UFMG UNIVERSIDADE FEDERAL DE MINAS GERAIS

Recommender Systems

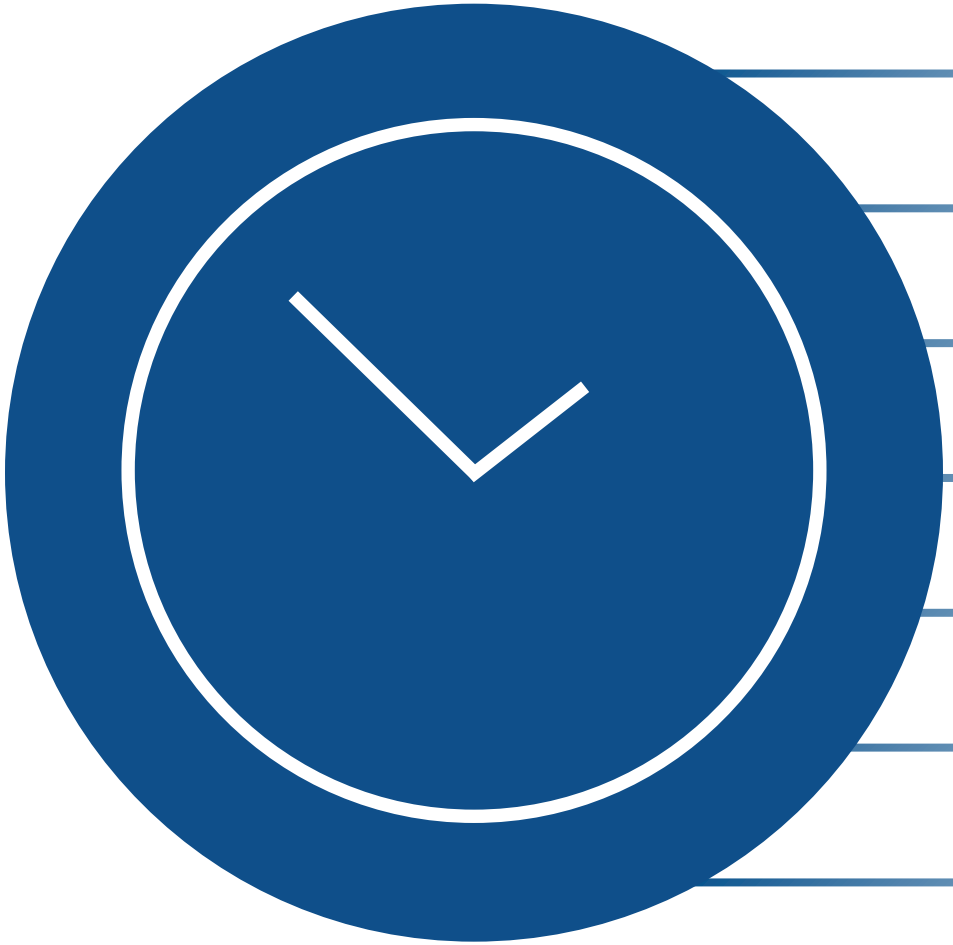
Introduction

Rodrygo L. T. Santos
rodrygo@dcc.ufmg.br

information overload

Credit: Keoni Cabral

Every minute of the day



Instagram *users share 65k photos*

Facebook *users share 240k photos*

Twitter *users post 575k tweets*

Netflix *users stream 452k hours*

YouTube *users stream 694k hours*

TikTok *users watch 167M videos*

Amazon *customers spend $283k*

Credit: Domo

Information overload

> *Global online content consumption is soaring in 2020, a new study of 10,000+ people in five countries says. The previous normal was just over 3h […] Average daily time spent consuming content is now ~7h, which includes phone, TV, and other forms of digital media.*
>
> ◦ Koetsier, Forbes 2020

The paradox of choice

*Which **digital camera** should I buy?*

*Which **destination** is the best for my holiday?*

*Which **investment** will support my children?*

*Which **movie** should I watch with my friends?*

*Which **book** should I buy for my next vacation?*

*Which **college degree** is best for my future?*

Credit: Anna Jumped

Search vs. recommendation

> *__Search__ is what you do when you're looking for something. __Discovery__ is when something wonderful that you didn't know existed, or didn't know how to ask for, finds you.*
>
> ◦ O'Brien, Fortune 2006

