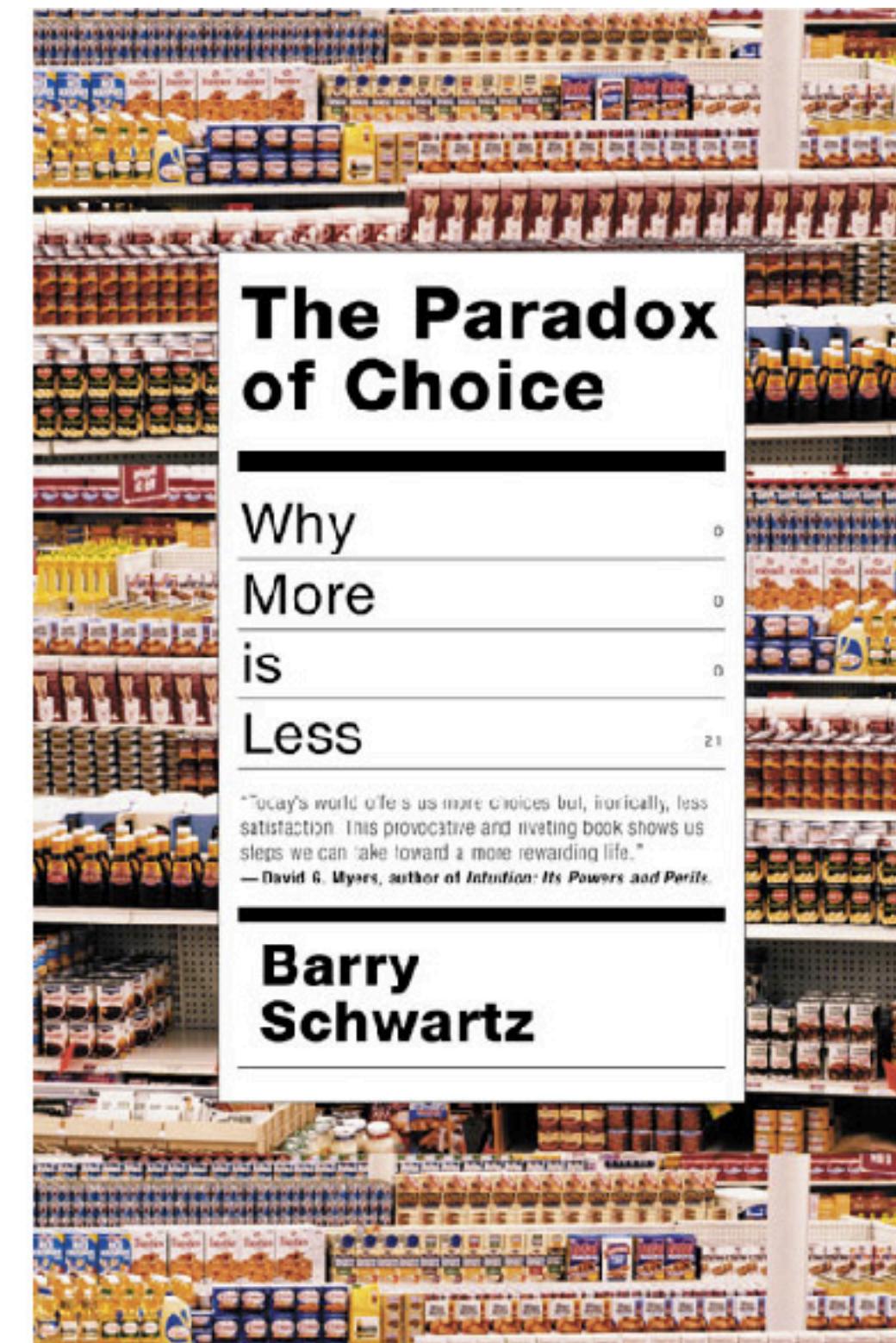


**MOST OF THE TIME WE
DO NOT WANT TO
CHOOSE**

“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

Because of that....

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

–Chris Anderson, from the Long Tail Book
- 2008



*"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

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

–UCSD Study, 2014

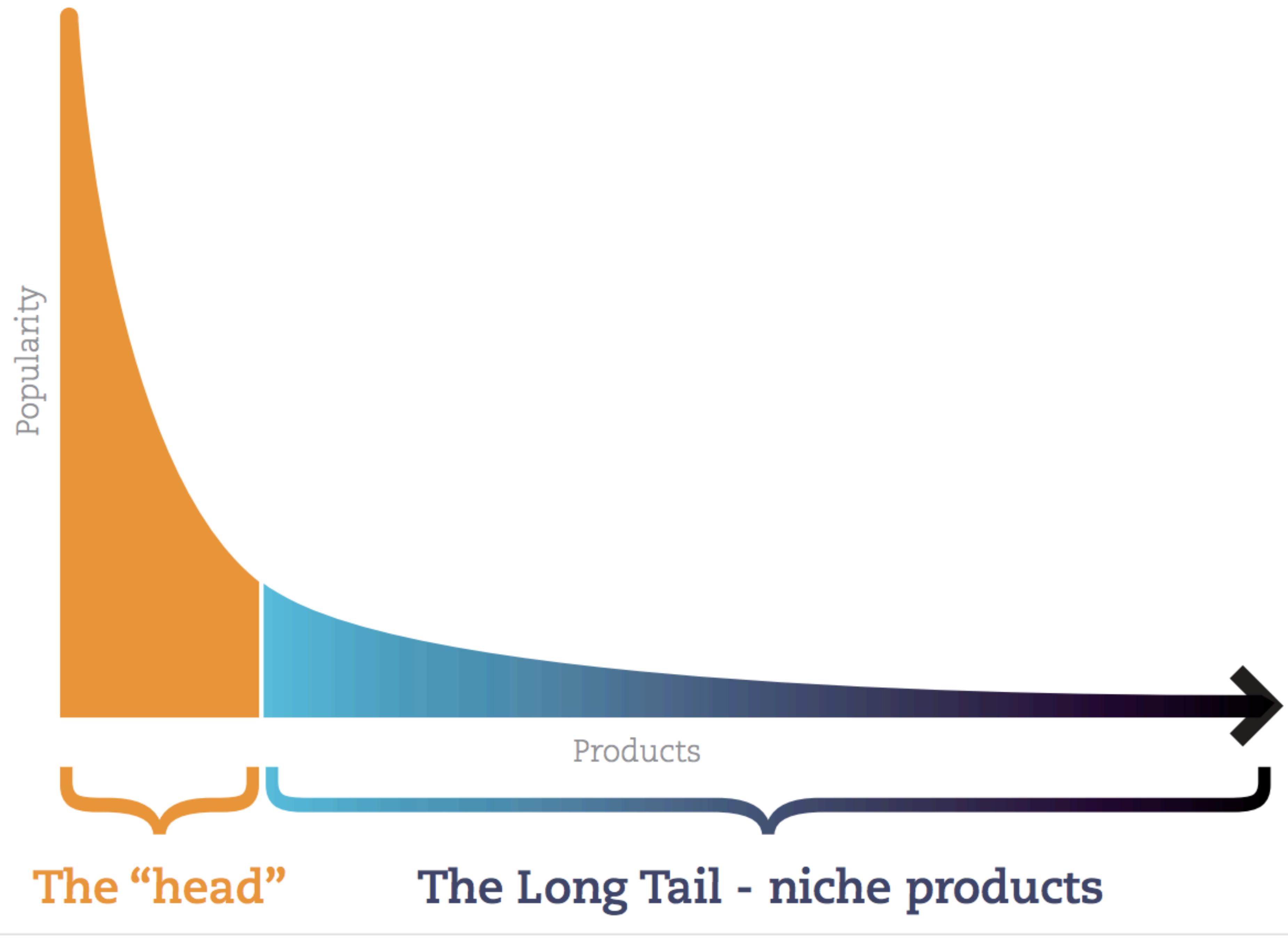
Today:

on facebook: More than 350 million photos get uploaded per day

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

Why business uses recommender systems?

- “**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



**Today,
RecSys models are really popular**

Lots of companies uses them



amazon.com.



movielens
helping you find the right movies

last.fm™
the social music revolution

Google™
News

You Tube

XBOX
LIVE

Walmart

facebook

LinkedIn

Telefónica
Investigación y Desarrollo

BBVA
DATA & ANALYTICS

Santander



ACM RecSys 2022

The 16th ACM Recommender Systems Conference will take place in Seattle from Sept. 18 - 23, 2022.

LATEST NEWS

Feb 2, 2022: The RecSys 2022 [Call for Papers](#) is out. Start preparing your great contributions, and mark your calendars!

