



UNIVERSITAT DE
BARCELONA



Master on Foundations of Data Science



Recommender Systems

Course Presentation



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

PhD in 2011 on Computer Vision and Machine Learning
at the University of Barcelona.

Graduate on Computer Science in 2007

Santi Seguí Email: santi.segui@ub.edu

Why recommender systems?

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

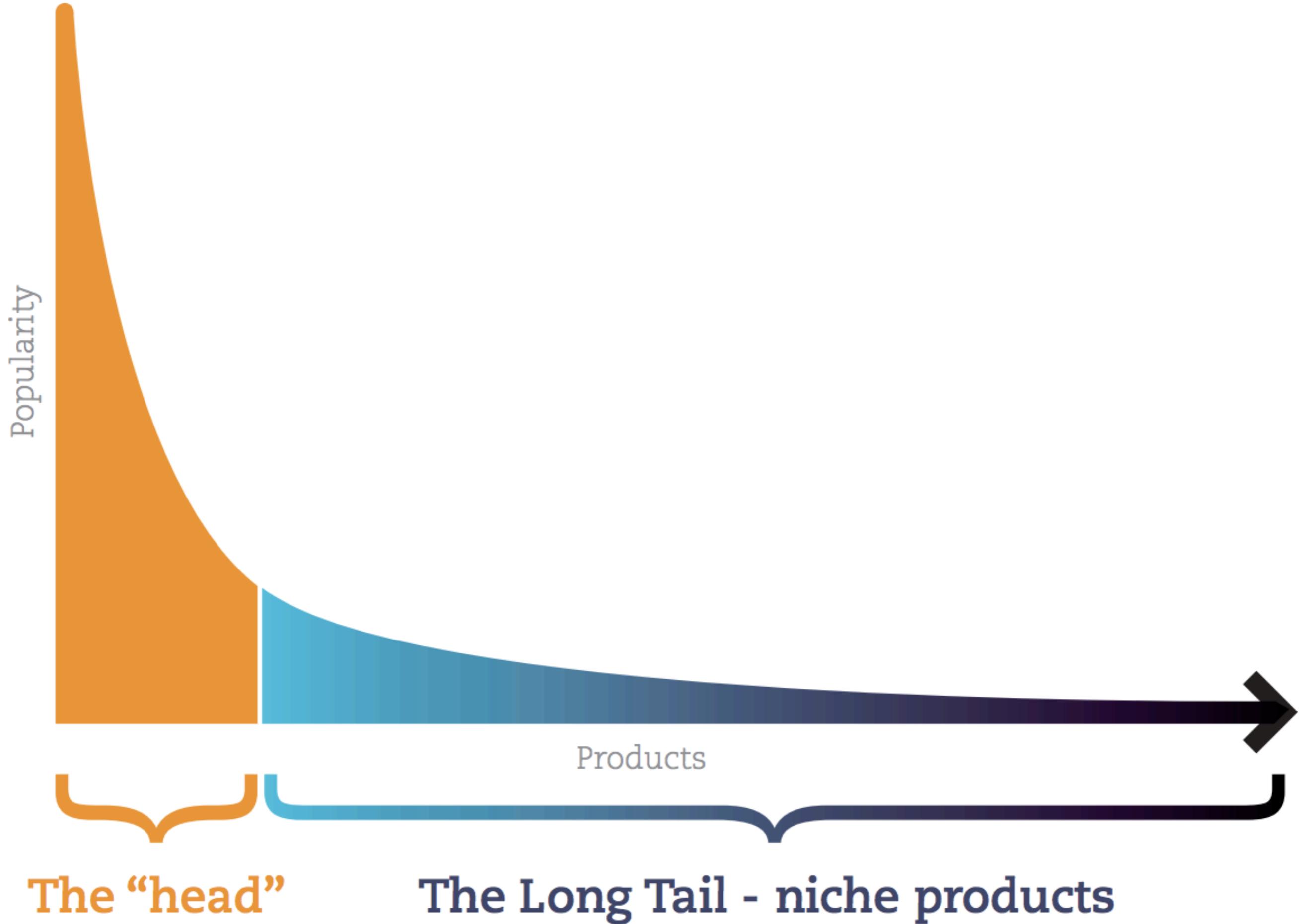
–Steve Jobs, 1997

“People read around 10 MB worth of material a day, hear 400MB a day, and see 1MB of information every second”