Credit: Bloomberg

The Netflix case

> *We think that the combined effect of **personalization** and **recommendations** save us **more than $1B** per year.*
>
> ◦ Gomes-Uribe & Hunt, TMIS 2015

Value for producers: revenue

Netflix: 66% of movies watched

Google News: 38% of news clicked

Amazon: 35% of products sold

The long tail

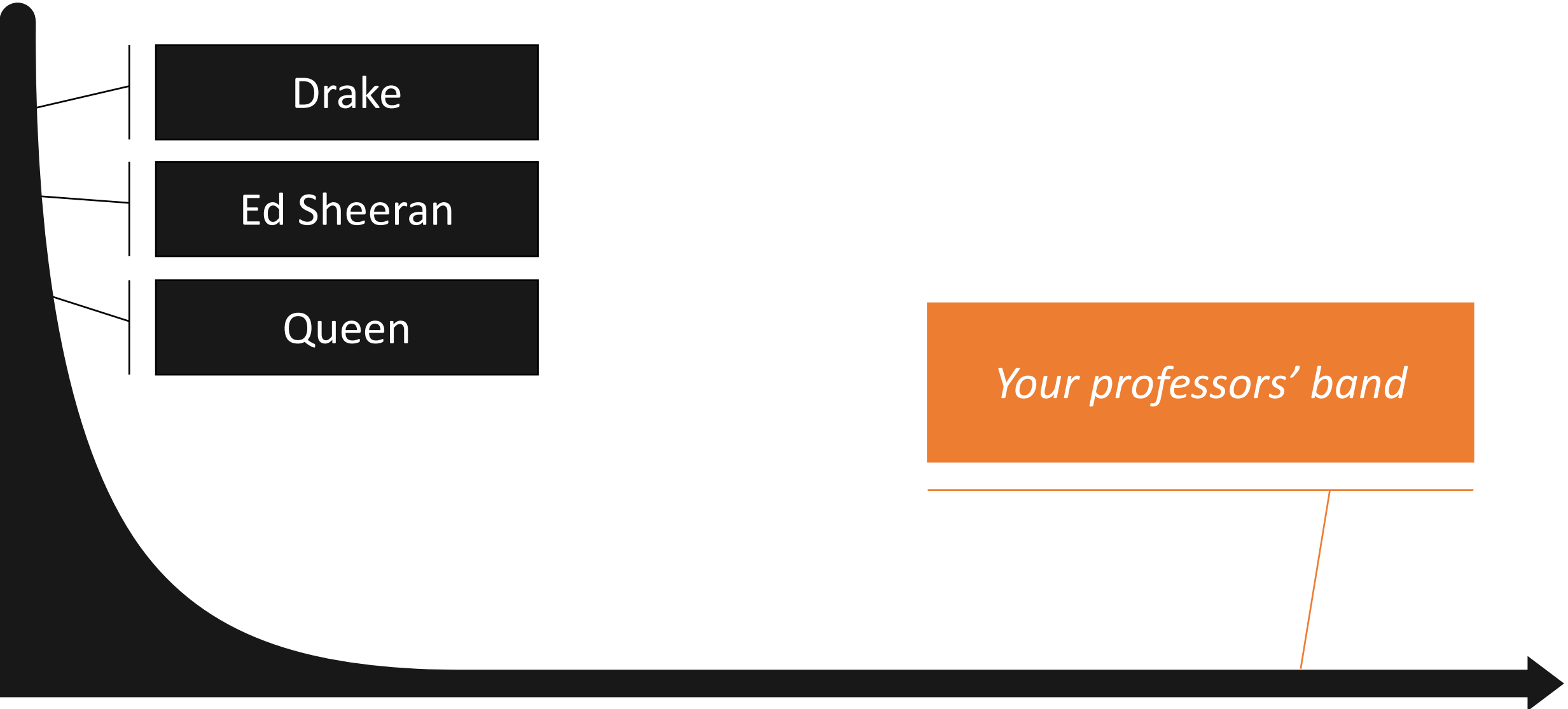
Concentration

- Selling a lot of a few items

Dissipation

- Selling less of lots of items

The long tail



Value for consumers: personalization

Recommendations tailored to individual needs

Credit: Matej Toman

Credit: Mike Seyfang

The Amazon case

> *If we have 4.5 million customers, we shouldn't have one store, we should have 4.5 million stores.*
>
> ◦ Jeff Bezos, Washington Post 1998

How did we get here?