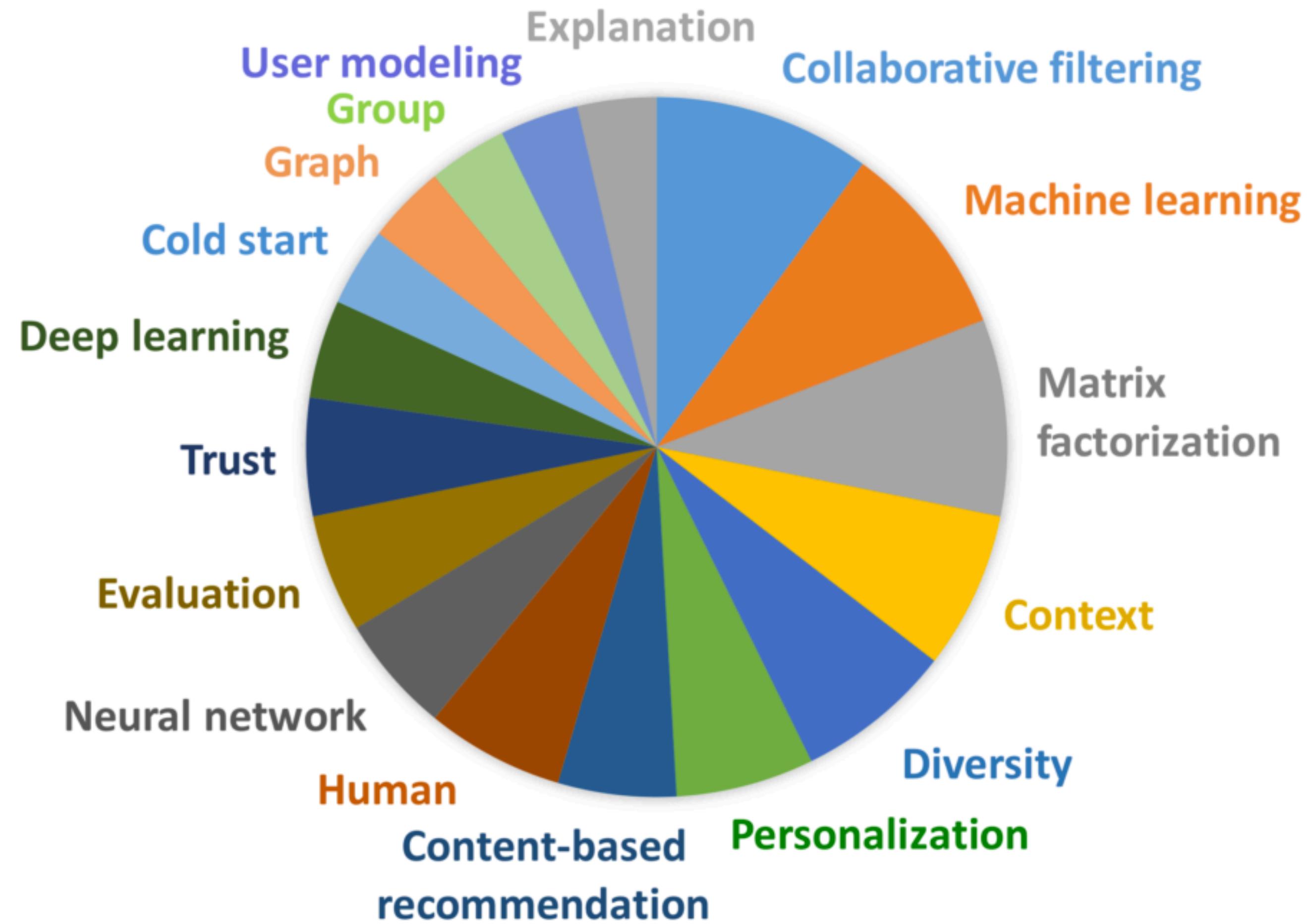
Sept 30, 2021: RecSys 2021 had a crammed [program](#) and the papers can be found in the [proceedings](#). We are looking forward to [RecSys 2022 in Seattle!](#)

SHORTCUTS TO CONFERENCES

- [RecSys 2022](#) (Seattle)
- [RecSys 2021](#) (Amsterdam)
- [RecSys 2020](#) (Online)
- [RecSys 2019](#) (Copenhagen)
- [RecSys 2018](#) (Vancouver)
- [RecSys 2017](#) (Como)
- [RecSys 2016](#) (Boston)
- [RecSys 2015](#) (Vienna)
- [RecSys 2014](#) (Silicon Valley)

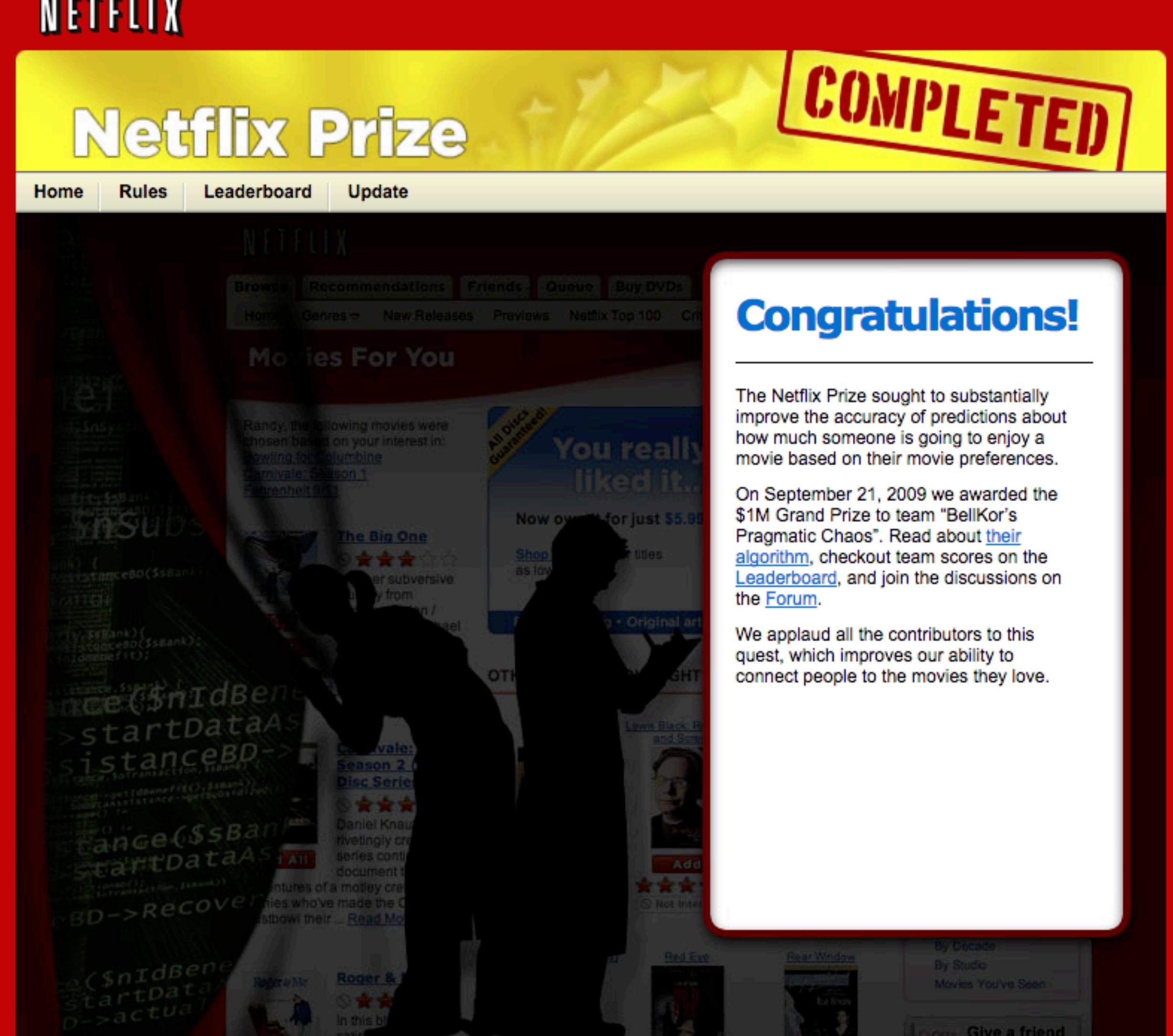


The ACM Conference Series on
Recommender Systems



Accepted Papers - Top Author Keywords

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"





About

The RecSys Challenge 2019 will be organized by [trivago](#), [TU Wien](#), [Politecnico di Milano](#), and [Karlsruhe Institute of Technology](#). trivago is a global hotel search platform focused on reshaping the way travelers search for and compare hotels, while enabling advertisers of hotels to grow their businesses by providing access to a broad audience of travelers via our websites and apps. trivago has established 55 localized platforms in over 190 countries and provides access to over two million hotels, including alternative accommodations, with prices and availability from over 400+ booking sites and hotel chains.

This year's challenge focuses on travel metasearch. The goal of this challenge is to develop a session-based and context-aware recommender system using various input data to provide a list of accommodations that will match the needs of the user.

