

Módulo 01: Introdução aos Sistemas de Recomendação



Aula 01: Motivação

Digitalização das lojas

Lojas Físicas

Espaço físico limitado (prateleiras)

Centenas de produtos

Clientes centralizados

Lojas Digitais

Espaço "ilimitado" (páginas na web)

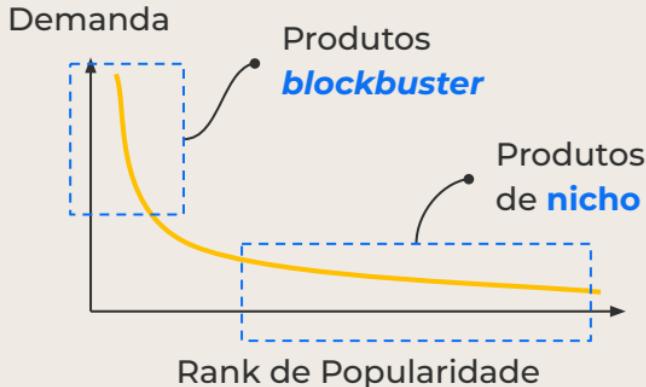
Milhares/milhões de produtos

Clientes descentralizados



Teoria da cauda-longa

A receita gerada por produtos de **nicho** pode ser tão alta quanto a receita de produtos **blockbuster**.



Como filtrar os
produtos mais
compatíveis com o
usuário a partir do
seu perfil de
consumo?



SISTEMAS DE
RECOMENDAÇÃO

E-COMMERCE

mercado
libre

americanas

olx

amazon.com

STREAMING

Y N

S Disney+

Empresas e
Setores

P

REDES SOCIAIS

in

Snap

Tik

G



C

yahoo!

BBC
NEWS

NOTÍCIAS

ifood
B. H

SERVIÇOS

Resultados publicados

↑ 38%

Aumento de **CTR** em relação ao algoritmo Top (*Google, 2006*)

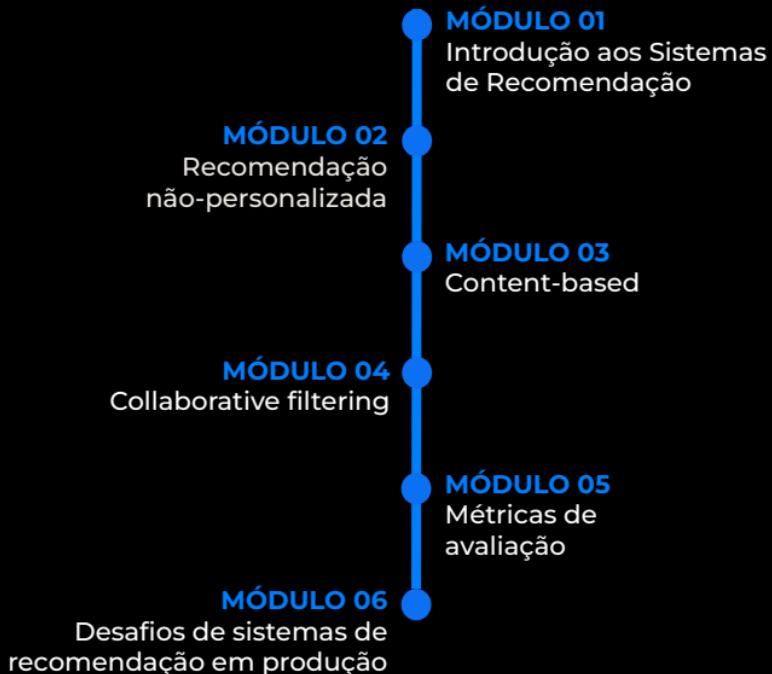
80%

Filmes assistidos através da recomendação (*Netflix, 2016*)

30%

Pageviews através da recomendação (*Amazon.com, 2017*)

Estrutura do curso





João Felipe Guedes

Senior Data Scientist @ Globo



MJV





Na próxima aula...

Elementos de um sistema de recomendação



Aula 02: **Elementos de um Sistema de Recomendação**

Matriz de preferências

Relaciona a entidade **usuário** com a entidade **item** através de um sinal de **feedback**.

$$\mathbf{R} = \left\{ \begin{array}{cccc} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nm} \end{array} \right\}$$

Feedback

Usuário u_1

Item i_2

Sinais de feedback

Feedback explícito

Registro **intencional** do usuário sobre sua preferência



X

Feedback implícito

Coletado através da **navegação do usuário**.



- Fácil obtenção
- Consumo ***near real time***
- Sinal **ruidoso**
- **Baixa qualidade** na preferência do usuário

- Difícil obtenção
- Distante do consumo
- Matriz altamente **esparsa**
- **Alta qualidade** na preferência do usuário

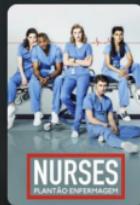
Matriz de similaridades

Trailers
Episódios
e mais
Similares

Indica o grau de similaridade entre 2 entidades (itens ou usuários) a partir dos **metadados ou dos consumos**.

$$\mathbf{I} = \begin{Bmatrix} i_{11} & i_{12} & \dots & i_{1M} \\ i_{21} & i_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ i_{M1} & i_{M2} & \dots & i_{MM} \end{Bmatrix}$$

$$\mathbf{U} = \begin{Bmatrix} u_{11} & u_{12} & \dots & u_{1N} \\ u_{21} & u_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & \dots & u_{NN} \end{Bmatrix}$$



Nurses -
Plantão
Enfermagem



House



The Night
Shift



Chicago Med:
Atendimento
de
Emergência



Transplant:
Uma Nova
Vida
2ª TEMPORADA



Hospital New
Amsterdam



ENFERMEIROS



DR. DEATH

Seed para uma recomendação

ITEM

Utiliza-se um **item** como base para recomendação de outros itens

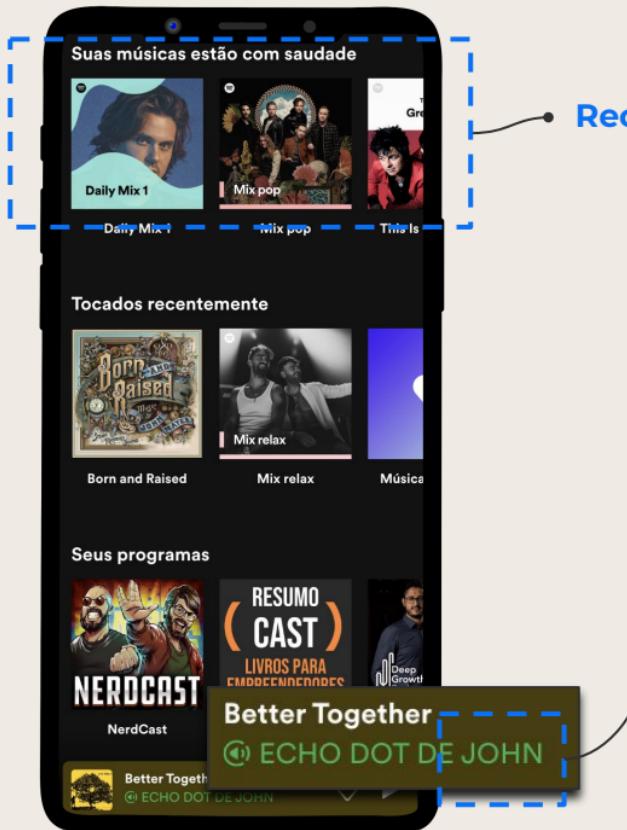
Seed

- "Similares à *{seed}*"
- "Itens parecidos com *{seed}*"
- "Porque você comprou *{seed}*"



Rec

Seed para uma recomendação



USER

Utiliza-se o perfil do **usuário** como base para recomendação de outros itens

- "Para **Você**"
- "**Suas** músicas estão com saudade"
- "Descubra"



Na próxima aula...