Ancient days

Systems exploration

Rapid commercialization

Research explosion

Back to reality

Ancient days

Manual recommendations

- Public knowledge: *”certain snakes are venomous”*
- Word of mouth: *”the new restaurant has good price”*
- Friends’ advice: *”the new Marvel series is great”*
- Expert critics: *”merlots pair well with steak”*

Systems exploration

Xerox PARC's Tapestry [Goldberg et al., CACM 1992]

- Introduced the idea of collaborative filtering

UMN's GroupLens [Resnick et al., CSCW 1994]

- Automated collaborative filtering for news

Systems exploration

MIT's Ringo [Shardanand and Maes, CHI 1995]

- Automated collaborative filtering for music

BellCore's MovieRecommender [Hill et al., CHI 1995]

- Automated collaborative filtering for movies

Rapid commercialization

Scale and value became key challenges

- New algorithms to reduce computation time
  - e.g., item-based correlations, dimensionality reduction
- New evaluation approaches to better model users
  - e.g., ranking-based evaluation, online evaluation

Research explosion

Multidisciplinary fields

- Artificial intelligence
- Information retrieval
- Data mining
- Security and privacy
- Business and marketing

Research explosion

Further fueled by the Netflix Prize (2006)

- Baseline (Cinematch) RMSE = 0.9525
- USD 1M prize for 10% improvement

51K contestants, 41K teams, 186 countries

- 44K submissions by 5K teams
- Grand prize winner RMSE = 0.8567 (2009)

Netflix Prize

COMPLETED

Home | Rules | Leaderboard | Update

Leaderboard

Showing Test Score. Click here to show quiz score

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos				
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22
3	Grand Prize Team	0.8582	9.90	2009-07-10 21:24:40
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31
5	Vandelay Industries !	0.8591	9.81	2009-07-10 00:32:20
6	PragmaticTheory	0.8594	9.77	2009-06-24 12:06:56
7	BellKor in BigChaos	0.8601	9.70	2009-05-13 08:14:09

Back to reality

> *[…]* ***improved predictions*** *of just how much a user would dislike a set of bad movies* ***did not help*** *the user or Netflix*
>
> ◦ Konstan, Recommender Systems: An Introduction (2011)

The ACM Conference Series on
Recommender Systems

HOME RECSYS 2020 PAST CONFERENCES HONORS BLOG CONTACT search...