–The Economist, November 2006

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

–Chris Anderson, from the Long Tail Book - 2008

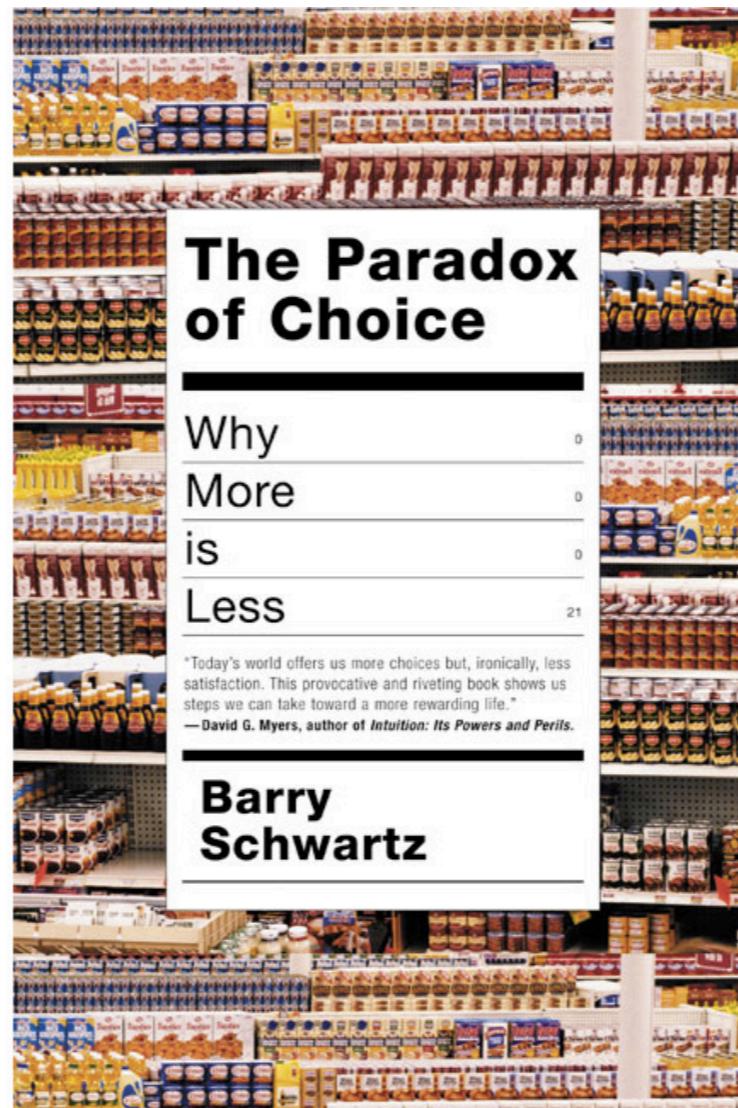


“In 2015, consumption will raise
to 74GB a day”

—UCSD Study, 2014



Information overload



Why Recommenders?

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.**



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

Labs (Thursday)	
16/02	Introduction
23/02	Non-Personalized Recommenders
02/03	Collaborative-Based Recommenders
09/03	Collaborative-Based Recommenders
16/03	Dimensionality Reduction for Recommender Systems
23/03	Content-Based Recommender Systems
30/03	Item-Based Recommender Systems
06/04	Exam Period
13/04	Holidays
20/04	Evaluation - A-B Testing
27/04	Graph Based Recommendations
04/05	Deep Learning and Recommender Systems
11/05	Context Aware Recommender System
18/05	Cold Start
25/05	Current Practices in Industry and Research

Course Evaluation

Assignment #1 : 30%

Assignment #2: 30%

Assignment #3: 20%

Final Exam : 20%

Who uses Recommender systems?

Who uses Recommender systems?



amazon.com.
The Amazon logo, which consists of the word 'amazon.com.' in black lowercase letters above a yellow curved arrow pointing from left to right.



movielens
helping you find the *right* movies

last.fm™
the social music revolution

Google™
News

YouTube

XBOX
LIVE

SCHIBSTED
MEDIA GROUP

Telefónica
Investigación y Desarrollo

BBVA
DATA & ANALYTICS

Santander

**Recommender systems
are everywhere**

A bit of History

- 1985 - First recommender systems
- 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

Top Rated Movies

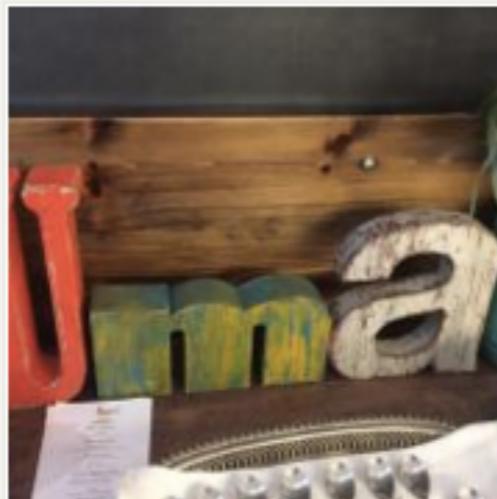
Top 250 as voted by IMDb Users

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](#)



Certificado de
excelencia

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

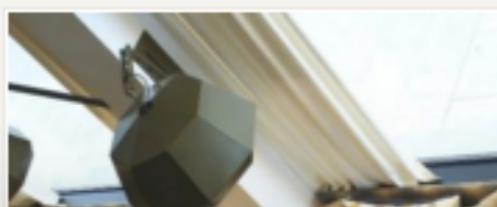
Precio: 80 € - 100 € | Mapa | Fotos de los visitantes (75)

Cocina:

[Fusión](#)

[Mediterránea](#)

[Española](#)



Informal

4 de 7.365 Restaurantes en Barcelona

●●●●● 215 opiniones



Netflix Prize

COMPLETED

[Home](#) | [Rules](#) | [Leaderboard](#) | [Update](#)

The Netflix Prize sought to substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences.

On September 21, 2009 we awarded the \$1M Grand Prize to team "BellKor's Pragmatic Chaos". Read about [their algorithm](#), checkout team scores on the [Leaderboard](#), and join the discussions on the [Forum](#).

We applaud all the contributors to this quest, which improves our ability to connect people to the movies they love.

[FAQ](#) | [Forum](#) | [Netflix Home](#)

© 1997-2009 Netflix, Inc. All rights reserved.

[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](#)

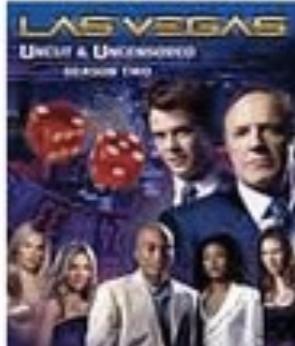
300

[Add](#)

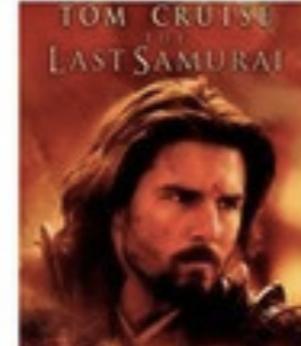
The Rundown

[Add](#)