Topologias de Sistemas de Recomendação



Aula 03: Topologias de Sistemas de Recomendação

Não-personalizado



Content-based



Collaborative Filtering

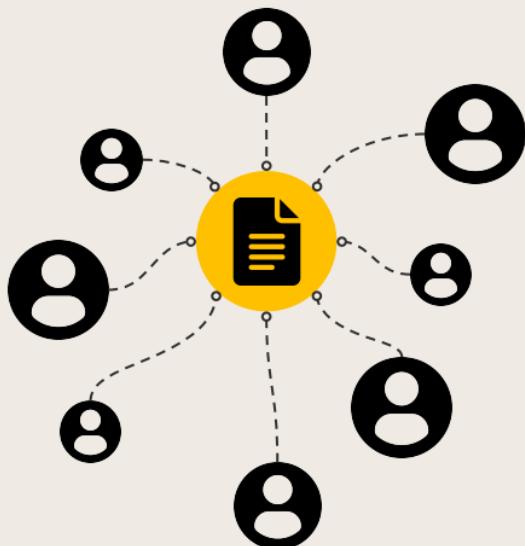


Recomendação Híbrida



Não-personalizado

Mesma recomendação para todos os usuários



- ***Cold-start*** de usuário
- Rápida atualização (streaming)
- Baixa complexidade
- Não considera o **perfil de consumo**

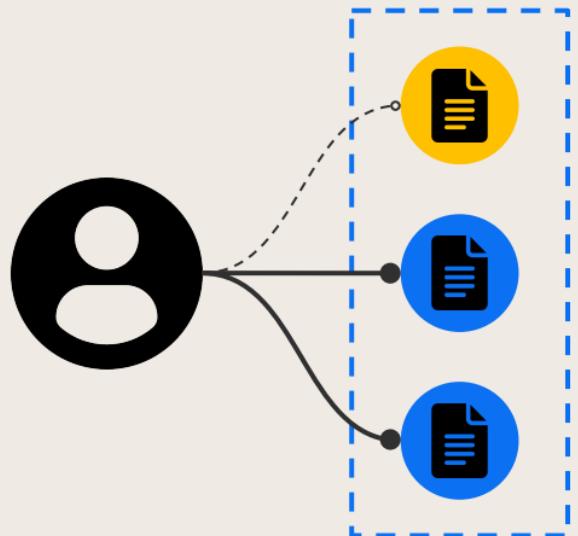
Não-personalizado

Mesma recomendação para todos os usuários



- ***Cold-start*** de usuário
- Rápida atualização (streaming)
- **Baixa complexidade**
- Não considera o **perfil de consumo**

Itens similares



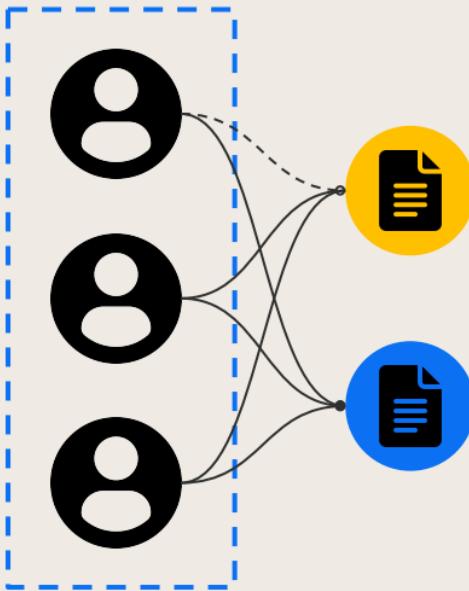
Foco nas **features** dos itens

- **Cold-start** de item
- **Alta** explicabilidade
- **Extração** de *features*
- Efeito **bolha**

● Item consumido ● Item recomendável

Collaborative filtering

Consumos similares

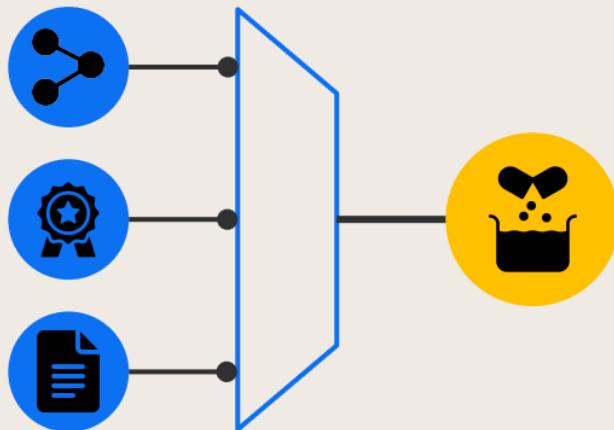


● Item consumido ● Item recomendável

Foco nos *sinais de feedback* dos usuários

- Desempenho
- Baixa explicabilidade
- Esparsidade
- Maior serendipidade

Recomendação Híbrida



Combina a saída de diferentes topologias

Trabalha as **vantagens** e **desvantagens** dos modelos de entrada

Maior complexidade de implementação

Mais **hiperparâmetros**

*Hybrid Recommender Systems:
Survey and Experiments
(R. Burke, 2002)*



Na próxima aula...

Exemplo prático de um dataset para recomendação



Aula 04: Exemplo prático de um dataset para recomendação

Análise exploratória do MovieLens





Obrigado.



Aula 07: Referências de estudo

Conferências



SIGIR
Special Interest Group
on Information Retrieval

SPONSORS

NETFLIX

Booking.com

Google

 **Meta**

 **Disney+**

 **amazon**

Blogs de pesquisa



research.netflix.com



research.atspotify.com



research.google

Frameworks em python





Obrigado.

Módulo 02: Recomendação Não-personalizada



Aula 08: Algoritmos de recomendação não-personalizada

O que é uma recomendação não-personalizada?

Não-personalizada

Mesma recomendação para todos os usuários ou itens

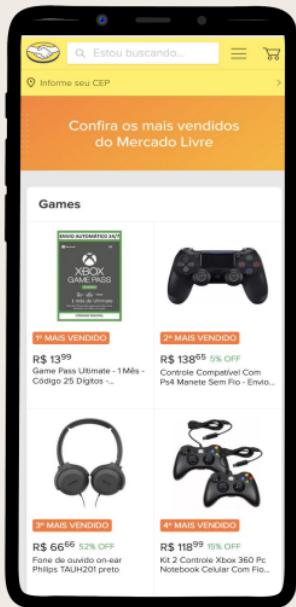
X

Personalizada

Recomendação orientada às características específicas do usuário ou do item

Exemplos de recomendação não-personalizada

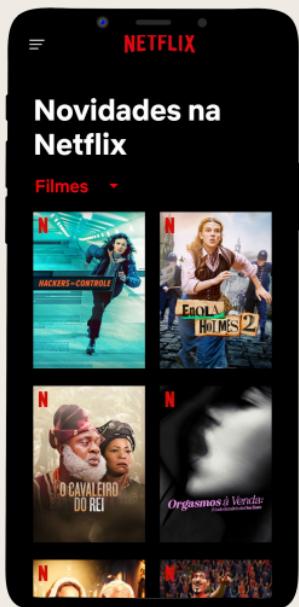
Top



Trending



Novidades



Por que uma recomendação não-personalizada?