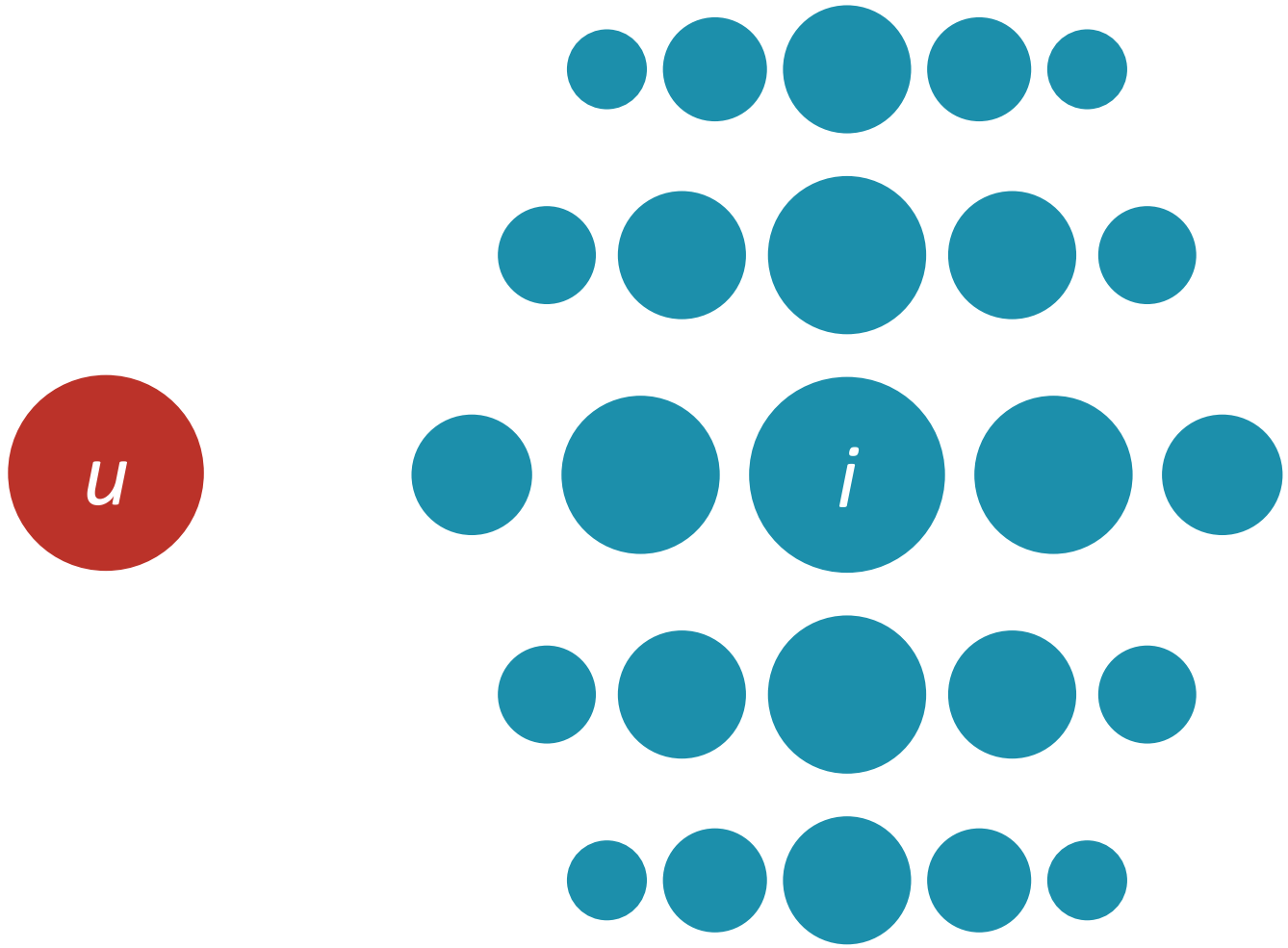
Recommender systems

> *“**Recommender systems** are software applications that aim to support users in their **decision-making** while interacting with large **information spaces**. They recommend **items** of interest to **users** based on **preferences** they have expressed, either explicitly or implicitly.*
>
> ◦ ACM RecSys