Bad Boys II

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

The Last Samurai

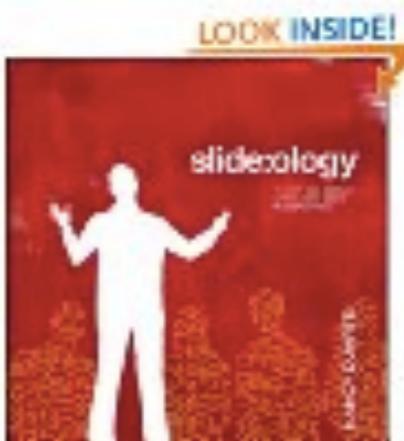


Star Wars: Episode III

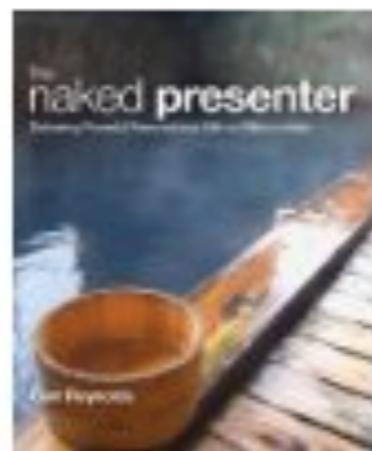
Robot Chicken: Season 3
(2-Disc Series)



Customers Who Bought This Item Also Bought



[slide:ology: The Art and Science of Creating Great Presentations](#) by Nancy Duarte
 (98)
\$23.09

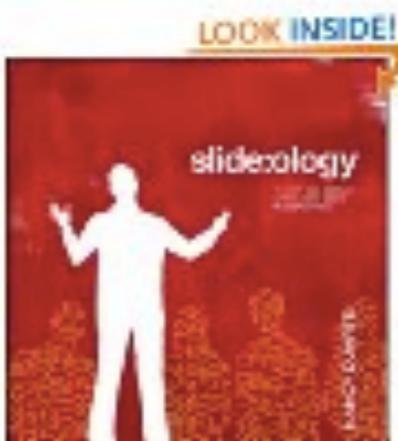


[The Naked Presenter: Delivering Powerful Presentations with Confidence and Conviction](#) by Garr Reynolds
\$16.49



[The Presentation Zen Way: Video Lessons on Simplifying Your Slides](#) by Garr Reynolds
\$26.39

Customers Who Bought This Item Also Bought



slide:ology: The Art and
Science of Creating
Great Slides... by Nancy Duarte

★★★★★ (98)

\$23.09

\$23.09

★★★★★ (86)

Learn more about this item



The Naked Prese...
Delivering Power
Present... by Gar
Reynolds

\$16.49

\$16.49

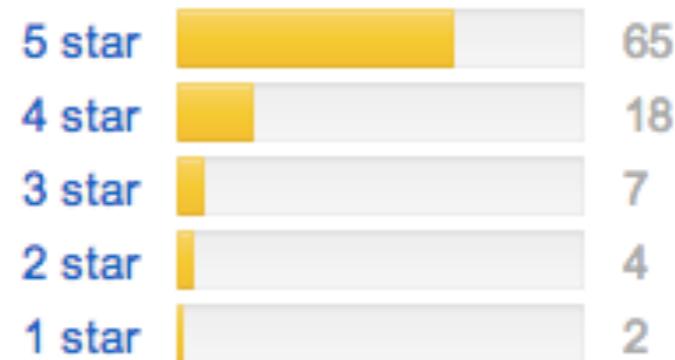
See all 96 reviews



Customer Reviews

★★★★★ (96)

4.5 out of 5 stars



[See all 96 customer reviews](#) ▾

1 Zen
ions on
eynolds

see all 96 reviews

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 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