1. Algoritmos de **menor complexidade**
2. Usuários **anônimos**
3. Problemas de ***Cold-start***
4. **Fallback** para a recomendação personalizada
5. **Baseline** para experimentos A/B de personalização



Na próxima aula...

Exemplo prático de recomendação Top N



Aula 09: Exemplo Prático de Recomendação Top N

Recomendação Top N





Na próxima aula...

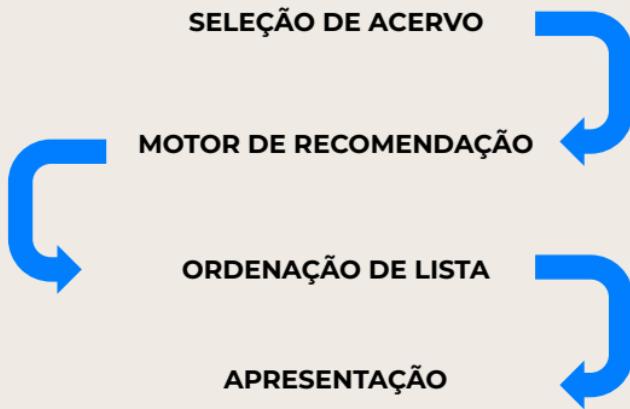
Seleção de acervo e regras de negócios



Aula 11:

Seleção de Acervo e Regras de Negócio

Etapas da recomendação

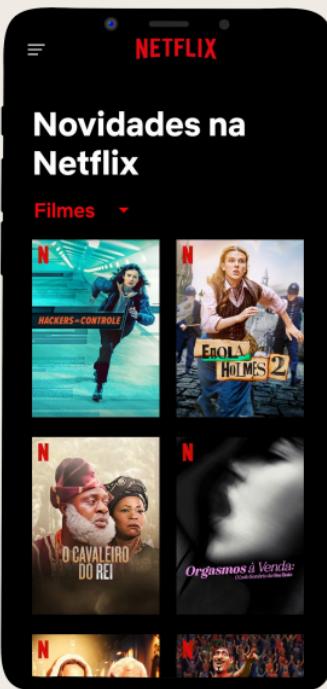


Filtros temporais

Qual deve ser a recência dos itens a serem recomendados ?

- Últimas 24h
- Último mês
- Sem recência

Alta vida útil



Baixa vida útil



Filtros de categoria

Qual categoria de itens deve ser filtrada para a oferta recomendada ?

- Área do produto
- Disponibilidade por país
- Faixa etária



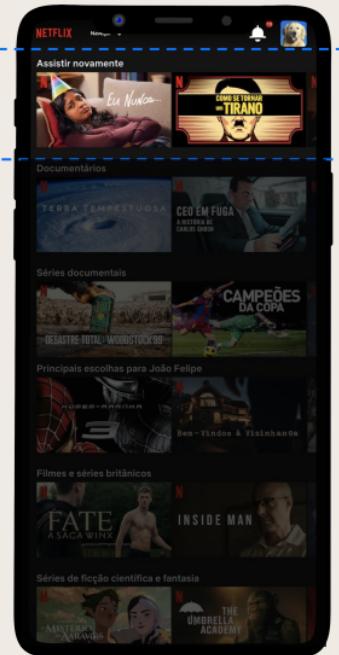
Filtros de consumo

Devemos considerar itens que o usuário já viu ?

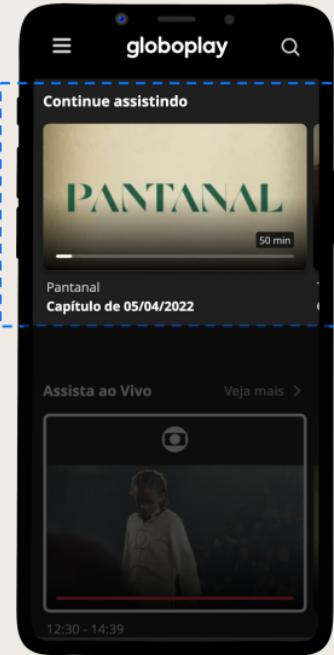
- *Compre novamente*
- *Continue assistindo*
- *Descubra*

Presentes em geral na recomendação personalizada

Consumo completo



Consumo parcial





Na próxima aula...

Exemplo prático de seleção de acervo



Aula 08: Exemplo Prático de Seleção de Acervo

Recomendação trending





Obrigado.

Módulo 03: Recomendação Content-based

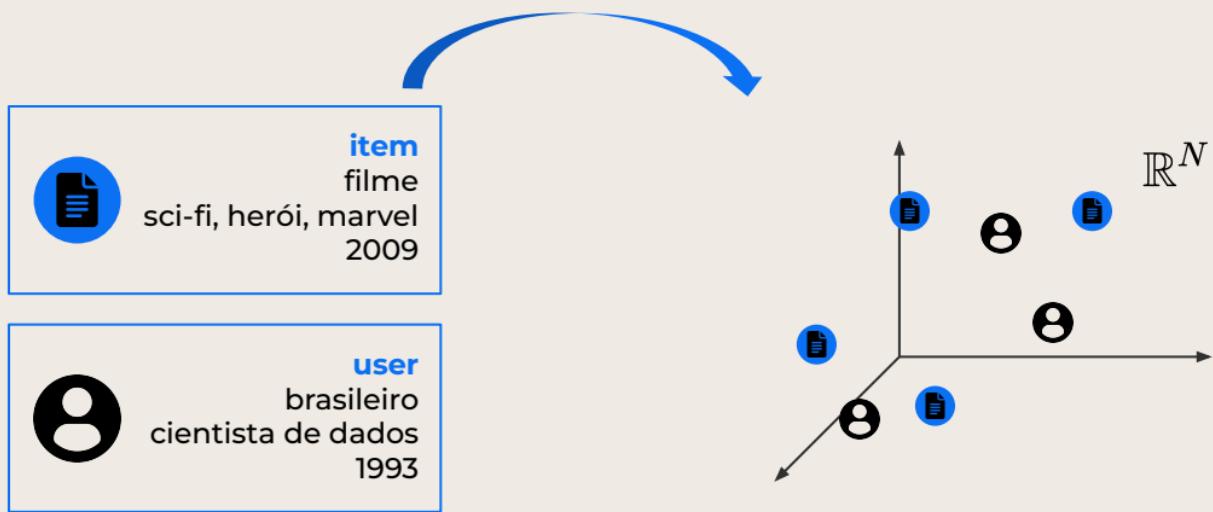


Aula 14:

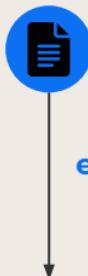
Representações Vetoriais e Similaridades

Representações vetoriais

O objetivo da representação vetorial é levar os **itens** e os **usuários** em para um espaço onde eles possam ser **comparados**.



Geradores de vetores



$$\mathbf{i}_a = [i_{a0}, i_{a1}, i_{a2}, i_{a3}, \dots, i_{aN}]^T$$

- One-hot encoding
- Principal Component Analysis
- Autoencoders
- Doc2vec
- Convolutional Neural Networks
- wav2vec
- Transformers

Em alguns casos os vetores são chamados de **embeddings** quando a informação é **compactada** via **machine learning**.