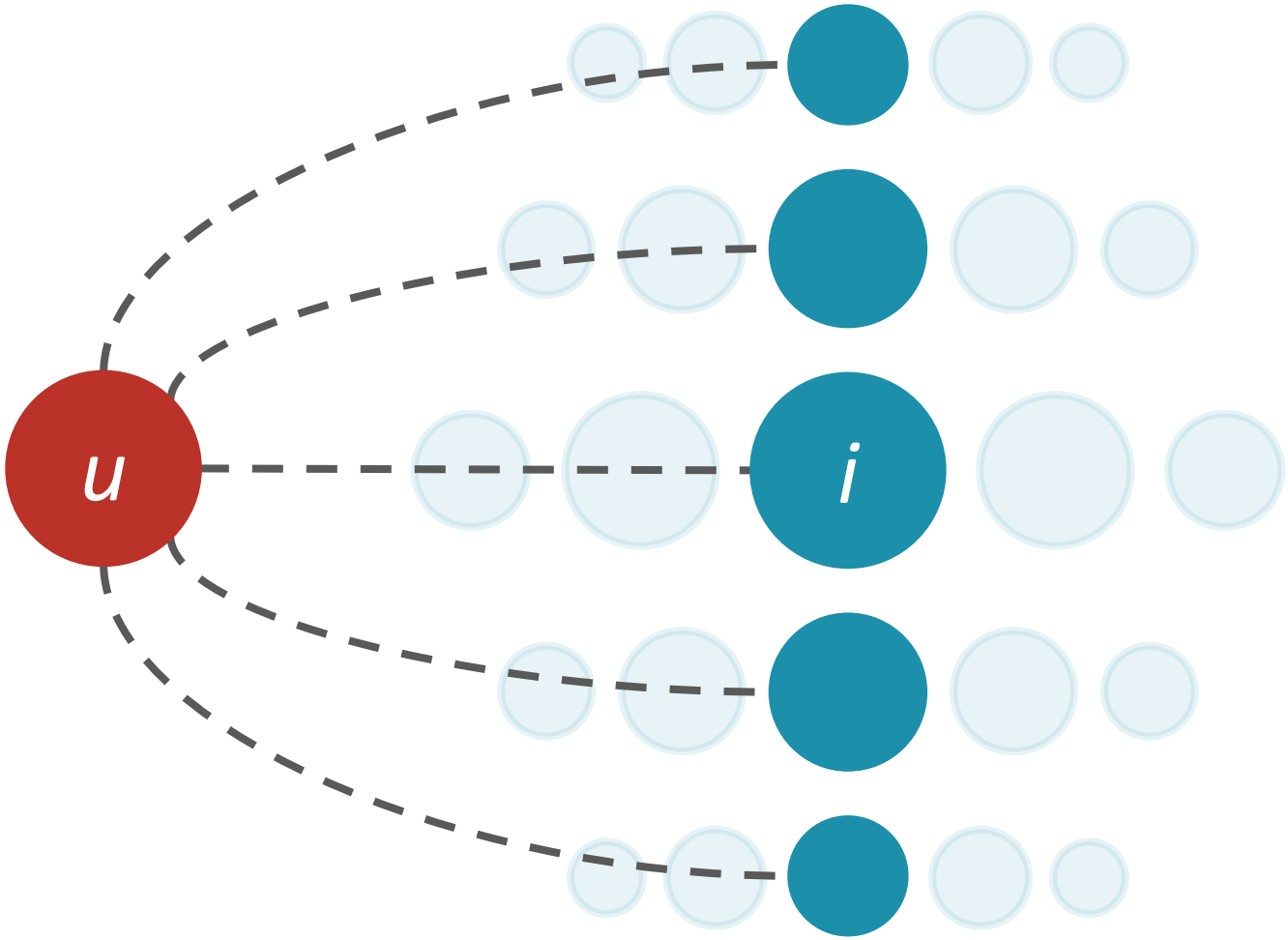
The recommendation problem

$$f(u, i)$$

The recommendation problem



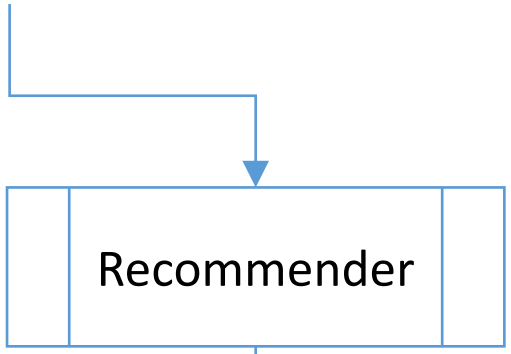
The recommendation problem



What to recommend?

How to recommend?