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

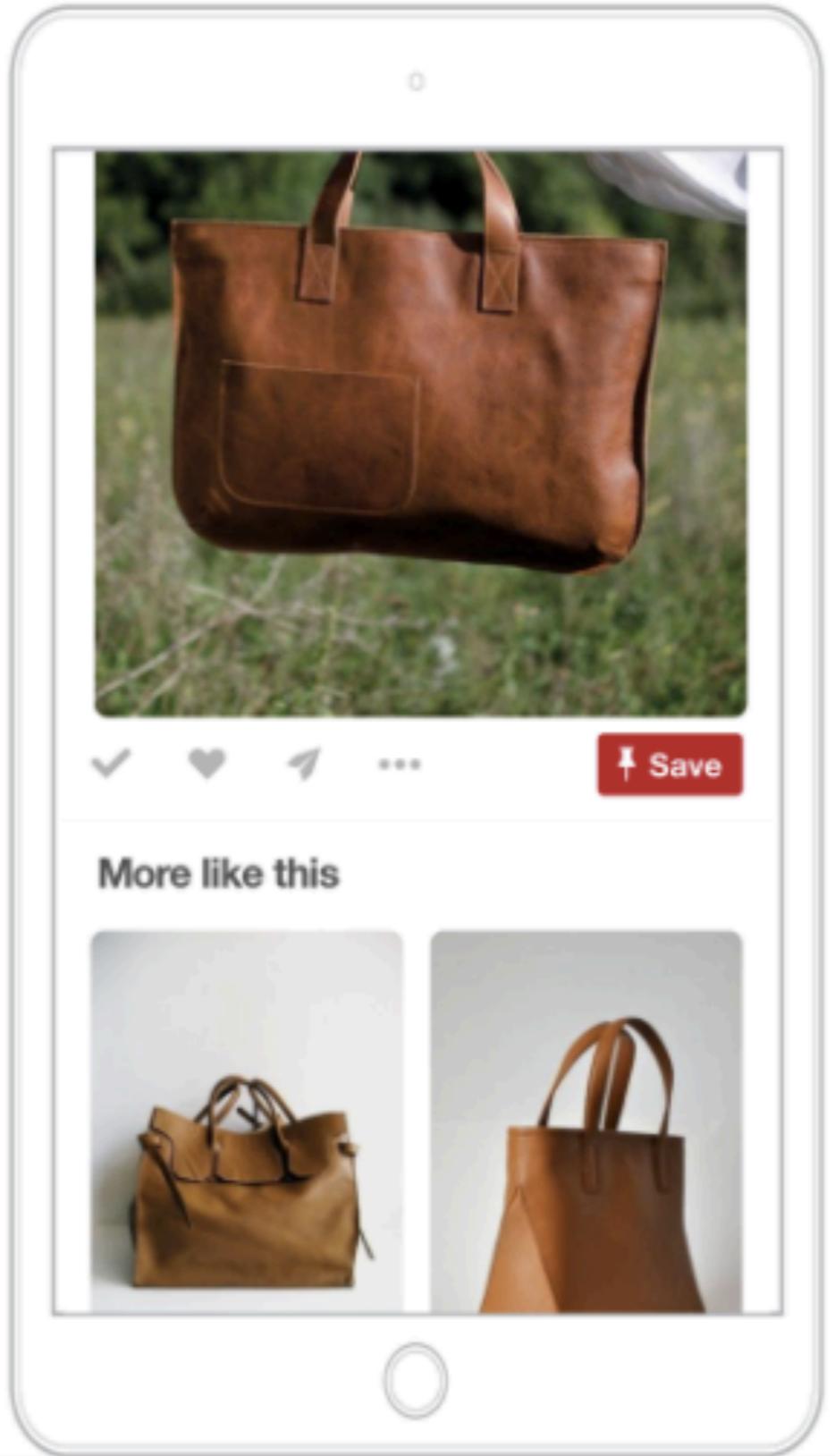


Figure 1. Related Pins recommendations for similar bags.



Recommender Systems

- 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

Non-Personalized methods

Conceptual goal:

Give me recommendation based on the users preferences
without using any information from the user itself.

Input:

Users ratings/Purchases/Behaviour

Content-based methods

Conceptual goal:

Give me recommendation based on the content (attributes) I have favored in my past ratings and actions

Input:

User ratings + item attributes

Collaborative-based methods

Conceptual goal:

Give me recommendation based on a collaborative approach that leverages the ratings and actions of my peers and myself

Input:

User ratings + community ratings

Knowledge-based methods

Conceptual goal:

Give me recommendation based on my explicit specification of the kind of content (attributes)

I want

Input:

User specifications + item attributes + domain knowledge

JAKE



Jake got a **beachball** & **sunglasses**

FUNMI



Funmi got a **polka-dot bikini**

NICK



Nick got a **beachball**...
What else might Nick like?

JAKE



Jake got a
beachball, sunglasses & flip flops

FUNMI



Funmi got a **bikini & flip flops**

NICK



Nick got a
beachball, sunglasses & flip flops

JOANNA



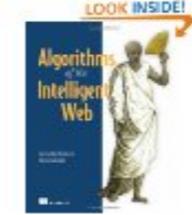
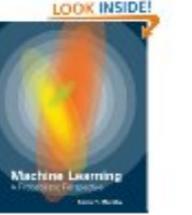
*What would you recommend
for Joanna?*

What is a **good** Recommender system?

Recommender systems in e-commerce

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

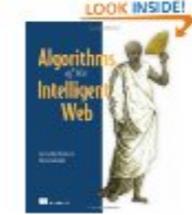
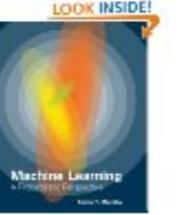
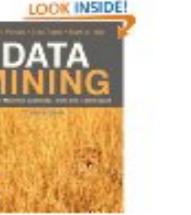
Customers Who Bought This Item Also Bought

 Recommender Systems Handbook Francesco Ricci Hardcover \$167.73	 Algorithms of the Intelligent Web Haralambos Marmanis 4.5 (14) Paperback \$26.76	 Programming Collective Intelligence: ... Toby Segaran 4.5 (91) Paperback \$25.20	 Machine Learning: A Probabilistic ... Kevin P. Murphy 4.5 (15) Hardcover \$81.00	 Data Mining: Practical Machine Learning ... Ian H. Witten 4.5 (29) Paperback \$42.61
---	---	---	---	---

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 Handbook Francesco Ricci Hardcover \$167.73	 Algorithms of the Intelligent Web Haralambos Marmanis ★★★★★ (14) Paperback \$26.76	 Programming Collective Intelligence: ... Toby Segaran ★★★★★ (91) Paperback \$25.20	 Machine Learning: A Probabilistic ... Kevin P. Murphy ★★★★★ (15) Hardcover \$81.00	 Data Mining: Practical Machine Learning ... Ian H. Witten ★★★★★ (29) Paperback \$42.61
---	---	---	---	---

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



Goals of the recommender systems?