Cálculo de similaridade

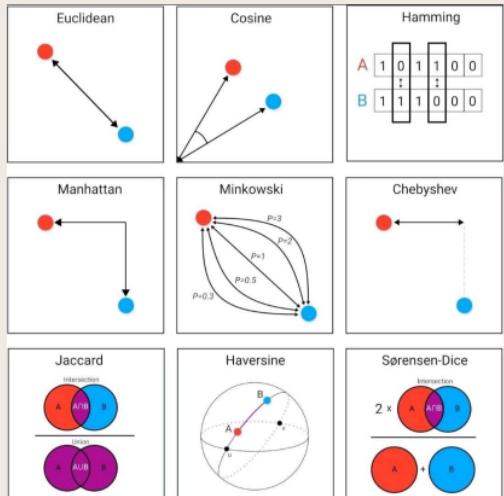
$$\mathbf{i}_a = [i_{a0}, i_{a1}, i_{a2}, i_{a3}, \dots, i_{aN}]^T$$

$\rho(\cdot)$

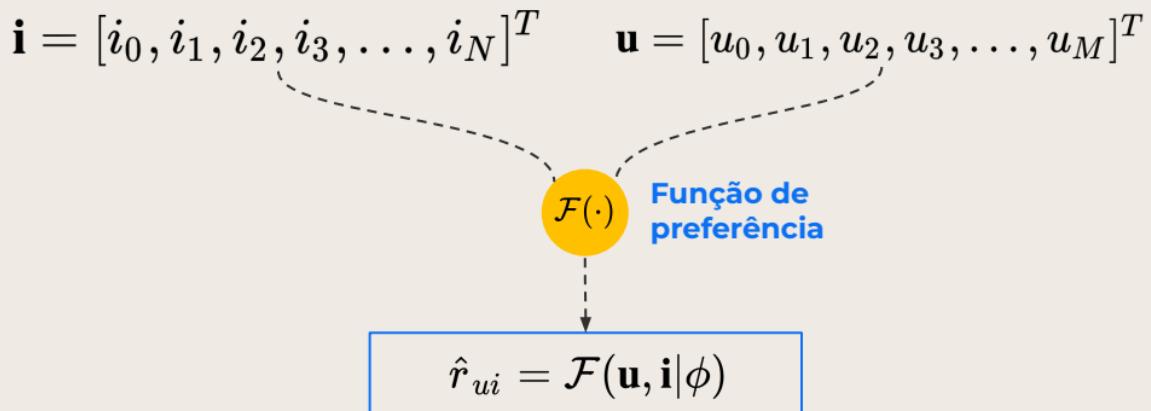
$$similarity_{ab} = \rho(\mathbf{i}_a, \mathbf{i}_b)$$

$$\mathbf{i}_b = [i_{b0}, i_{b1}, i_{b2}, i_{b3}, \dots, i_{bN}]^T$$

- Distância euclidiana
- Similaridade cosseno
- Distância Jaccard
- Distância Chebyshev



Função de preferência





Na próxima aula...

Exemplo prático de recomendação por similaridade



Aula 10:

Exemplo Prático de Recomendação por Similaridade (parte 1)

Recomendação de jogos da Steam





Na próxima aula...

Geração de embeddings com PCA



Aula 11:

Geração de Embeddings com PCA

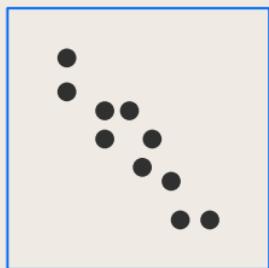
Covariância

$$cov(x, y) = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

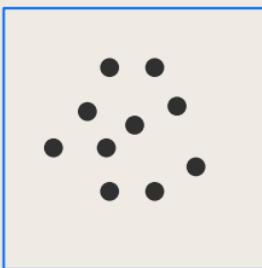
$$cov(x, y) < 0$$

$$cov(x, y) \approx 0$$

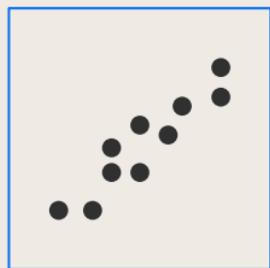
$$cov(x, y) > 0$$



Features
Dependentes
(redundantes)

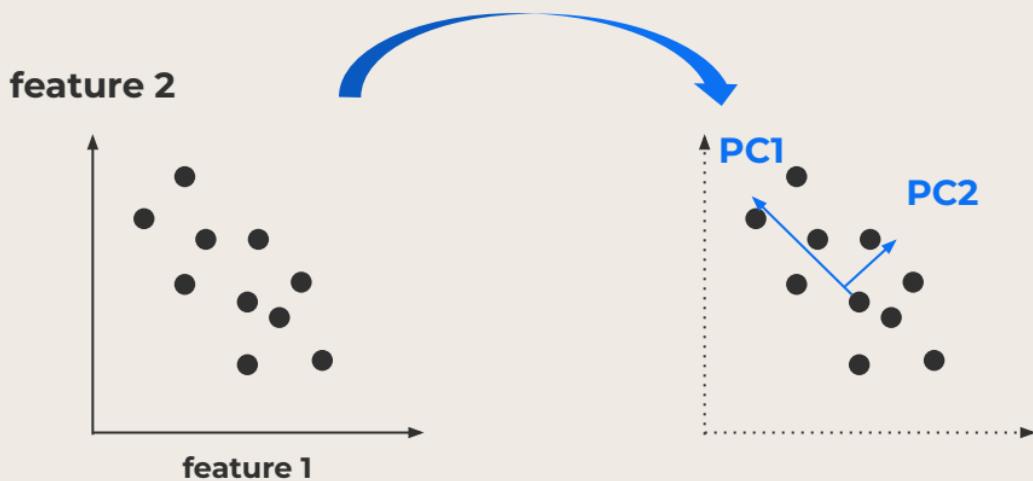


Features
Independentes

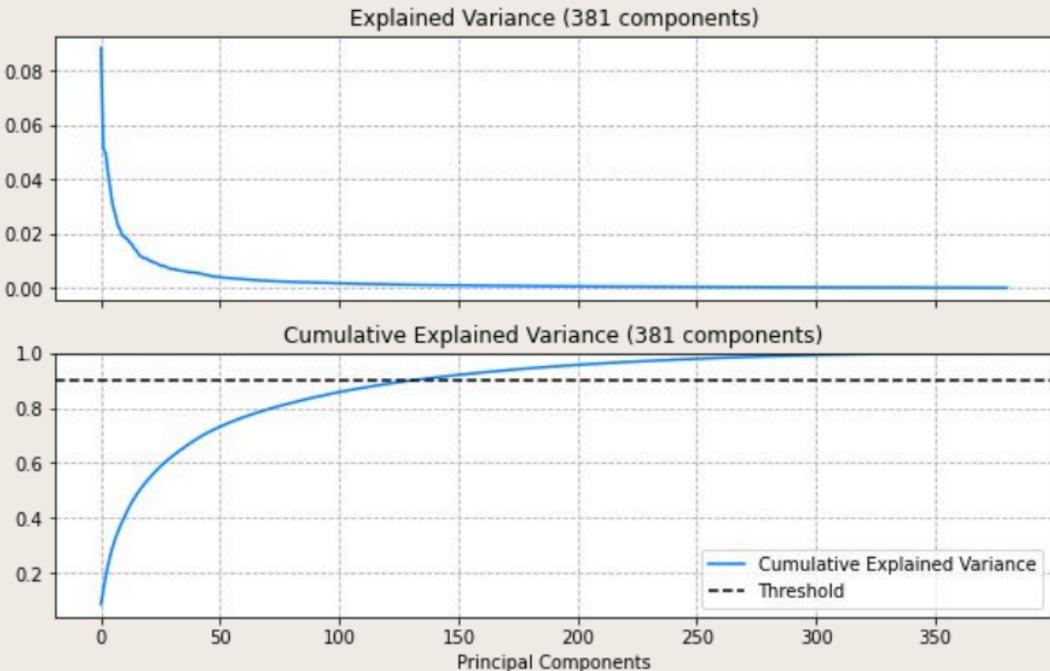


Motivação da PCA

A PCA explora a **covariância** entre as *features* para criar um novo conjunto de coordenadas **ortogonais** (linearmente independentes) onde **cada coordenada explica uma parte da variância dos dados**.