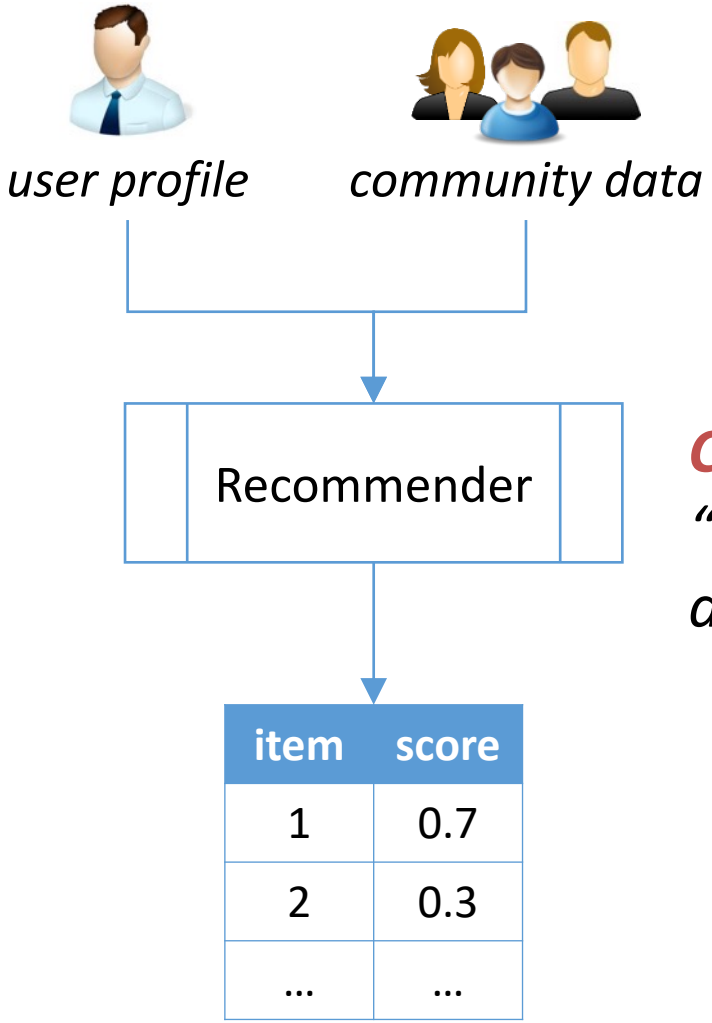
user profile



item	score
1	0.7
2	0.3
…	…

Personalized recommendation

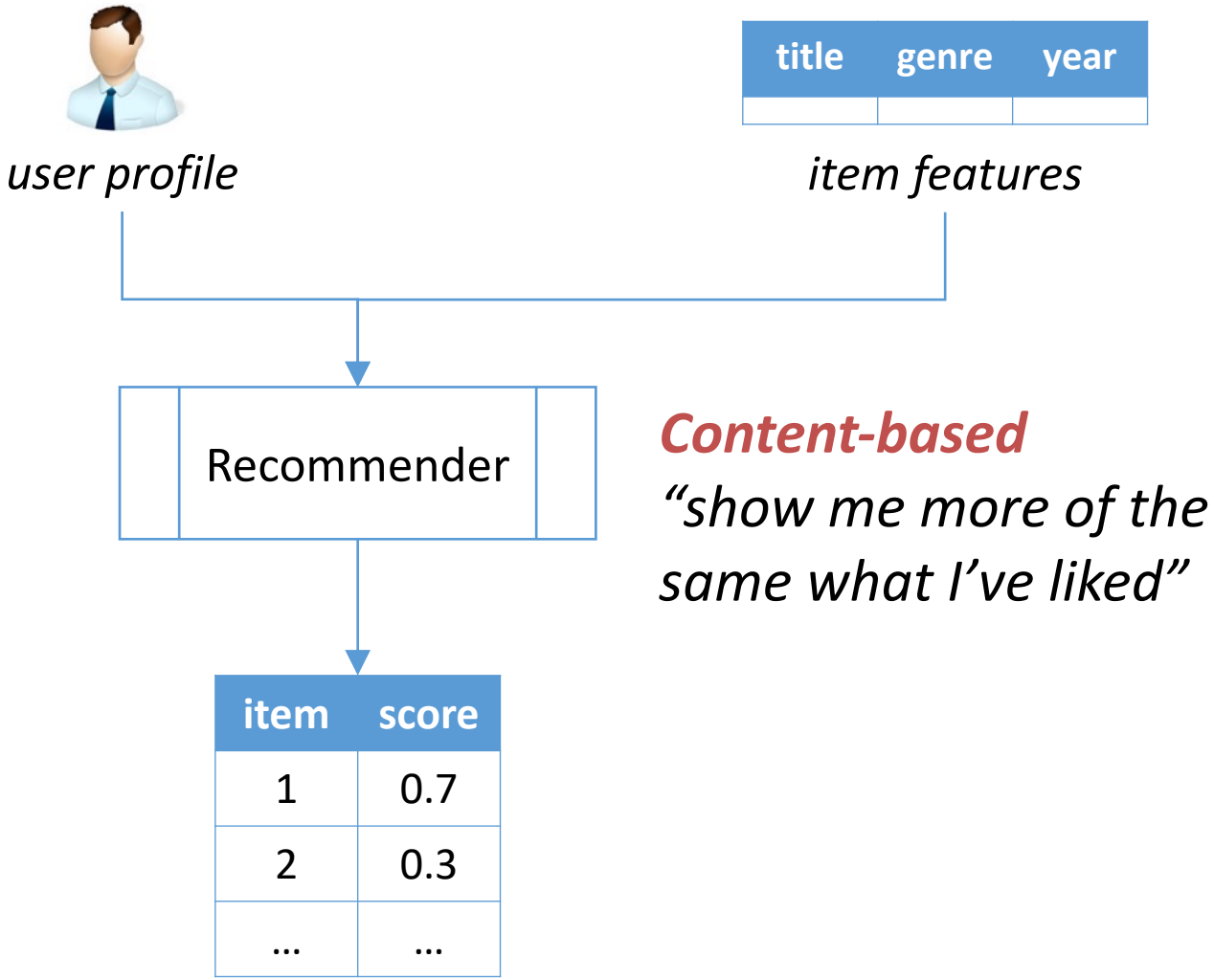
How to recommend?