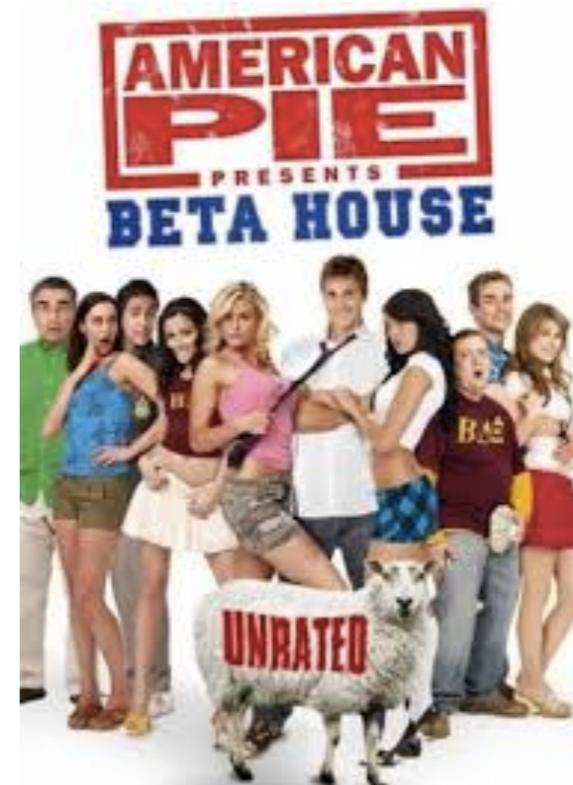
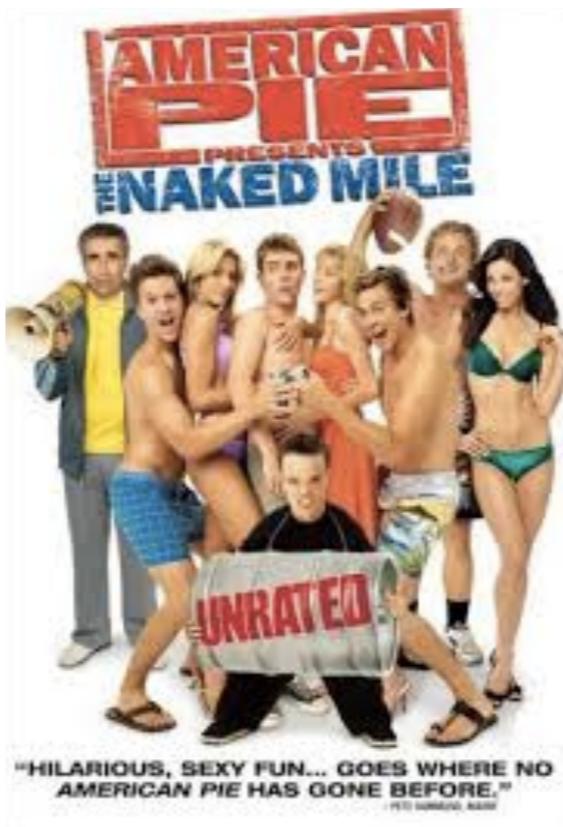
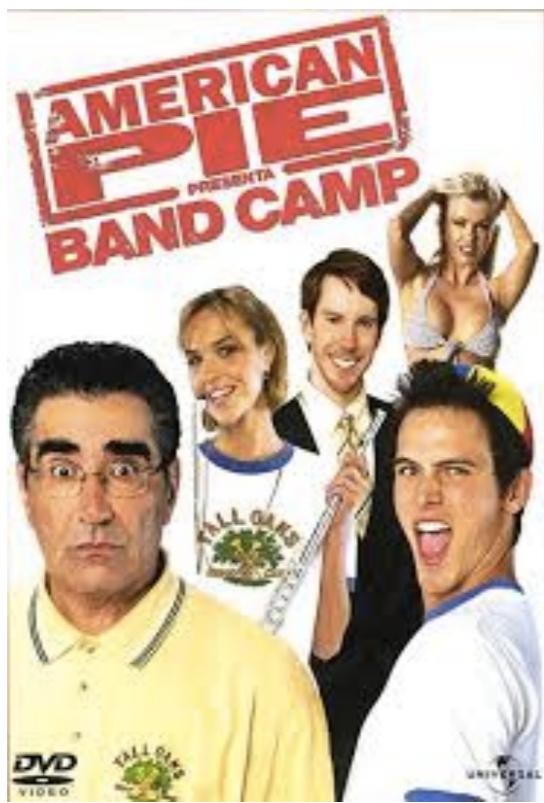
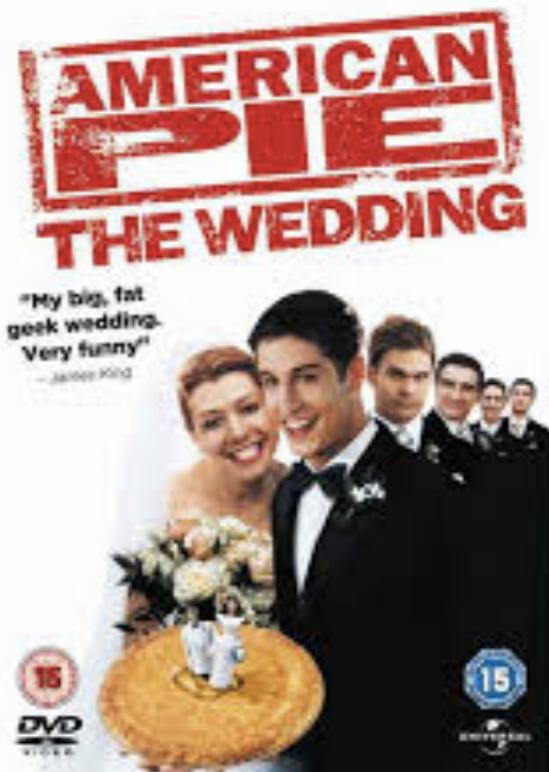
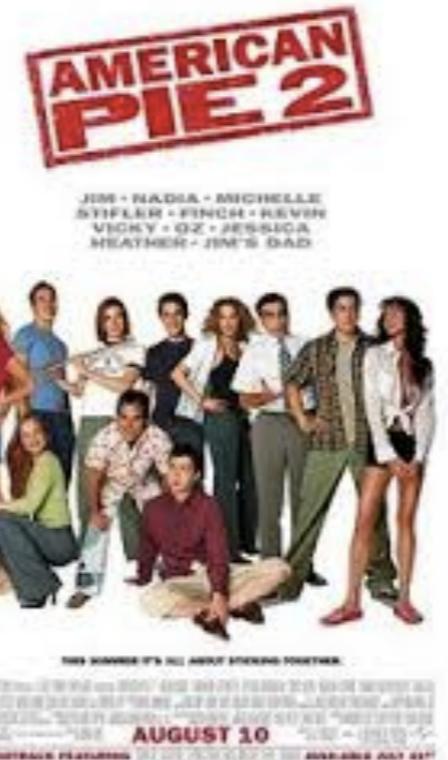
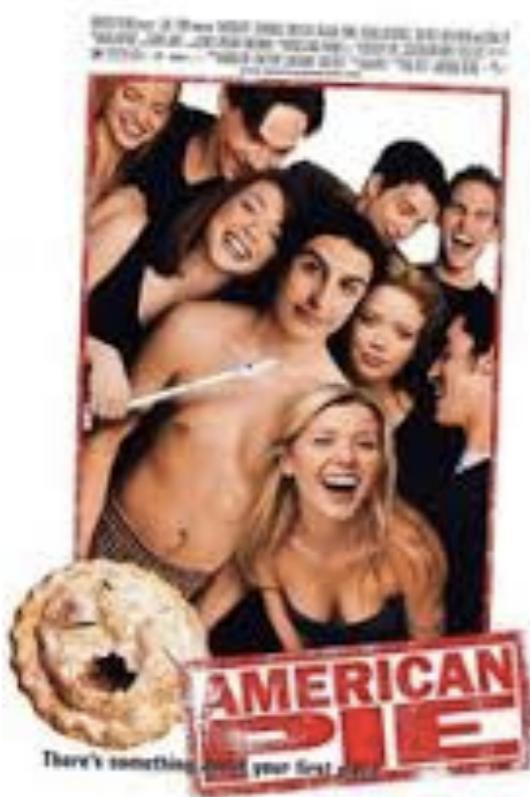
What do you think?