Benchmark Challenge: Running

Spotify Million Playlist Dataset Challenge

A dataset and open-ended challenge for music recommendation research

By  Spotify 16.4k 839 23 70 23 Follow

[Overview](#) [Leaderboard](#) [Notebooks](#) [Discussion](#) [Insights](#) [Resources](#) [Submissions](#) [Rules](#)

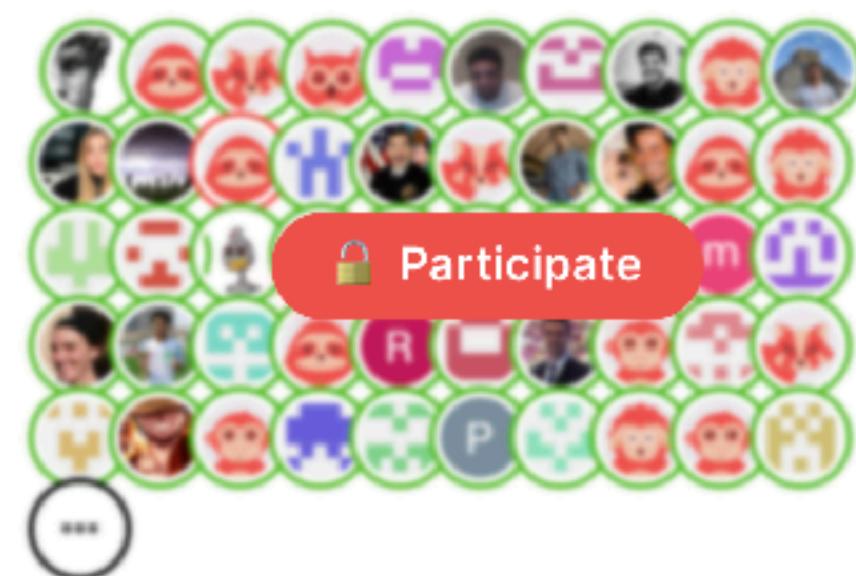
[Participate](#)

-  Summary
-  Background
-  Dataset
-  Task
-  Evaluation
-  Challenge Dataset
-  Submission Format



Spotify® Million Playlist Dataset

PARTICIPANTS



Updates

Competition submission now closed

Thank you to all participants in the challenge, we will compute the final leaderboard and post an update shortly. The submission portal is now closed.

Important Dates

- June 1st - Test set released, submissions open
- June 15th - End of challenge
- June 22nd - Announcement of winners
- July 8th - Paper submissions deadline
- August 5th - Paper acceptance notification
- August 19th - Camera-ready due

Challenge paper now available

We have released a pre-print of the challenge paper, which goes into more details about the challenge, dataset and metrics used. If you use this dataset for any publication, please cite [this paper](#).

Dataset change lists available

We are aware that the dataset size is huge, and as it is frequently changing, instead of downloading the whole dataset and validation set each time they are updated (~80GB) you can now just download the list of *engaged with Tweet IDs* and *user IDs* of data that has been deleted. Please remove any data that contain these from your local copy. You can find these lists [here](#).

Introduction

Twitter is what's happening in the world and what people are talking about right now. On Twitter, life comes to life as conversations unfold, showing you all sides of the story. From breaking news and entertainment to sports, politics and everyday interests, when things happen in the world, they happen first on Twitter.

On the platform, users post and engage with (in the form of Likes, Replies, Retweets and Retweets with comments) content known as "Tweets". This challenge aims to evaluate novel algorithms for predicting different engagement rates at a large scale, and push the state-of-the-art in recommender systems. Following the success and advancements in the domain of top-K recommendations, we aim to encourage the development of new approaches by releasing the largest real-world dataset to predict user engagements. The dataset comprises of roughly 200 million public engagements, along with user and engagement features, that span a period of 2 weeks and contain public interactions (Like, Reply, Retweet and Retweet with comment), as well as 100 million pseudo negatives which are randomly sampled from the public follow graph. While sampling the latter pool of Tweets, we take special care about preserving user privacy.

The submitted methods will be evaluated on a held-out test set generated from more recent Tweets on the platform, and the evaluation metrics will include precision-recall area under curve (PR-AUC) and cross-entropy loss. Participants will also be provided with a validation set, for which the

Booking.com Data Challenge

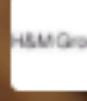
Booking.com Challenge - WebTour 2021
ACM WSDM workshop



Featured Prediction Competition

H&M Personalized Fashion Recommendations

Provide product recommendations based on previous purchases

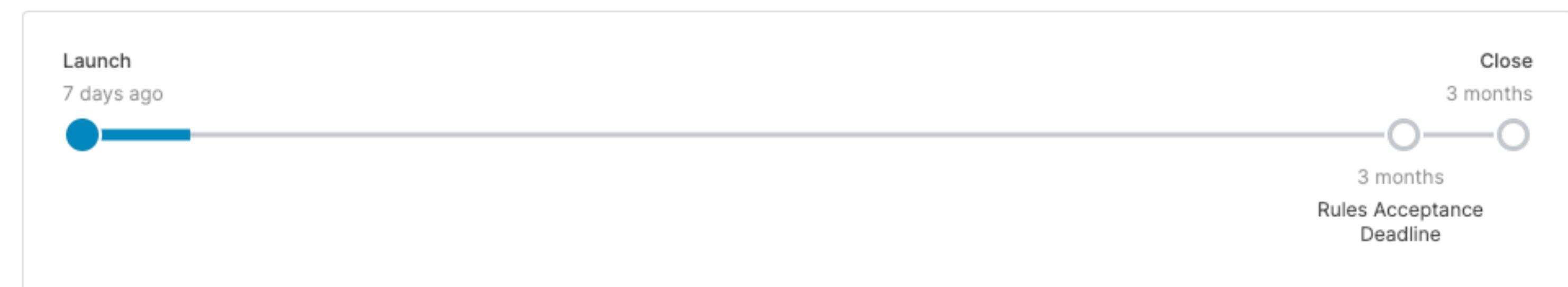
 H&M Group · 355 teams · 3 months to go (3 months to go until merger deadline)

\$50,000
Prize Money

Overview Data Code Discussion Leaderboard Rules [Join Competition](#) ...

Overview

Description	H&M Group is a family of brands and businesses with 53 online markets and approximately 4,850 stores. Our online store offers shoppers an extensive selection of products to browse through. But with too many choices, customers might not quickly find what interests them or what they are looking for, and ultimately, they might not make a purchase. To enhance the shopping experience, product recommendations are key. More importantly, helping customers make the right choices also has a positive implications for sustainability, as it reduces returns, and thereby minimizes emissions from transportation.
Evaluation	In this competition, H&M Group invites you to develop product recommendations based on data from previous transactions, as well as from customer and product meta data. The available meta data spans from simple data, such as garment type and customer age, to text data from product descriptions, to image data from garment images.
Timeline	There are no preconceptions on what information that may be useful – that is for you to find out. If you want to investigate a categorical data type algorithm, or dive into NLP and image processing deep learning, that is up to you.
Prizes	



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

Monday	Thursday	
14 Feb Introduction	17 Feb Non-Personalized	
21 Feb Collaborative Based Recsys	24 Feb Collaborative Based Recsys	MOVIELENS CF
28 Feb Factorization Models	3 March Content Based Models	
7 March Factorization with side features +Factorization Machines	10 March Deep Learning for RecSys	MOVIELENS WITH TF
14 March RecSys Challenge	17 March Context Based Models Evaluation /Learning to Rank	
21 March Graph Based Models	24 March Bias (popularity)	DEEP LEARNING
28 March Project Graphs and Bias	31 March Quiz	

Course Evaluation

Assignment #1 : 25% - Collaborative RecSystem

Assignment #2: 25% - Content RecSystems

Assignment #3: 25% - Graph / DL RecSystems

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

A bit of History

1st Generation

- Knowledge-Based
- Content-Based
- Collaborative Filtering
- Hybrid

2nd Generation

- Matrix Factorization
- Web Usage Mining Based
- Personality Based

3rd Generation

- Collaborative Filtering using DL
- Deep Content Based
- Combined modeling of Users and Items

Recommender systems
are everywhere

IMDb Charts

Top Rated Movies

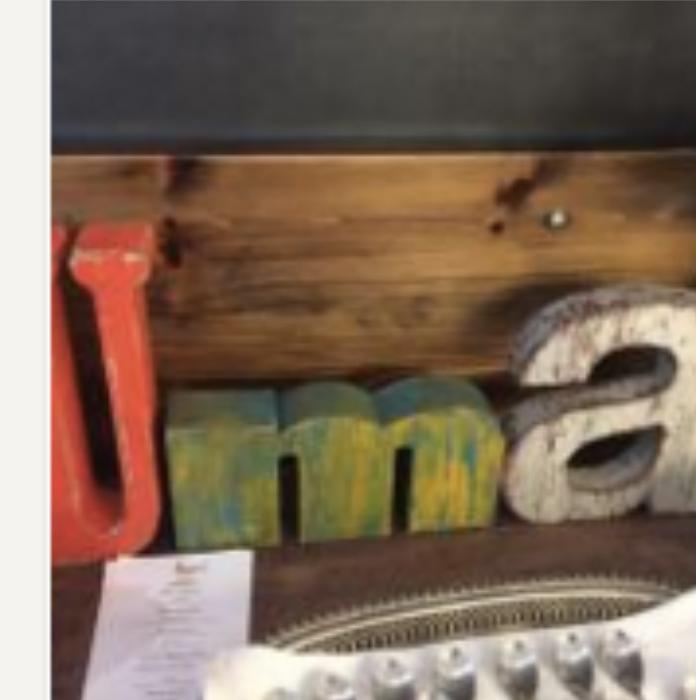
Top 250 as voted by IMDb Users

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)