Collaborative filtering

"tell me what's popular among my peers"

How to recommend?

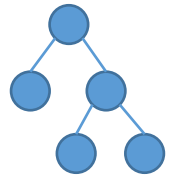


Content-based

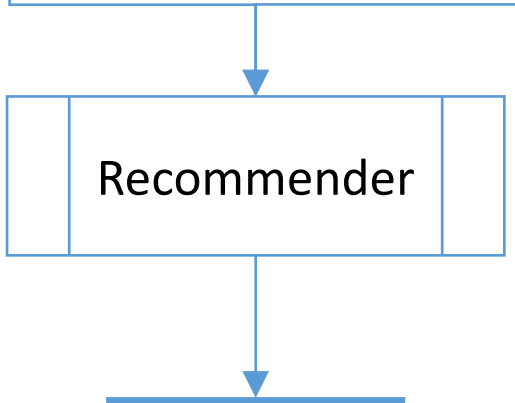
"show me more of the same what I've liked"

How to recommend?

user profile



knowledge models

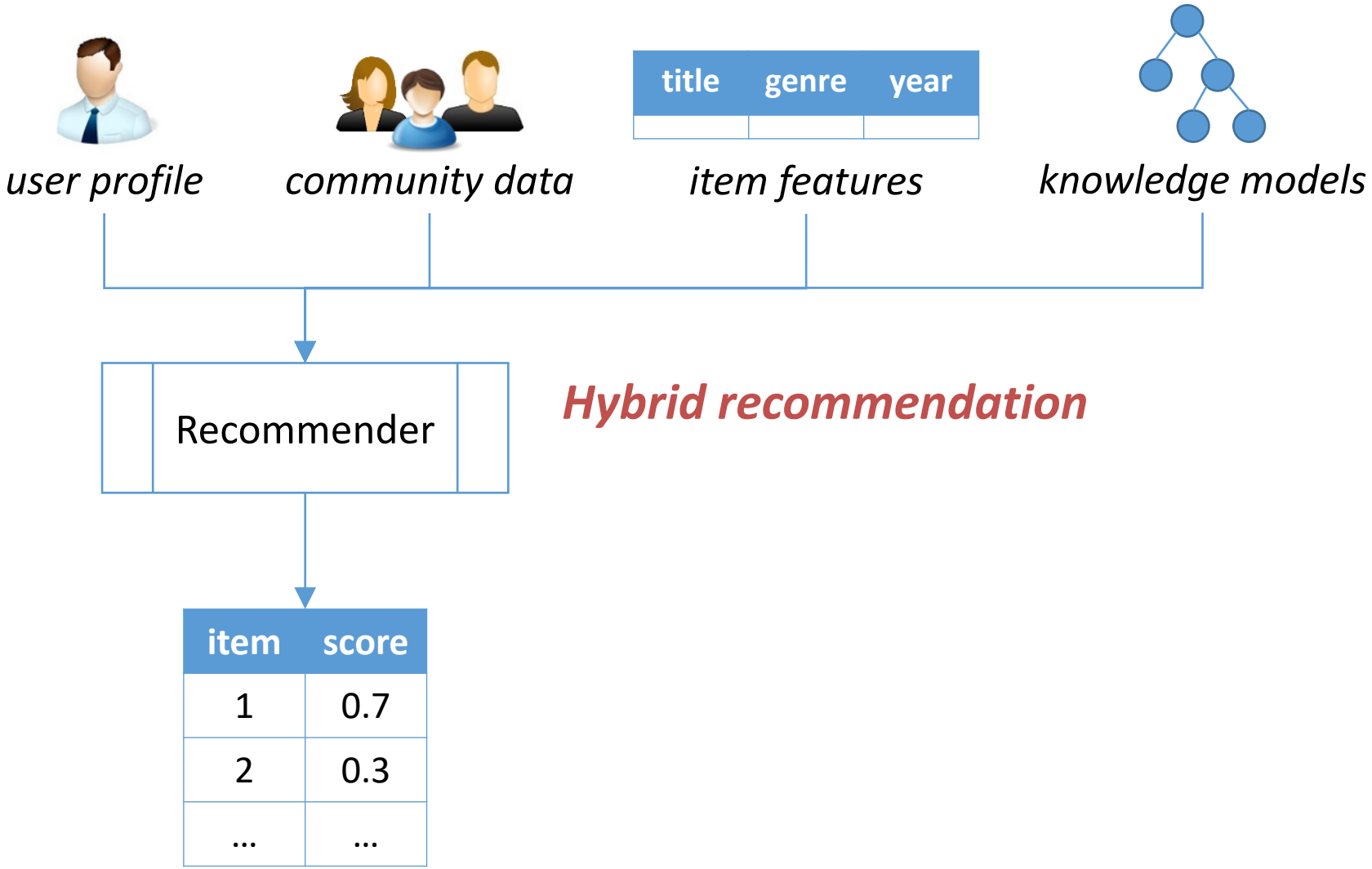


Knowledge-based

"tell me what fits based on my needs"

item	score
1	0.7
2	0.3
…	…

How to recommend?



Course scope

Focus on algorithms

- Collaborative
- Content-based
- Knowledge-based
- Hybrid

Course scope

Focus on evaluation

- Methodology and metrics

A taste of advanced topics

- Diversity and novelty
- Context-awareness
- Machine-learned models

Out-of-scope

We have dedicated courses for:

- Information retrieval
- Machine learning
- Data mining

Course goals

At the end, you should be able to:

- Identify potential application domains
- Implement basic recommender systems
- Critique a design to identify potential strengths and weaknesses and to compare alternatives

Textbooks

Recommender Systems: An Introduction

- Jannach, Zanker, Felfernig, Friedrich (2011)

Recommender Systems: The Textbook

- Aggarwal (2016)

Recommender Systems Handbook (3rd edition)

- Ricci, Rokach, Shapira (2022)

Other relevant material

General background

- Algorithms and data structures
- Basic statistics
- Basic linear algebra

Advanced readings

- Google Scholar is your friend

Course grading (tentative)

Assignments: 10

Paper seminars: 10

Research challenges: 40

Exams: 40

Course attendance

“

O que é necessário para ser aprovado em uma dada atividade acadêmica curricular?

É necessário obter nota final igual ou superior a 60, em uma escala de 0 a 100, bem como a indicação de assiduidade, a qual deve ser igual ou superior a 75% (art. 12 das NGG).

Pre-course survey

Fill in a short survey describing your past experience and expectations related to the course

- https://forms.gle/7mcatGc5LtAFM2ta7