Goals of the recommender systems?

1st: Increase product sales or revenue

ok,
but this is the business-centric goal.

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
- Total number of sales
- Promotion of certain items
- Customer return rates
- Customer satisfaction and loyalty

**But, recommender
systems can change
user behaviors**

One critical point!!

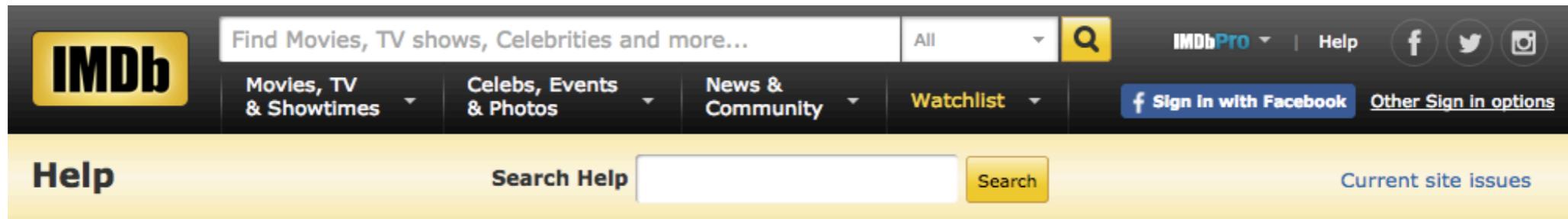
Trust is critical!!

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

Be careful!!

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

Be transparent!



The screenshot shows the top navigation bar of the IMDb website. It features the IMDb logo, a search bar with placeholder text "Find Movies, TV shows, Celebrities and more...", a dropdown menu set to "All", a search icon, and links for "IMDbPro", "Help", and social media icons for Facebook, Twitter, and Instagram. Below the main menu, there are four categories: "Movies, TV & Showtimes", "Celebs, Events & Photos", "News & Community", and "Watchlist". To the right of these are "Sign In with Facebook" and "Other Sign in options". A yellow banner at the bottom left says "Help", and a search bar with a "Search" button is on the right. At the very bottom, there's a link for "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!

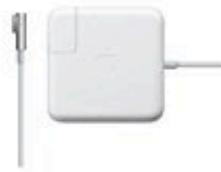


Note:

- 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

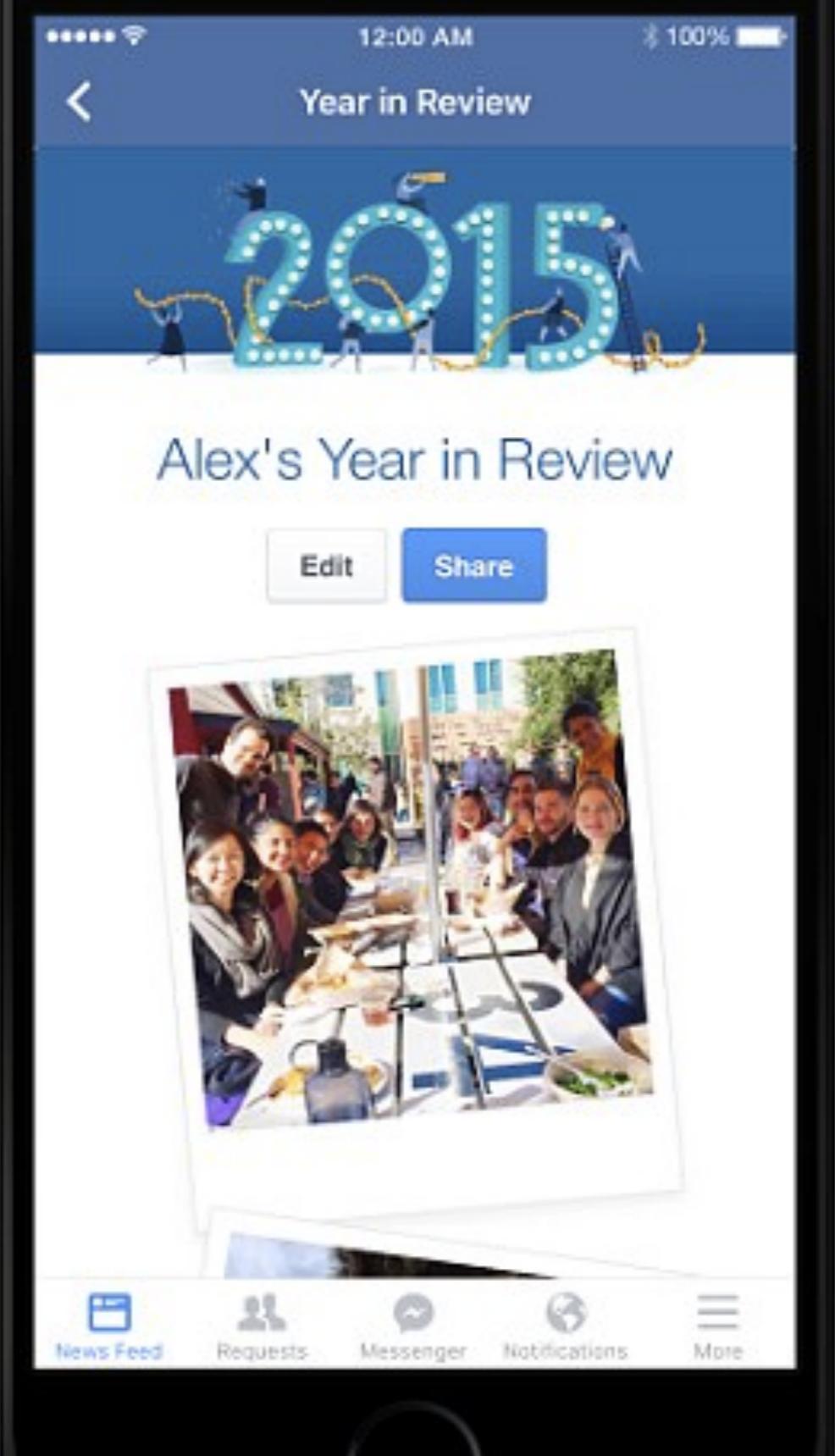
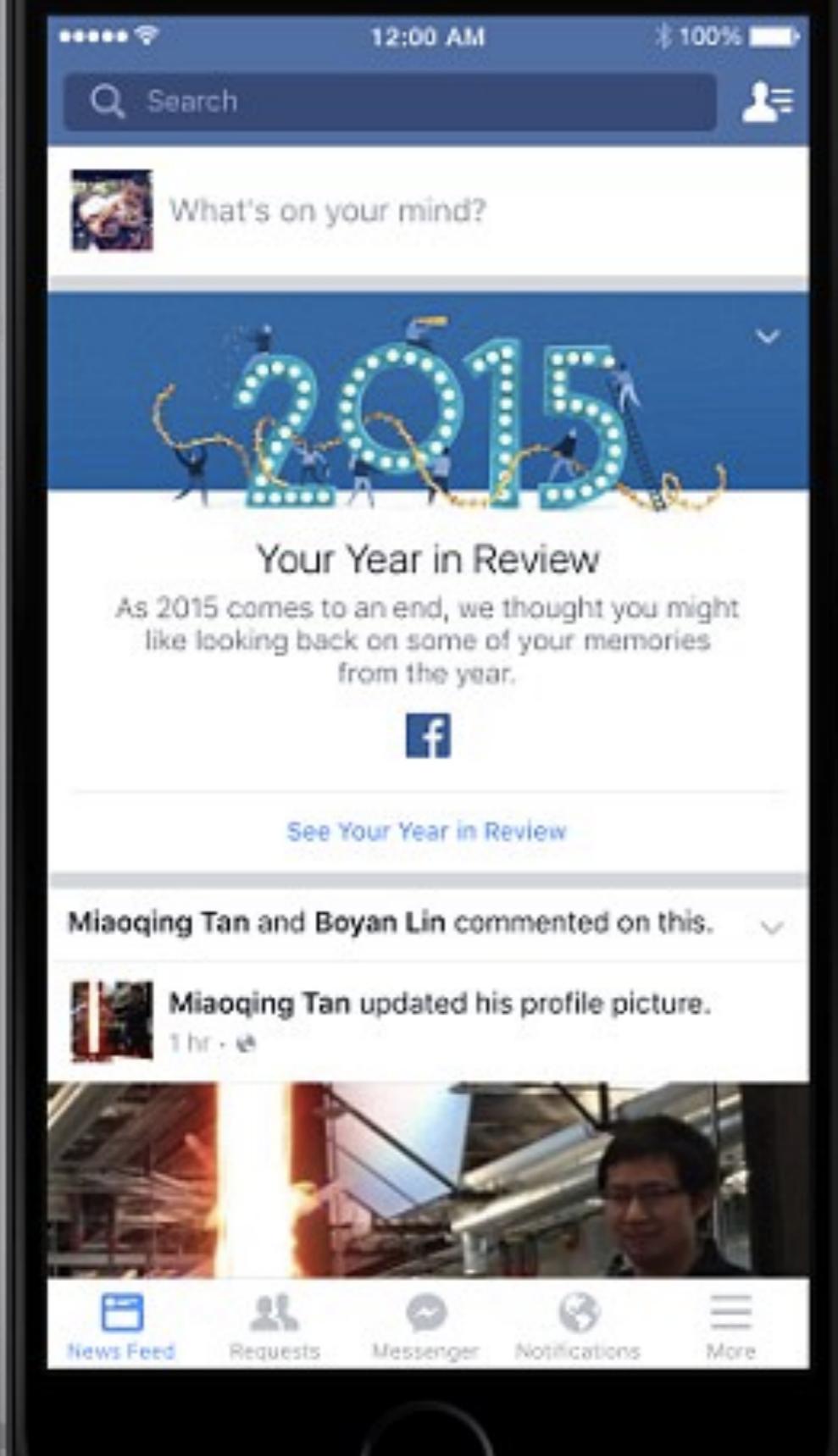
Añadir a la Lista de deseos

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!!



Which data recommender systems use?

Two types of data

Explicit



Filling out a form



Making a purchase



Favoriting an item

Implicit



Browsing activity



Searches



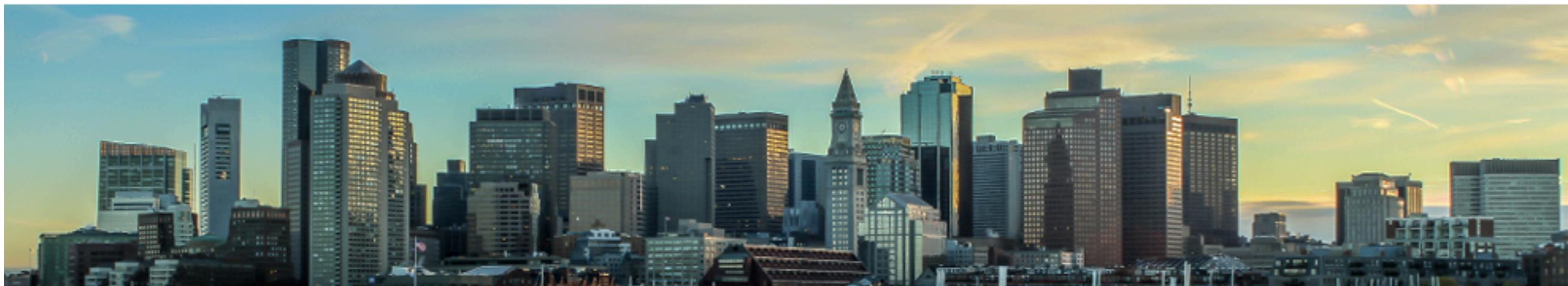
Location

Basic notation

- Some basic notation we will use throughout the course:
 - $m \times n$ rating matrix denoted by R , where m is the total number of users and n is the total number of items.
 - The (i, j) entry of R indicates the rating of user i to the item j , and it is denoted by r_{ij}
 - When the entry (i, j) , is estimated, or predicted, it is denoted by \hat{r}_{ij} with a “hat” symbol indicating that it is estimated.