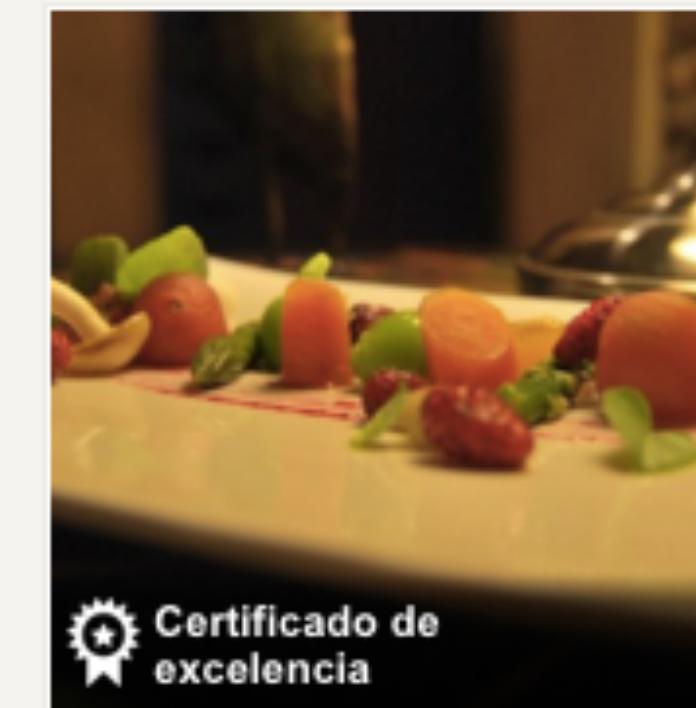
Cocina:

[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

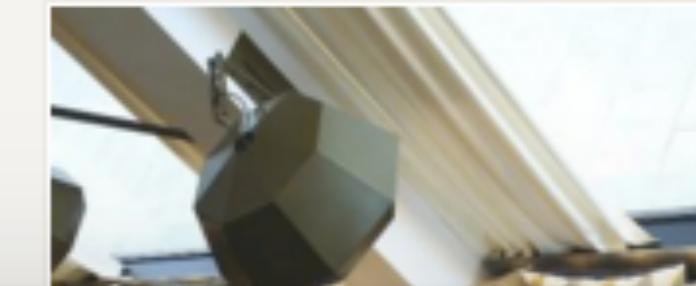
Precio: 80 € - 100 € | Mapa | Fotos de los visitantes (75)

Cocina:

[Fusión](#)

[Mediterránea](#)

[Española](#)

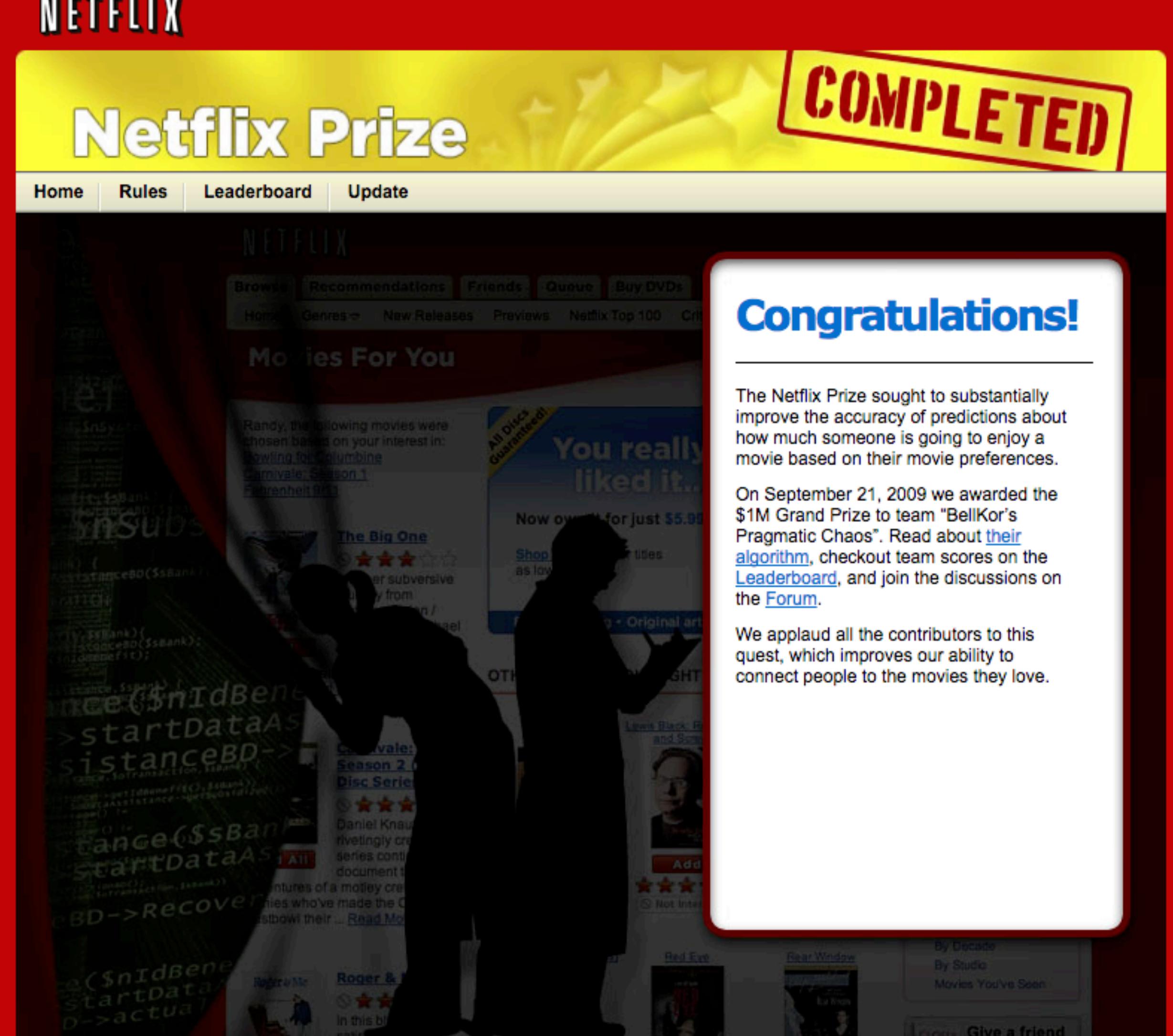


Informal

4 de 7.365 Restaurantes en Barcelona

●●●● 215 opiniones

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 ❤](#)

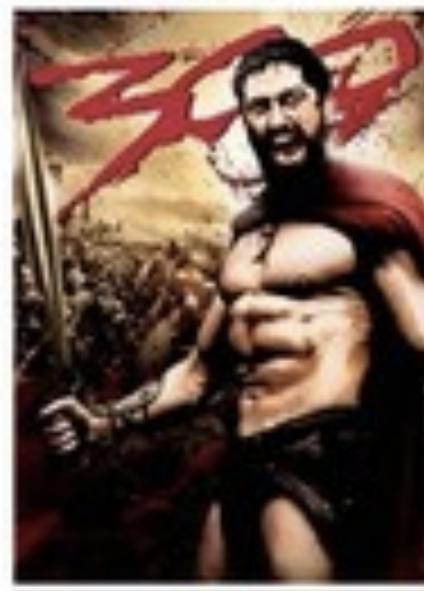
Congratulations! Movies we think You will ❤

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

Spider-Man 3

[Add](#)

300

[Add](#)

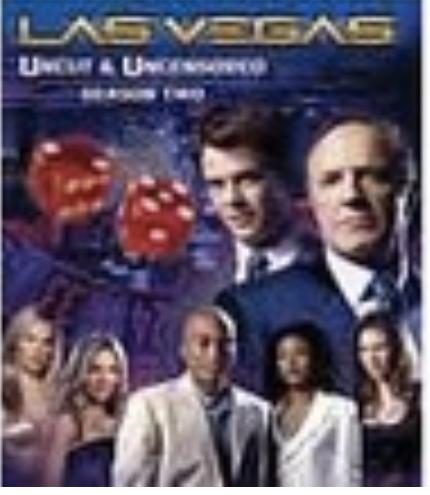
The Rundown

[Add](#)

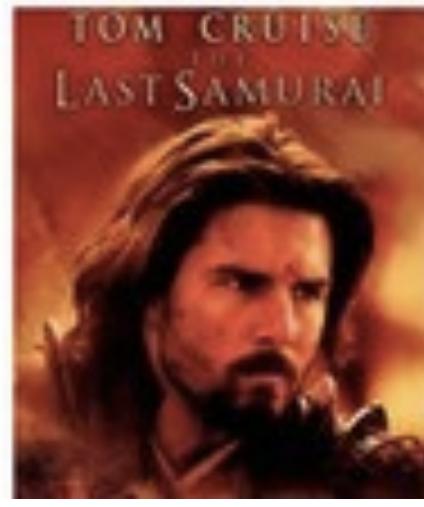
Bad Boys II

[Add](#)

Las Vegas: Season 2 (6-Disc Series)



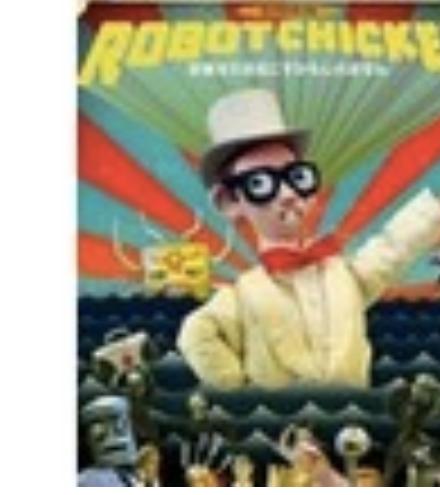
The Last Samurai



Star Wars: Episode III



Robot Chicken: Season 3 (2-Disc Series)





Browse ▾

DVD

Search



Joshua ▾

Top Picks for Joshua



Trending Now



Because you watched Narcos



New Releases





AMAZON Deliver to
Torna a configurar els plug-ins

All

All Today's Deals Customer Service Gift Cards Sell Registry

Discover our Beauty selection

You are on amazon.com. You can also shop on Amazon Spain for millions of products with fast local delivery. Click here to go to amazon.es.