Redução de dimensionalidade



Simplificando a matemática

Onde queremos chegar?

$$X'_{k \times n} = \overbrace{T_{k \times m}}^{\text{Transformação}} \cdot \underbrace{X_{m \times n}}_{\text{Dados originais (dimensão m)}}$$

Dados transformados com $k << m$

O que precisamos encontrar?

Matriz de transformação T

Solução da PCA

Solução simplificada

T é a matriz formada pelos **primeiros k autovetores** da matriz de covariância dos dados originais **ordenados pelos autovalores** de forma decrescente.

$$\mathbf{C}_X = Cov(\mathbf{X}) = \begin{pmatrix} cov(x_1, x_1) & cov(x_1, x_2) & \dots & cov(x_1, x_n) \\ cov(x_2, x_1) & cov(x_2, x_2) & \dots & cov(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(x_n, x_1) & cov(x_n, x_2) & \dots & cov(x_n, x_n) \end{pmatrix}$$

$$\mathbf{C}_x \cdot \mathbf{T} = \mathbf{\Lambda} \cdot \mathbf{T}$$



Na próxima aula...

Exemplo prático de recomendação por similaridade (parte 2)



Aula 12:

Exemplo Prático de Recomendação por Similaridade (parte 2)

Recomendação de jogos da Steam com PCA





Obrigado.

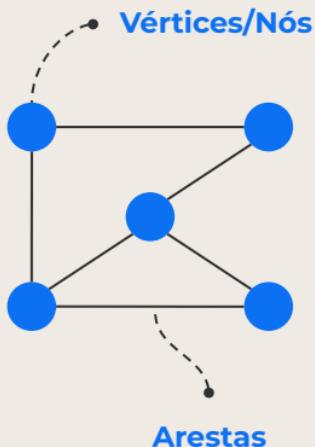
Módulo 04: Collaborative Filtering



Aula 19: Grafos de Itens e Usuários

O que é um grafo?

Abstração matemática que permite representar o relacionamento entre pares de objetos.

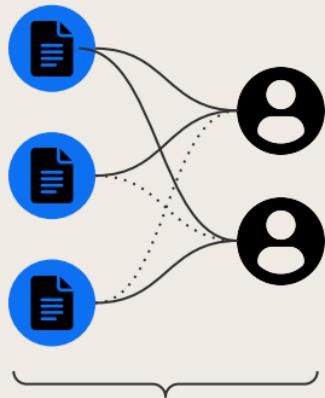


Exemplos de grafos

- Redes Sociais
- Links da web
- Citações em artigos
- Rede Neural

Grafos em sistemas de recomendação

User-Item

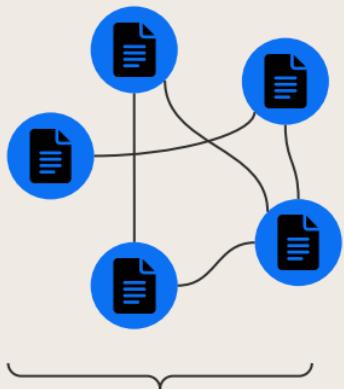


Arestas
Preferência
Pageview
Videoview

- Vértice do tipo Usuário ou Item
- Grafo bipartido
- Predição de arestas

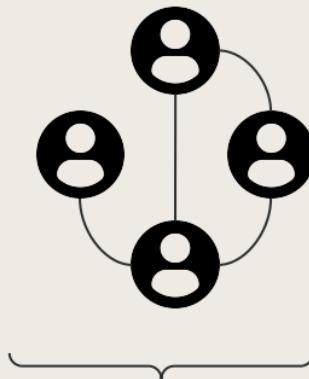
Grafos em sistemas de recomendação

Item-Item



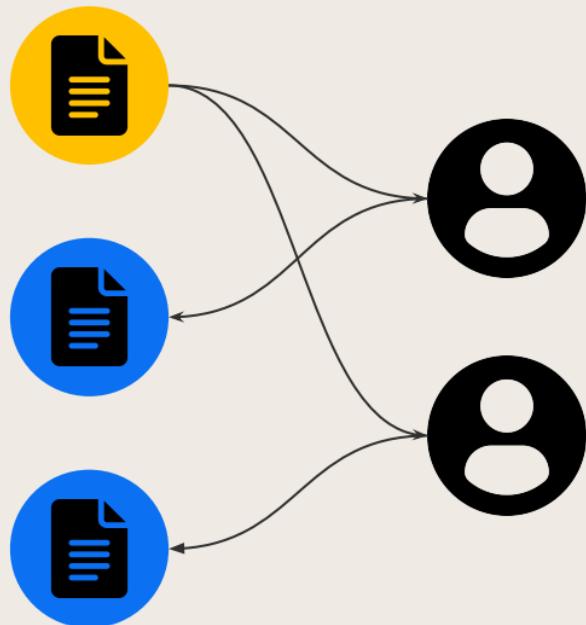
Arestas
Similaridade
Co-visitação

User-User



Arestas
Itens em comum
Friendship

"Quem comprou também comprou..."



Algoritmo Co-visitation

1. Vizinhos de um item-alvo (usuário)
2. Vizinhos do usuário (itens consumidos)
3. Score: quantidade de usuários em comum



Na próxima aula...

Exemplo prático de recomendação por co-visitation



Aula 14:

Exemplo Prático de Recomendação por Co-visitation

Recomendação por Co-visitation





Na próxima aula...

Recomendação com Item KNN



Aula 23:

Recomendação com Nearest Neighbors

Matriz de preferências estimada

$$\hat{r}_{ui} = \mathcal{F}(\mathbf{u}, \mathbf{i} | \phi)$$

$$\mathbf{R} = \begin{Bmatrix} r_{11} & ? & \dots & r_{1m} \\ ? & r_{22} & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & ? & \dots & \hat{r}_{nm} \end{Bmatrix}$$

{ "Target" }

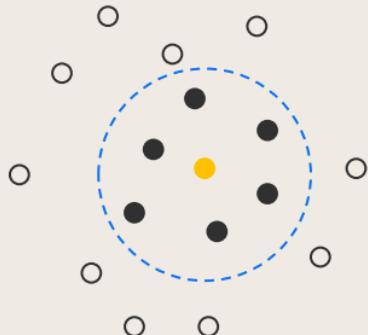
$$\hat{\mathbf{R}} = \begin{Bmatrix} \hat{r}_{11} & \hat{r}_{12} & \dots & \hat{r}_{1m} \\ \hat{r}_{21} & \hat{r}_{22} & \dots & \hat{r}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_{n1} & \hat{r}_{n2} & \dots & \hat{r}_{nm} \end{Bmatrix}$$

{ "Predicted" }

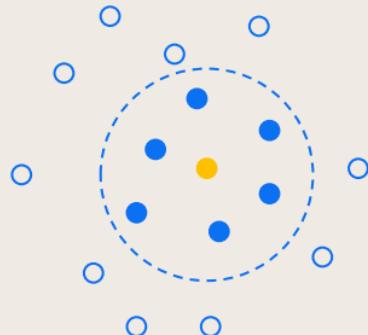
Modelos de *Nearest Neighbors*

Nos modelos baseados em **vizinhança**, procuramos estimar a preferência a partir da **similaridade** entre itens ou entre usuários

K **usuários** mais próximos



K **itens** mais próximos



ItemKNN

Média de *feedbacks*
(bias) do item i

K vizinhos de i que o
usuário u já consumiu

Feedback do usuário u
para o item j

$$\hat{r}_{ui} = b_i + \frac{\sum_{j \in N_u^k(i)} sim(i,j) \cdot (r_{uj} - b_j)}{\sum_{j \in N_u^k(i)} sim(i,j)}$$

Similaridade entre
os itens i e j

UserKNN

Média de *feedbacks* (bias) do usuário u

$$\hat{r}_{ui} = b_u + \frac{\sum_{v \in N_i^k(u)} sim(u,v) \cdot (r_{vi} - b_v)}{\sum_{v \in N_i^k(u)} sim(u,v)}$$

K vizinhos de u que consumiram o item i

Similaridade entre os usuários u e v

Feedback do usuário v para o item i

ItemKNN ou UserKNN ?