Innovation on Recommender Systems





10th ACM Conference on Recommender Systems

Boston, MA, USA, 15th-19th September 2016

The ACM Recommender Systems conference (RecSys) is the premier international forum for the presentation of new research results, systems and techniques in the broad field of recommender systems. Recommendation is a particular form of information filtering, that exploits past behaviors and user similarities to generate a list of information items that is personally tailored to an end-user's preferences. As RecSys brings together the main international research groups working on recommender systems, along with many of the world's leading e-commerce companies, it has become the most important annual conference for the presentation and discussion of recommender systems research. RecSys 2016, the tenth conference in this series, will be held in Boston, MA, USA. It will bring together researchers and practitioners from academia and industry to present their latest results and identify new trends and challenges in providing recommendation components in a range of innovative application contexts. In addition to the main technical track, RecSys 2016 program will feature keynote and invited talks, tutorials covering state-of-the-art in



[RECSYS 2016 \(BOSTON\)](#)

[About the Conference](#)

[Program](#)

[Keynotes](#)

[Workshops](#)

[Tutorials](#)

[Doctoral Symposium](#)

[RecSys Challenge](#)



\$60,000 • 1,717 teams

Santander Product Recommendation

Wed 26 Oct 2016

Merger and Entry Deadline

Wed 21 Dec 2016 (6.5 days to go)

Dashboard

- Home
- Data
- Make a submission

Information

- Description
- Evaluation
- Rules
- Prizes
- Timeline

Forum

Kernels

- New Script
- New Notebook

Leaderboard

[Competition Details](#) » [Get the Data](#) » [Make a submission](#)

Can you pair products with people?

Ready to make a downpayment on your first house? Or looking to leverage the equity in the home you have? To support needs for a range of financial decisions, [Santander Bank](#) offers a lending hand to their customers through personalized product recommendations.



Public Leaderboard

1. Tom Van de Wiele
2. In Public Leaderboard We Trust
3. yoniko
4. idle_speculation
5. Alejo y Miro
6. Jared Turkewitz and BreakfastPirate
7. colun
8. Jack (Japan)
9. Sameh Faidi
10. DLmuvr

Under their current system, a small number of Santander's customers receive many recommendations while many others rarely see any resulting in an uneven customer experience. In their second competition, Santander is challenging Kagglers to predict which products their existing customers will use in the next month based on their past behavior and that of similar customers.

With a more effective recommendation system in place, Santander can better meet the individual needs of all customers and ensure their satisfaction no matter where they are in life.

Disclaimer: This data set does not include any real Santander Spain's customer, and



Expedia Hotel Recommendations

Which hotel type will an Expedia customer book?

\$25,000 · 1,974 teams · 8 months ago

[Overview](#)[Data](#)[Kernels](#)[Discussion](#)[Leaderboard](#)[More](#)[Submit Predictions](#)

Overview

[Description](#)[Evaluation](#)[Prizes](#)[Timeline](#)

Planning your dream vacation, or even a weekend escape, can be an overwhelming affair. With hundreds, even thousands, of hotels to choose from at every destination, it's difficult to know which will suit your personal preferences. Should you go with an old standby with those pillow mints you like, or risk a new hotel with a trendy pool bar?



Expedia wants to take the proverbial rabbit hole out of hotel search by providing personalized hotel recommendations to their users. This is no small task for a site with hundreds of millions of visitors every month!

Currently, Expedia uses search parameters to adjust their hotel recommendations, but there aren't enough customer specific data to personalize them for each user. In this competition, Expedia is challenging Kagglers to contextualize customer data and predict the likelihood a user will stay at 100 different hotel groups.

The data in this competition is a random selection from Expedia and is not representative of the overall statistics.

Complex Task & Research

Cross-Domain Recommendation

Context-aware Recommendation

Constraint-Based Recommendation

Recommendation and Privacy

Group Recommendation

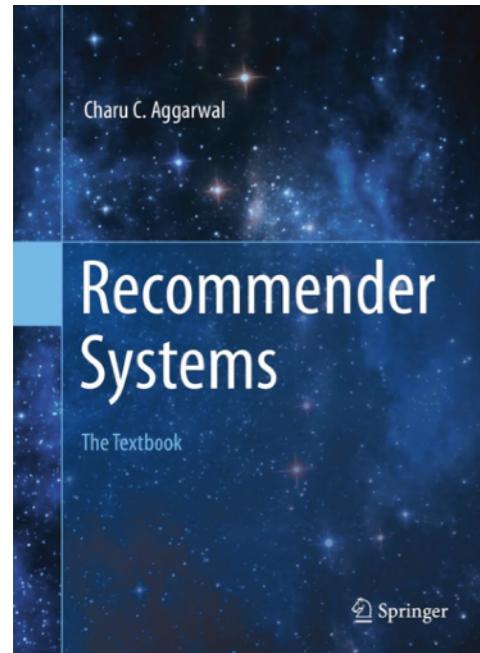
Recommendation as Intelligent Task Routing

Impact of Recommendations

New Domains for Recommendation

Recommendation and Social Networks

Bibliography



+



+

The Google logo, consisting of the word "Google" in its signature multi-colored font (blue, red, yellow, green).