Beauty picks

Shop now

Shop by Category

Computers & Accessories	Video Games
Baby	Toys & Games

Shop now

AmazonBasics

See more

Sign in for the best experience

Sign in securely

We ship over 45 million products around the world

Shop top categories

See more

Find your ideal TV

See more

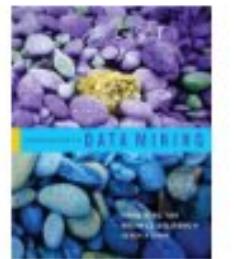
Get fit at home

Explore now

Computers & Accessories

Shop now

Frequently Bought Together



Price for both: \$158.15

Add both to Cart

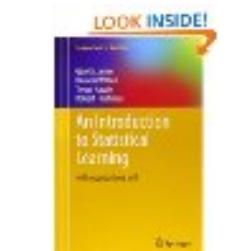
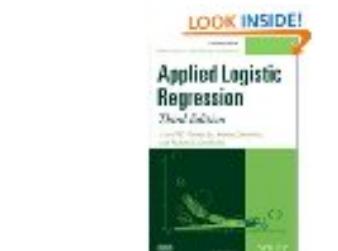
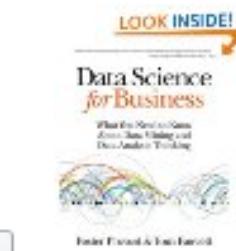
Add both to Wish List

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...](#)

› Foster Provost

 102
#1 Best Seller in Data Mining
Paperback
\$37.99 ✓Prime

[Data Mining: Practical Machine Learning Tools...](#)

› Ian H. Witten

 52
Paperback
\$40.65 ✓Prime

[Data Mining: Concepts and Techniques, Third...](#)

› Jiawei Han

 28
Hardcover
\$60.22 ✓Prime

[Regression Analysis by Example](#)

› Samprit Chatterjee

 9
Hardcover
\$92.39 ✓Prime

[SAS Statistics by Example](#)

Ron Cody

 10
Perfect Paperback
\$44.37 ✓Prime

[Applied Logistic Regression](#)

David W. Hosmer Jr.

 9
Hardcover
\$62.33 ✓Prime

[An Introduction to Statistical Learning:...](#)

› Gareth James

 56
#1 Best Seller in Mathematical & Statistical...
Hardcover
\$72.79 ✓Prime

Page 1 of 15

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...](#)

› Foster Provost

102
#1 Best Seller in Data Mining

Paperback
\$37.99

[Data Mining: Practical Machine Learning Tools...](#)

› Ian H. Witten

52
Paperback
\$40.65

[Data Mining: Concepts and Techniques, Third...](#)

› Jiawei Han

28
Hardcover
\$60.22

[Regression Analysis by Example](#)

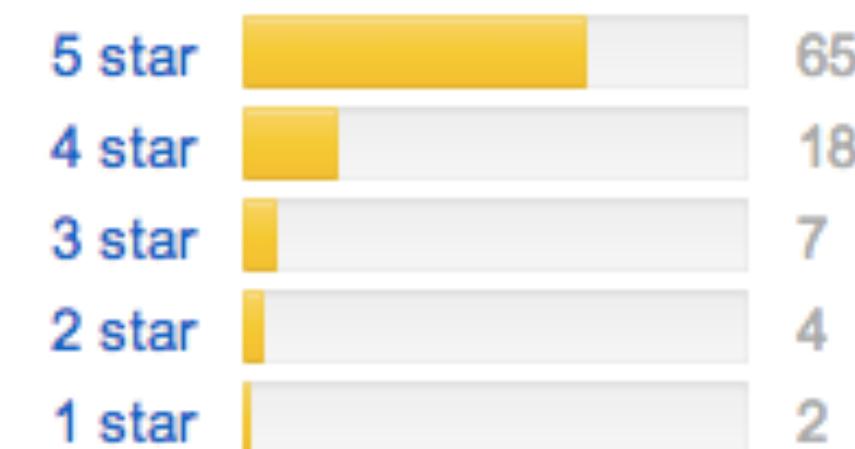
› Samprit Chatterjee

9
Hardcover
\$92.39

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

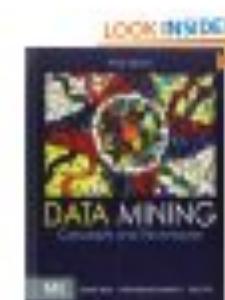
\$37.99 Prime



[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 Prime



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

› Ian H. Witten

★★★★★ 52

\$40.65 Prime

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.

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](#)

Which
information is
used?

Jobs You May Be Interested In

Powered by
LinkedIn®



Senior Data Analyst Job

Thomson Reuters - Bangalore, KA



Data Scientist/ Senior Data Scientist

HeadHonchos.com - Bangalore - IN



Hiring Computer Scientist (Java) for...

Adobe - Noida

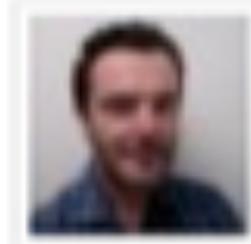


People You May Know



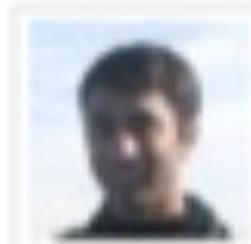
Rajeev Gupta, MBA Candidate '13
at Kelley School of Business

[Connect](#)



Alex Vauthey, Sr. Director of Engineering at LinkedIn

[Connect](#)



Bikram Bhatia, Graduate Research Assistant at University of Illinois at Urbana-Champaign

[Connect](#)

[See more »](#)

What to Watch

My Subscriptions

Music

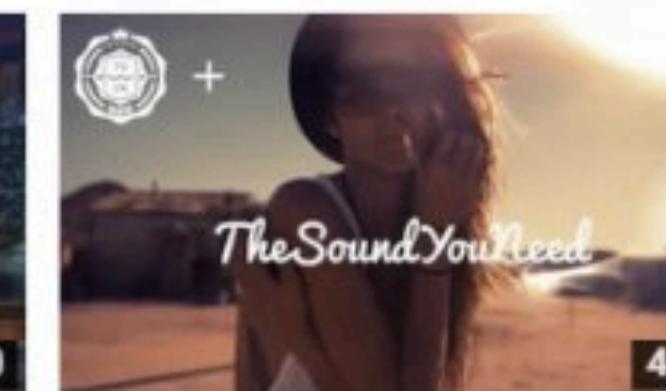
Recommended



[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 Darklordabc
88,923 views • 8 months ago



[READING MEAN COMMENTS w/ JAMES CORDEN // Grace Helbig](#)
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 Association](#)
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](#)



by Google™



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](#)

A screenshot of a Gmail draft window. At the top, there are three buttons: "Send", "Save Now", and "Discard". To the right of these is a timestamp: "Draft autosaved at 6:01 PM (1 minute ago)". Below the buttons is a "To:" field containing email addresses for "Dopey", "Grumpy", and "Doc". To the right of the "To:" field is a link "Also include: Happy, Sneezy, Sleepy" with underlined links for each character name. Below the "To:" field are "Add Cc" and "Add Bcc" links. Under the "Subject:" label is the subject line "New in Labs: Suggest more recipients". At the bottom of the draft area are two links: "Attach a file" and "Add event invitation", each preceded by its respective icon.