A quantidade de usuários e itens pode impactar na eficiência do algoritmo devido ao **tamanho da matriz de similaridades**

$$\mathbf{I} = \left\{ \begin{array}{cccc} i_{11} & i_{12} & \dots & i_{1M} \\ i_{21} & i_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ i_{M1} & i_{M2} & \dots & i_{MM} \end{array} \right\} \quad \mathbf{U} = \left\{ \begin{array}{cccc} u_{11} & u_{12} & \dots & u_{1N} \\ u_{21} & u_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & \dots & u_{NN} \end{array} \right\}$$



Na próxima aula...

Exemplo prático com Nearest Neighbors



Aula 16: Exemplo Prático com Nearest Neighbors

Recomendação com ItemKNN





Na próxima aula...

Recomendação com Fatoração de Matrizes



Aula 27:

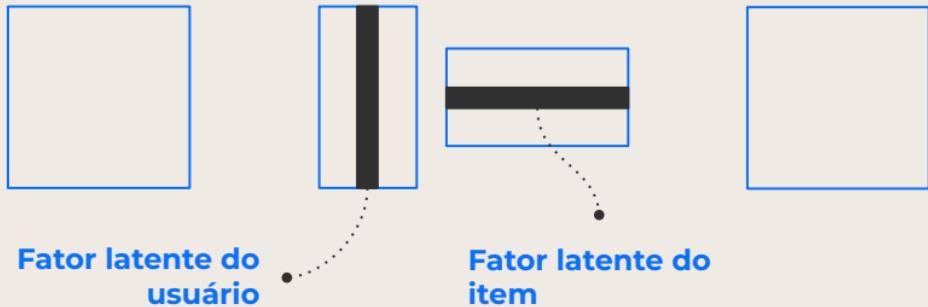
Recomendação com Fatoração de Matrizes

Fatoração de matrizes

Na fatoração de matrizes, procuramos representar a **matriz de preferências** como uma **multiplicação de diversas matrizes** contendo **fatores latentes** dos usuários e itens.

$$\mathbf{R}_{U \times I} \approx \mathbf{P}_{U \times k} \mathbf{Q}_{k \times I} = \hat{\mathbf{R}}_{U \times I}$$

$$\mathbf{R} \quad \approx \quad \mathbf{P} \quad \mathbf{Q} \quad = \quad \hat{\mathbf{R}}$$



$$\mathbf{p}_u = [p_1, p_2, \dots, p_k]^T$$

$$\mathbf{q}_i = [q_1, q_2, \dots, q_k]$$

Função de preferência

$$\mathbf{q}_i = [q_0, q_1, q_2, q_3, \dots, q_k] \quad \mathbf{p}_u = [p_0, p_1, p_2, p_3, \dots, p_k]$$

$\mathcal{F}(\cdot)$

Função de
preferência

$$\hat{r}_{ui} = \mathcal{F}(\mathbf{u}, \mathbf{i} | \phi) = b + b_{\mathbf{u}} + b_{\mathbf{i}} + \mathbf{p}_u^T \mathbf{q}_i$$

Processo de otimização

Encontrar as matrizes P e Q que **minimizam a distância** entre a matriz de preferência original e a estimada (considerando apenas **entradas conhecidas**).

$$\hat{\mathbf{R}} = \mathbf{P}\mathbf{Q}$$


$$\mathbf{P}^*, \mathbf{Q}^* = \min_{P,Q} \|\mathbf{R} - \hat{\mathbf{R}}\|_2 + \lambda_p \|\mathbf{P}\| + \lambda_q \|\mathbf{Q}\|$$

Parâmetros
ótimos

Parâmetros a
serem treinados

Regularização

Algoritmo ALS (Alternating Least Squares)

Ideia: treinar uma matriz de cada vez (de forma **alternada**)

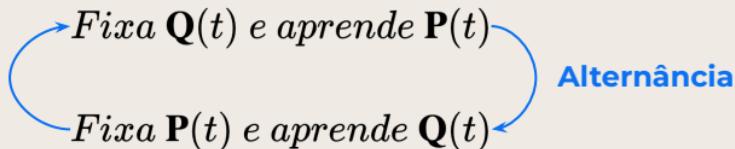
$$\mathbf{P}^*, \mathbf{Q}^* = \min_{P,Q} \|\mathbf{R} - \hat{\mathbf{R}}\|_2 + \lambda_p \|\mathbf{P}\| + \lambda_q \|\mathbf{Q}\|$$

Não há solução analítica se P
e Q forem variáveis
ao mesmo tempo

1. Inicializar P e Q aleatoriamente

$$\mathbf{P}(t=0) = \text{random}(U, k) \quad \mathbf{Q}(t=0) = \text{random}(k, I)$$

2. Loop até convergência



Algoritmo FunkSVD

Ideia: similar ao ALS, porém incrementando o número de fatores para ser mais **rápido**.

• Solução do Netflix Prize Challenge

$$P \quad Q \quad = \quad \hat{R}$$



Resolve ALS
para **$k = 1$**



Resolve ALS
para **$k = 2$**

⋮
⋮
⋮



Resolve ALS
para **$k = N$**



Na próxima aula...

Exemplo prático com fatoração de matrizes



Aula 17:

Exemplo Prático com Fatoração de Matrizes

Recomendação por FunkSVD





Obrigado.

Módulo 05:

Métricas de

Avaliação



Aula 30: Métricas de Acurácia

Documentação de métricas

Referência

Evaluating Recommendation Systems

Evaluating Recommendation Systems

Guy Shani and Asela Gunawardana

Abstract Recommender systems are now popular both commercially and in the research community, where many approaches have been suggested for providing recommendations. In many cases a system designer that wishes to employ a recommendation system must choose between a set of candidate approaches. A first step towards selecting an appropriate algorithm is to decide which properties of the application to focus upon when making this choice. Indeed, recommendation systems have a variety of properties that may affect user experience, such as accuracy, robustness, and speed. In this paper we review how to evaluate recommendation systems based on a set of properties that are relevant for the application. We focus on comparative studies, where a few algorithms are compared using some evaluation metric, rather than absolute benchmarking of algorithms. We describe experimental settings appropriate for making choices between algorithms. We review these types of experiments, starting with an offline setting, where recommendation approaches are compared without user interaction, then reviewing user studies, where a small group of subjects experiment with the system and report on the experience, and finally describe large scale online experiments, where real user populations interact with the system. In each of these cases we describe types of questions that can be answered, and suggest protocols for experimentation. We also discuss how to draw trustworthy conclusions from the conducted experiments. We then review a large set of properties, and explain how to evaluate systems given relevant properties. We also survey a large set of evaluation metrics in the context of the property that they evaluate.