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

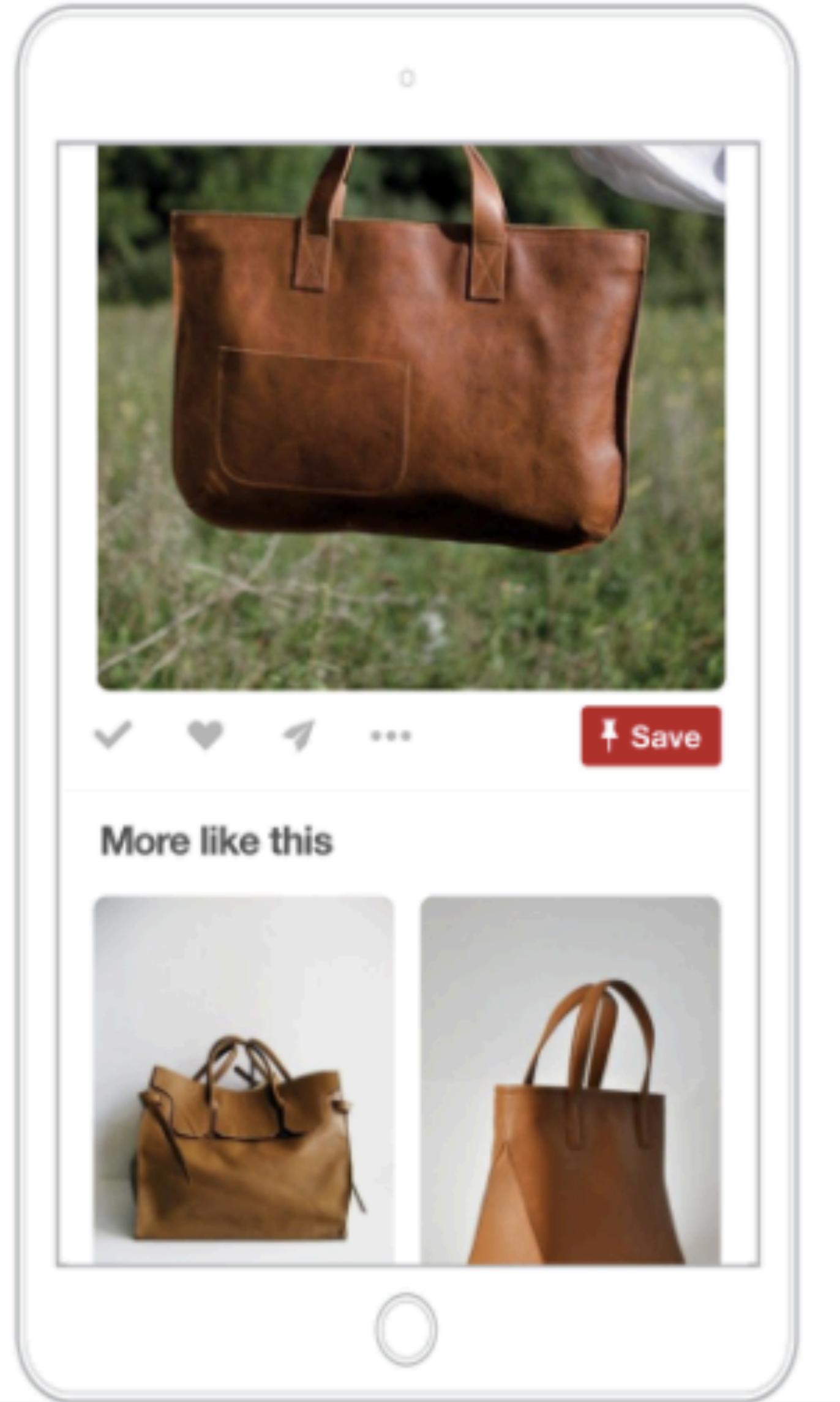


Figure 1. Related Pins recommendations for similar bags.



20 Answers



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).

Which data recommender systems use?

Two types of data

Explicit



Filling out a form



Making a purchase



Favoriting an item

Implicit



Browsing activity



Searches



Location

How recommender systems are designed?

- Two primary ways in which recommender systems may be formulated:
 1. **Prediction problem**
 2. **Ranking problem**, also referred as top-k recommendation problem

Types of Recommender Systems

Non-Personalized methods

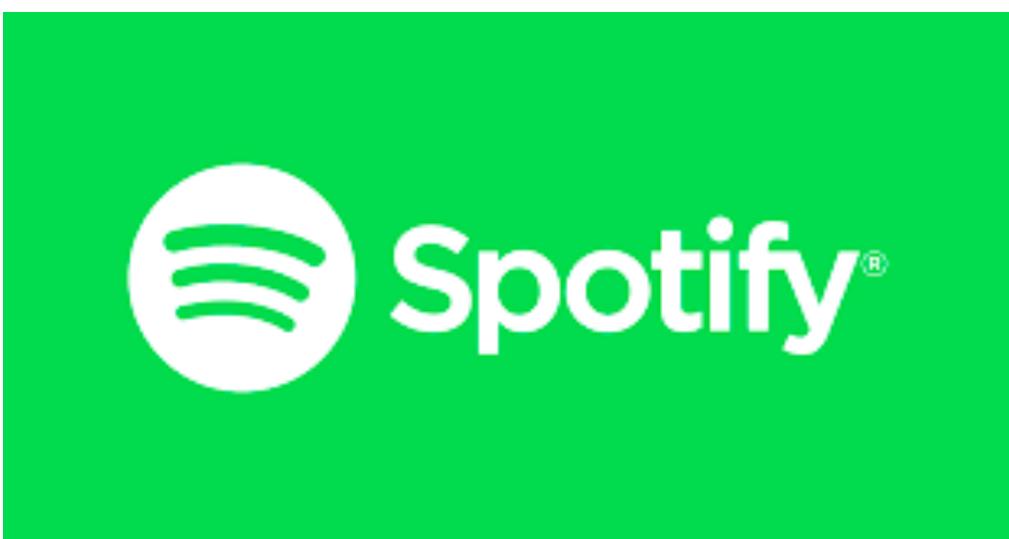
Content-based methods

Collaborative-based methods

Knowledge-based methods

HOMEWORK:

**What do 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:



What is a good Recommender System?

Goals of the recommender systems?

What do you think?

Goals of the recommender systems?