Relembrando a matriz de preferências estimada

$$\hat{r}_{ui} = \mathcal{F}(\mathbf{u}, \mathbf{i} | \phi)$$

$$\mathbf{R} = \left\{ \begin{array}{cccc} r_{11} & ? & \dots & r_{1m} \\ ? & r_{22} & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & ? & \dots & \hat{r}_{nm} \end{array} \right\}$$



"Target"

$$\hat{\mathbf{R}} = \left\{ \begin{array}{cccc} \hat{r}_{11} & \hat{r}_{12} & \dots & \hat{r}_{1m} \\ \hat{r}_{21} & \hat{r}_{22} & \dots & \hat{r}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_{n1} & \hat{r}_{n2} & \dots & \hat{r}_{nm} \end{array} \right\}$$



"Predicted"

Acurácia de feedback

$$e_{ui} = r_{ui} - \hat{r}_{ui}$$

$\mathcal{K} \rightarrow$ Conjunto de pares user-item conhecidos

Mean Absolute Error

$$MAE = \frac{1}{|\mathcal{K}|} \sum_{(u,i) \in \mathcal{K}} |e_{ui}|$$

Root Mean Squared Error

$$RMSE = \sqrt{\frac{1}{|\mathcal{K}|} \sum_{(u,i) \in \mathcal{K}} e_{ui}^2}$$

Problemas

- Não considera o **tamanho** da lista recomendada
- Não considera o **ranking** do item na lista
- Depende de **feedback explícito**

Acurácia de classificação

Considerando uma lista recomendada de K itens:

	RECOMENDADO	NÃO-RECOMENDADO
CONSUMIU	True-positive (TP)	False-negative (FN)
NÃO CONSUMIU	False-positive (FP)	True-negative (TN)

Precision@k

$$Precision@K = \frac{\#TP}{\#TP + \#FP}$$

Recall@k

$$Recall@K = \frac{\#TP}{\#TP + \#FN}$$

Problemas

- Não considera o ranking do item na lista
- Não são deriváveis

Acurácia de classificação - Exemplo

Catálogo					
Consumido					
Recomendado					

- $\#TP = 1$ -----> Recomendou e consumiu
- $\#FP = 2$ -----> Recomendou mas não consumiu
- $\#FN = 1$ -----> Consumiu mas não recomendou

$$Precision@4 = \frac{\#TP}{\#TP + \#FP} = \frac{1}{1+2}$$

$$Recall@4 = \frac{\#TP}{\#TP + \#FN} = \frac{1}{1+1}$$

Acurácia de ranqueamento

$rank_u \rightarrow$ Ranking do item consumido pelo usuário u $\mathcal{U} \rightarrow$ Conjunto de usuários que receberam a recomendação

MRR@k

$$MRR@K = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{rank_u}$$

Problemas

-

Não são deriváveis

Acurácia de ranqueamento - exemplo

Recomendado				
Consumido (User 1)				
Consumido (User 2)				

$$MRR@4 = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{rank_u} = \frac{1}{2} \left(\frac{1}{1} + \frac{1}{4} \right) = \frac{5}{8}$$



Na próxima aula...

Métricas de negócio



Aula 31:

Métricas de Negócio

Item Space Coverage



Item Space Coverage

$$ISC@K = \frac{|\cup_{u \in \mathcal{U}} \hat{\mathcal{R}}_u|}{|\mathcal{I}|}$$

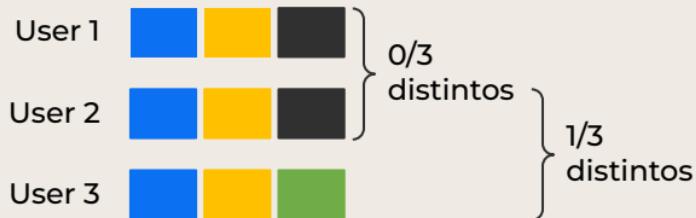
"Fração dos itens do catálogo que apareceram em alguma recomendação"



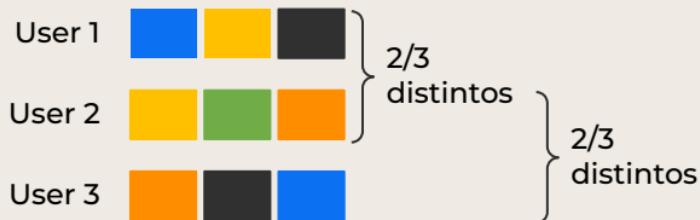
$$ISC@4 = \frac{3}{5}$$

Personalização

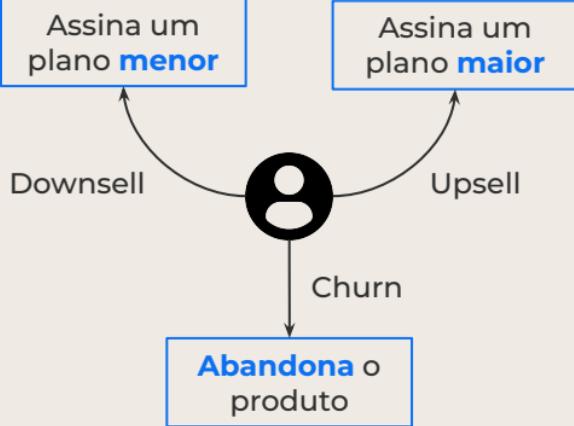
Baixa Personalização



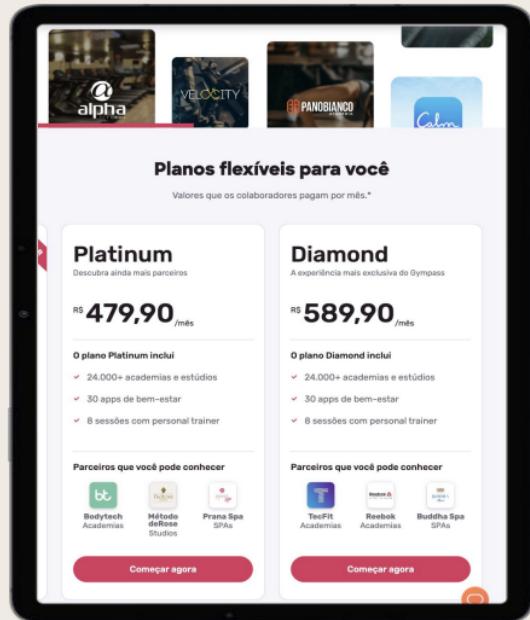
Alta Personalização



Upsell, downsell e churn



Next Best Offer (NBO)



Playtime

Sessão do usuário



Característico em
produtos de vídeo

- Playtime por sessão
- Playtime semanal
- Playtime mensal

Time to action

Sessão do usuário



Tempo dentro de uma **sessão** para um usuário realizar uma **ação de interesse**

- Time to play
- Time to purchase

Ticket Médio

Sessão do usuário



Valor médio de **vendas** em um determinado período

- Ticket médio por sessão
- Ticket médio semanal
- Ticket médio mensal



Na próxima aula...

Exemplo prático de avaliação da recomendação



Aula 20: Exemplo Prático de Avaliação da Recomendação

Métricas de Acurácia e Cobertura





Obrigado.

Módulo 06: Desafios de Sistemas de Recomendação em Produção



Aula 34: Design de Interfaces

Recomendação vs. User Interface

Como a **interface do usuário** impacta na qualidade da **recomendação**?