1st: Increase product sales or revenue

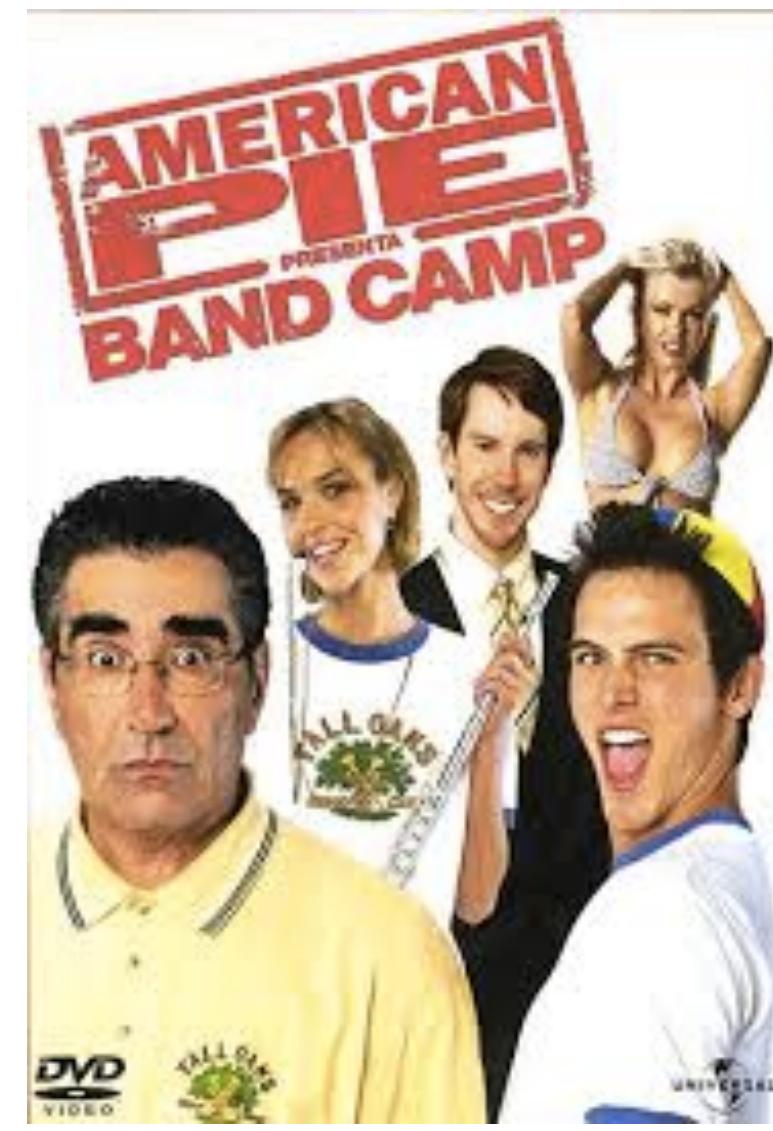
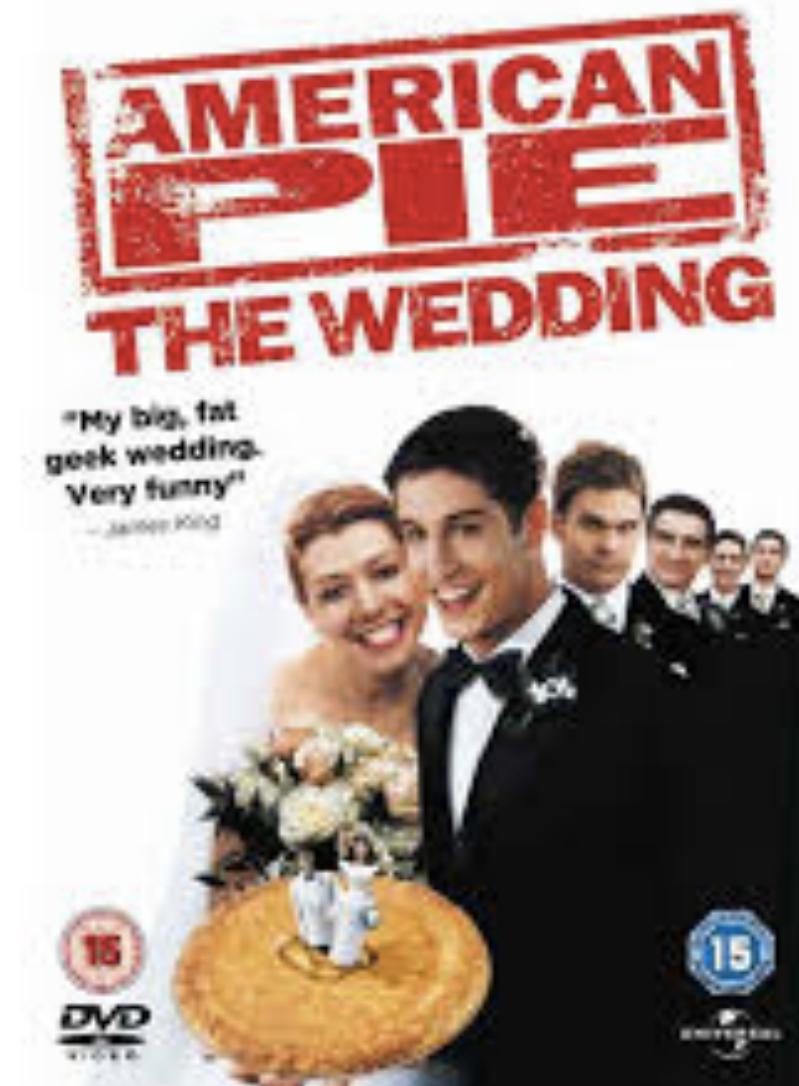
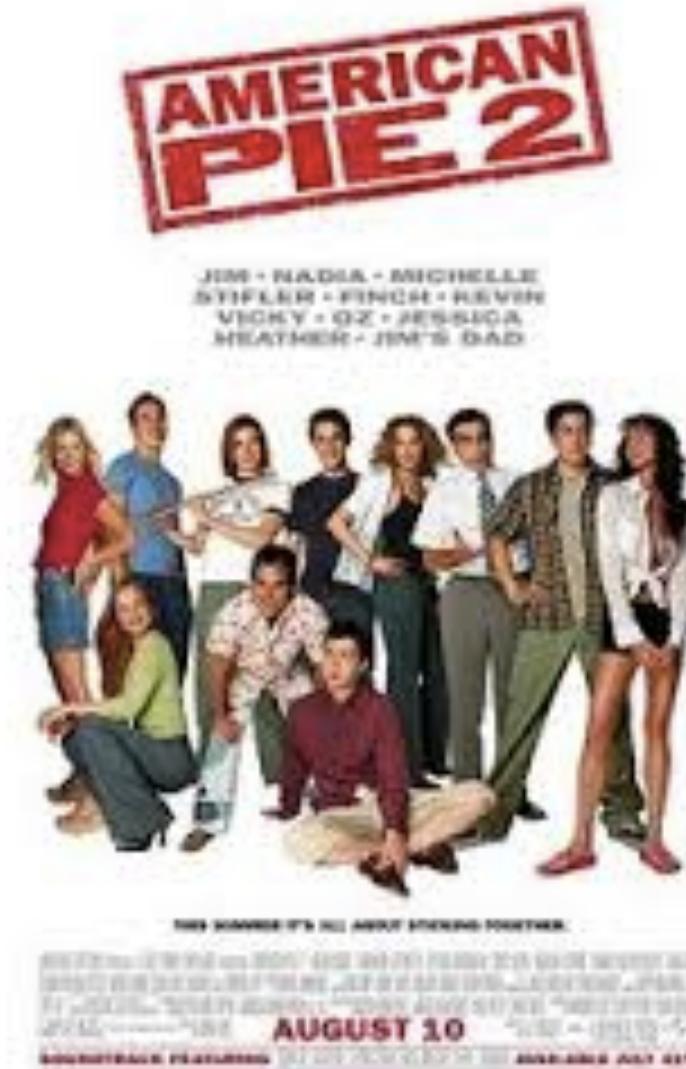
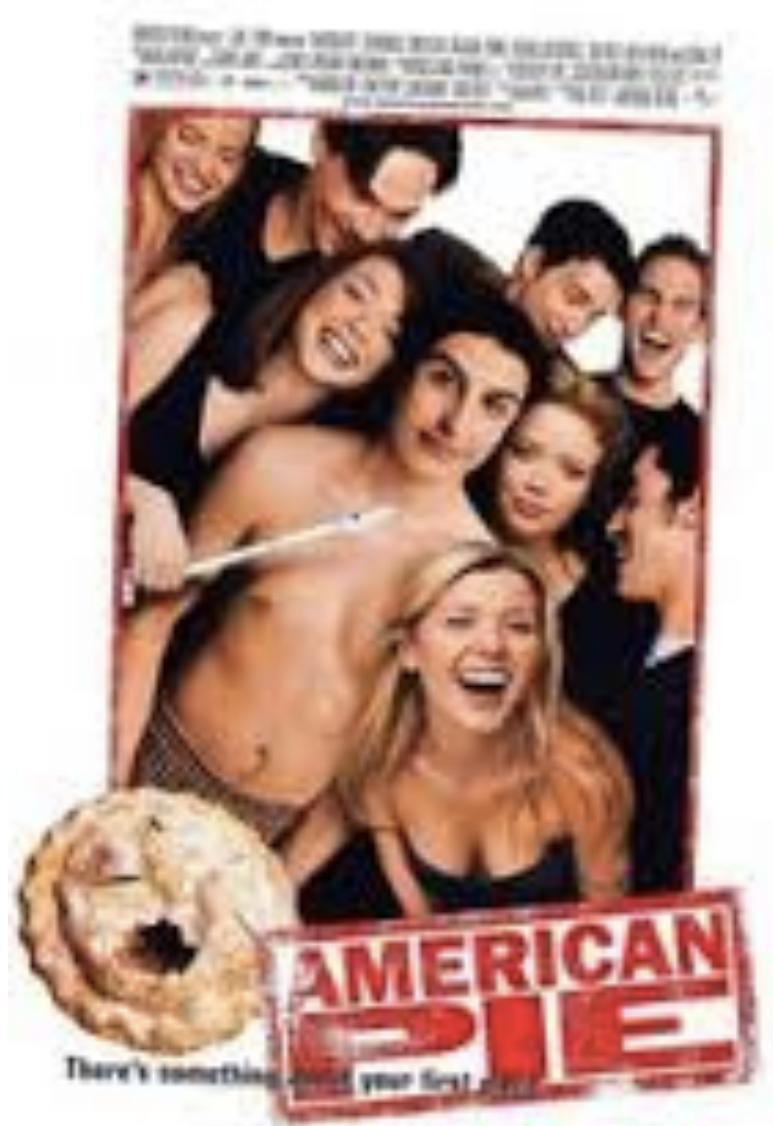
ok,
but this is the business-centric goal not easy to be codified.

Recommender systems in e-commerce

- One Recommender Systems research question
- **What should be in that list?**

Customers Who Bought This Item Also Bought





Recommender systems in e-commerce

- Another question both in research and practice
- How do we know that these are **good recommendations?**

Customers Who Bought This Item Also Bought



Recommender systems in e-commerce

- This might lead to ...
 - What is a **good** recommendation?
 - What is a **good** recommendation **strategy**?
 - What is a **good** recommendation **strategy for my business**?

Customers Who Bought This Item Also Bought



How can we evaluate a recommender?

How can we **evaluate** a recommender system?

The common operational and technical goals are:

Relevance

Novelty

Serendipity (pleasant and unexpected)

Recommendation diversity

Which are the **measures** in the practice?

- Accuracy / Root Mean Square Error
- Some metrics related to ranking (MAP, NCGD,...)
- Total number of sales
- Promotion of certain items
- Customer return rates
- Customer satisfaction and loyalty

CAUTION

RecSys
can change
user behaviors

Evaluating Recommender systems

- This is one of the **most critical** step when building a recommender system.
- A single criterion cannot capture many goals of the designer.
- ***Online*** (such as AB-Testing) and ***offline*** methods are used

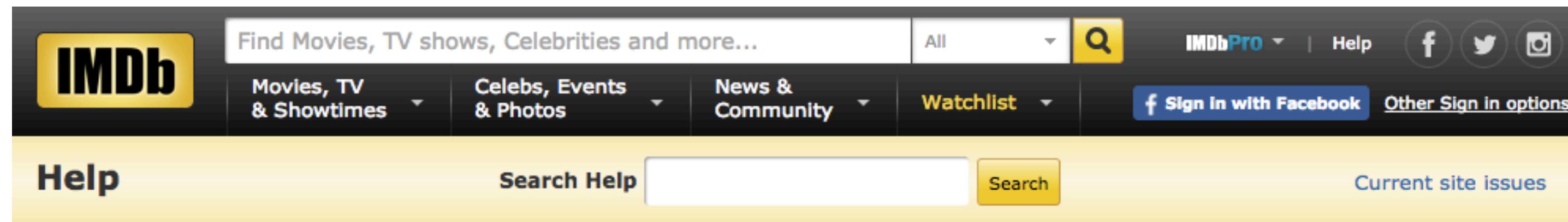
One critical point!!

Trust is critical!!



Are you recommending me
what is **good for me** or what is **good for you?**

Be transparent!

A screenshot of the IMDb website's help section. The header includes the IMDb logo, a search bar, and links for IMDbPro, Help, and social media. Below the header, a yellow navigation bar contains the word "Help", a search input field, and a "Search" button. To the right of the search bar is a link to "Current site issues".

Personalized Recommendations Frequently Asked Questions

IMDb makes personalized recommendations to help you discover movies and TV shows that you will love.

Where can I find my Watchlist and my ratings?

Find your [Watchlist](#) using the *Watchlist* link in the top right of the menu at the top of every page on the IMDb site. Find [your ratings](#) by using the *Your Ratings* item when you hover over your account name in the menu bar.

How does IMDb choose personalized recommendations?

First, we take all of the movies and TV shows that you have either [rated](#) or added to your [Watchlist](#). Then, we compare your data to ratings made by other users. We can then find movies and TV shows that people with similar tastes to you like. For each recommendation, you can see a list of the movies or TV shows upon which the recommendation was based. You have either rated these titles highly, or added them to your Watchlist.

How does IMDb know what I "showed interest in"?

When you give a movie a positive rating or add a movie to your Watchlist, we track that as a movie that you are interested in.

Why do I see items in my recommendations that I've already rated?

If we do not have any recommendations for you (either because you have logged out or because you have not rated many items), IMDb will show you a list of items that many people have seen so that you can enter ratings to get recommendations. When this happens, the widget will be titled *Rate and Get Recommendations* instead of *Recommended for You*.

Where can I find personalized recommendations?

You can see your personalized recommendations at the bottom of your [Watchlist](#) or the bottom of your [Ratings History](#). If you are logged in they are also visible on the IMDb home page in the section *Recommended for You*.

You are recommending a movie that I really do not want to see. How can I make it go away?

Click on the 'no' button.

How can I improve my personalized recommendations?

To improve your personalized recommendations, [find and rate the movies that you love](#).

I've just rated a whole bunch of titles. How long will it take for those ratings to impact my recommendations?

New ratings will have an immediate impact on your recommendations. After rating and adding titles to your Watchlist, reload the page to update your recommendations.

How do I suggest features or changes to personalized recommendations?

Please give us feedback via the [contact form](#).

Be transparent!



Note:

- If you're unsure why we're recommending a particular item, go to [Your Amazon.com](#), then click **Why Recommended?** beneath the item. We'll tell you which of your previous ratings, likes or purchases influenced our suggestion so you can revise your ratings or exclude those purchases.
- You can [tell us](#) about items you like, even if you didn't purchase it on Amazon.com. Locate the item on Amazon.com, then write a review and add a rating. We'll use that rating to help provide you better recommendations.

Be transparent!

10.



Apple MC747Z/A - Adaptador de corriente para MacBook y 13" MacBook Pro (45W), blanco

de Apple (21 de octubre de 2010)

Valoración media de los clientes: ★★★★☆ (13)

En stock

Precio: EUR 79,00

[22 de 2^a mano y nuevo\(s\) desde EUR 68,52](#)

Añadir a la cesta

Añadir a la Lista de deseos

Ya lo tengo No me interesa ★★★★☆ Valorar este producto

Recomendado porque has añadido Apple MD592Z/A - Adaptador de corriente para MacBook Air a tu cesta ([Cambiar esto](#))

11.



Caja de 20 paquetes de toallitas Zig Zag secamanos doble capa

de Clim Profesional (13 de noviembre de 2014)

Valoración media de los clientes: ★★★★☆ (3)

En stock

Precio: EUR 25,90

Ofrecido por [CLIM Profesional productos de limpieza](#)

Añadir a la cesta

Añadir a la Lista de deseos

Ya lo tengo No me interesa ★★★★☆ Valorar este producto

Recomendado porque has añadido Apple MD592Z/A - Adaptador de corriente para MacBook Air a tu cesta ([Cambiar esto](#))

12.



DELL Vostro 3559 2.3GHz i5-6200U - Ordenador portátil de 15.6" (i5-6200U, 1366 x 768 Pixeles, teclado Español QWERTY), color Negro

de Dell (10 de marzo de 2016)

Valoración media de los clientes: ★★★★☆ (1)

En stock

Precio: EUR 566,00

Ofrecido por [CanalPC Informática](#)

Añadir a la cesta

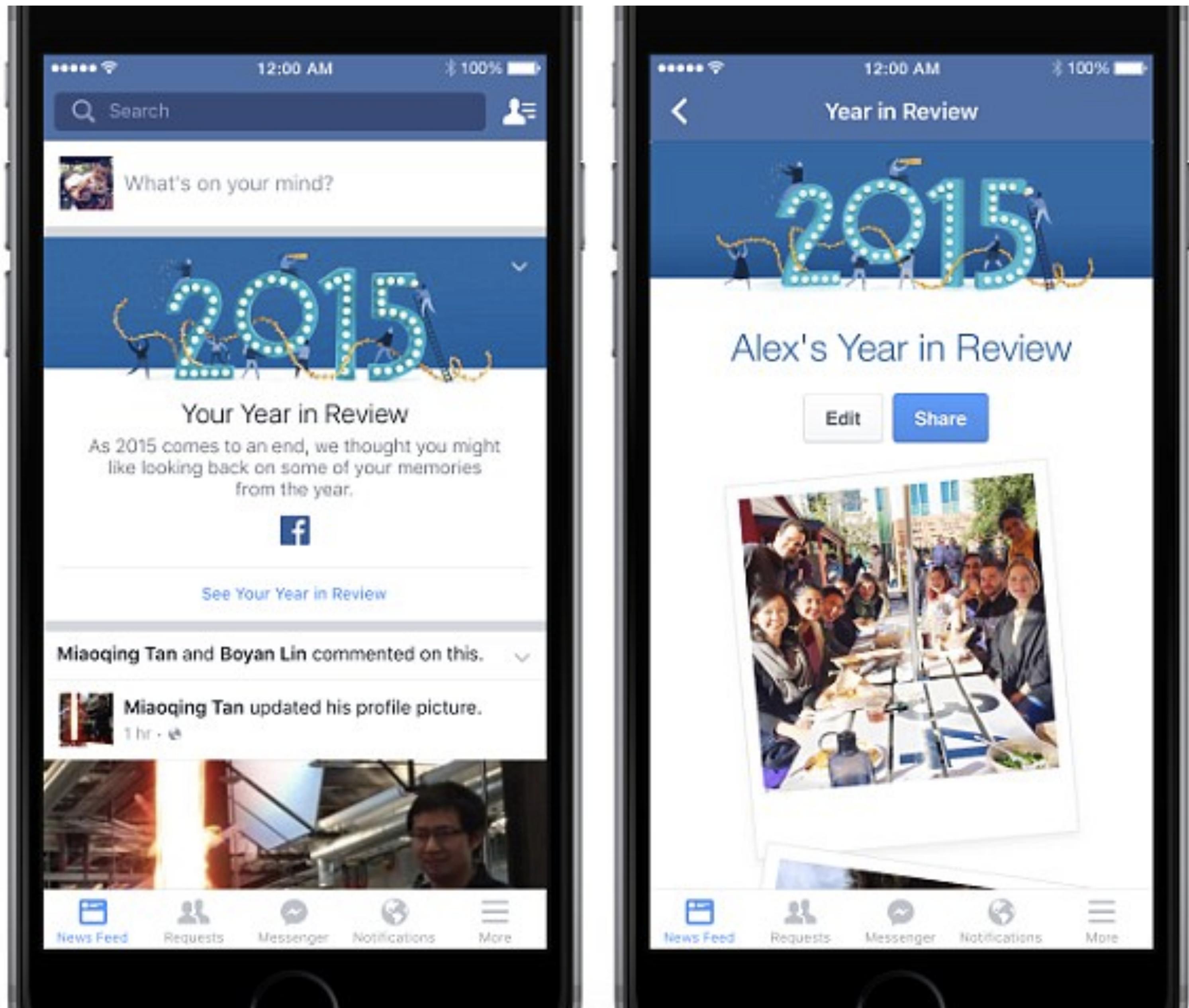
Añadir a la Lista de deseos

Ya lo tengo No me interesa ★★★★☆ Valorar este producto

Recomendado porque has añadido Apple MD592Z/A - Adaptador de corriente para MacBook Air a tu cesta ([Cambiar esto](#))

Another critical point!!

Not everything is allowed!!



<http://www.forbes.com/sites/amitchowdhry/2014/12/29/facebook-year-in-review-apology/#5c74535a3cb0>

Facebook Apologizes For Its 'Year In Review' Approach



Amit Chowdhry, CONTRIBUTOR
[FULL BIO ▾](#)

Opinions expressed by Forbes Contributors are their own.

This month, [Facebook](#) launched a customizable slideshow feature called *Year In Review*. To increase the engagement in *Year In Review*, Facebook automatically created a slideshow with suggested photos for each of their users. Facebook's algorithms decided the suggested photos based on the month each photo was uploaded and the number of "likes" they received. The default tagline for the *Year In Review* slideshow is "It's been a great year! Thanks for being a part of it." Unfortunately, many Facebook users had a tumultuous year instead of a "great" year and *Year In Review* served as a reminder of traumatic circumstances.

This year, web designer Eric Meyer lost his six-year-old daughter [Rebecca](#) to brain cancer so he did not have any interest in creating a *Year In Review* slideshow to share with his Facebook friends. However, Facebook's algorithms automatically created a slideshow for him with his daughter's face in the center of the first slide with some holiday-themed clip art. Facebook displayed the slideshow at the top of his News Feed on December 24th, asking him if he wanted to customize and share his *Year In Review*.

"I know, of course, that this is not a deliberate assault. This inadvertent algorithmic cruelty is the result of code that works in the overwhelming majority of cases, reminding people of the awesomeness of their years, showing them selfies at a party or whale spouts from sailing boats or the marina outside their vacation house," said Meyer in a [blog post](#). "But for those of us who lived through the death of loved ones, or spent extended time in the hospital, or were hit by divorce or losing a job or any one of a hundred crises, we might not want another look at this past year."