Interface antiga da Amazon.com

The screenshot shows the homepage of Amazon.com from August 27, 1999. The top navigation bar includes links for WELCOME, BOOKS, MUSIC, VIDEO, TOYS & GAMES, ELECTRONICS, e-CARDS, and AUCTIONS. Below the navigation is a search bar with the placeholder "All Products" and a "GO" button. To the right of the search is a banner for "In Books London Calling" featuring a book cover for "The Golden Child". A message below the banner states: "Hello! Shopping at Amazon.com is 100% secure--guaranteed. Already a customer? [Sign in.](#)". On the left, there's a sidebar titled "BROWSE" with categories like Books, Music, Video, Auctions, and Electronics. The main content area has several sections: "In Books London Calling", "In Electronics MiniDisc the Magnificent", and "Amazon.com 100 Hot Books Updated Hourly" which lists three Harry Potter books by J.K. Rowling. The footer displays the date "Friday, August 27, 1999".

Posição da oferta - viewport



Nome da oferta

Explicabilidade



Novidades



Layout da oferta

Imagens e thumbnails

Shorts

Thumbnail	Title	Subtitles	View Count	Age
	Jordan Peterson Will Leave You...	to learn to play	1,2 mi de visualizações	• há 1 mês
	RONALDO CUTUCA IMPRENSA E...	RONALDO FALA DO APOIO DE NEYMAR A BOLSONARO	2 mi de visualizações	• há 1 mês
	She is mastering her skills in desserts...	CILLIAN MURPHY	8,1 mi de visualizações	• há 1 mês
	RONALDO leva bronca de ROMÁRIO...	mas quem decide	4,5 mi de visualizações	• há 1 mês
	just Jordan Peterson ANNIHILATING...	JORDAN PETERSON MAKES FEMINIST BOW HER HEAD...	3,6 mi de visualizações	• há 1 mês

The Big Bang Theory Season 6 | Bloopers vs Actual Scene
8:17

Tetiana Ivanova: How to become a Data Scientist in ...
PyData
56:25

ELE FAZ A ALEGRIA DO CARDIOLOGISTA!
15:38

CASIMIRO REAGE: 7 CAMADAS DE CARNE! X...
Cortes do Casimiro [OFICIAL]
614 mil visualizações • há 1 mês

ABU DHABI | FORMULA 1 | QUALIFYING HIGHLIGHTS
5:57

The Coopers
6,3 mi de visualizações • há 1 anno

PyData
259 mil visualizações • há 6 anos

Cortes do Casimiro [OFICIAL]
614 mil visualizações • há 1 mês

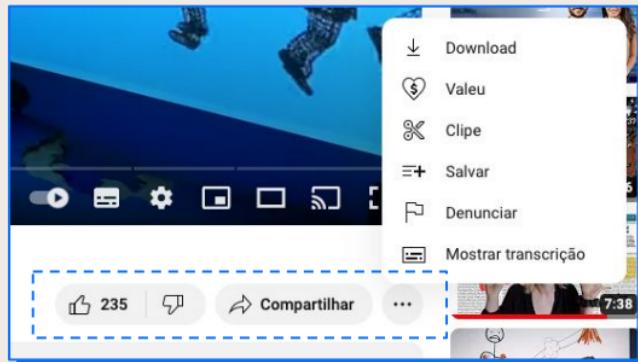
Abu Dhabi Grand Prix
FORMULA 1
3,3 mi de visualizações • há 8 dias

Metadados dos itens

Mecanismos de feedback



A screenshot of a Twitter thread. The first tweet is from Yann LeCun (@ylecun) on Nov 26, stating: "Four interesting presentations at NeurIPS co-authored by @randall_balestr and me with various sets of coauthors." The second tweet is from Randall Balestrero (@randall_balestr) on Nov 26, stating: "Interested in (A) self-supervised learning <-> spectral embedding (B) data-augmentation's unfair per-class impact (C) DA implicit/explicit regularizer (D) fast orthogonal/unitary weight learning? Come see us #NeurIPS2022 4:30/Hall J". Below the tweets are engagement metrics: 1 like, 4 retweets, 53 favorites, and a share icon.





Na próxima aula...

Experimentação em sistemas de recomendação



Aula 35: Experimentação em Sistemas de Recomendação

Fatores que influenciam na qualidade da recomendação



Algoritmos

Modelos
Hiperparâmetros
Seleção de Acervo
Métricas de Avaliação



User Interface

Posicionamento
Layout
Nome
Feedback



Sazonalidade

BlackFriday
Natal
Copa do Mundo
Lançamentos

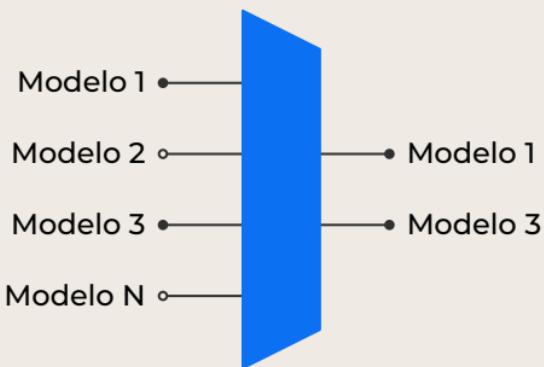
Experimentos

A **cultura de experimentação** deve estar fortemente presente na construção de um **sistema de recomendação** devido às **incertezas do ambiente**



Experimentos offline

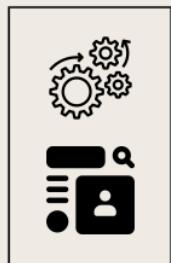
Foco na realização de muitos experimentos com **dados passados** para **selecionar modelos** promissores



- Baixo custo: **sem modificações de código** em produção
- Baixo risco: **sem interações** com usuários em tempo real
- Altamente dependente da **amostra de dados**

Casos de uso

Foco na realização de poucos experimentos com
usuários selecionados para **validar modelos**



- Alto custo: entrevistas, agendamento
- Baixo risco: poucos usuários envolvidos
- *Feedbacks qualitativos* dos usuários

Experimentos online

Foco na realização de poucos experimentos com **dados reais** para validar modelos em **produção**

CONTROLE



ALTERNATIVA

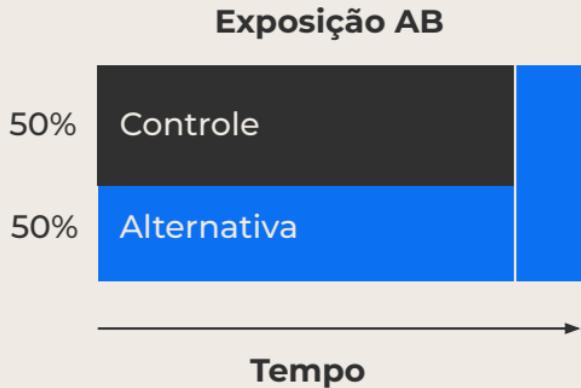


- Alto custo: modificação de código em produção
- Alto risco: **muitos usuários** envolvidos
- Maior **fidelidade** ao consumo do usuário

Recomendação não-personalizada,
"no-rec", algoritmos simples

Multi armed bandits (MAB)

MAB é uma técnica baseada em **aprendizado por reforço** onde a exposição de uma alternativa varia de acordo com a sua métrica





Na próxima aula...

Arquitetura de um sistema de recomendação



Aula 36: Arquitetura de um Sistema de Recomendação

Como juntar os diferentes componentes de um sistema de recomendação?



Arquitetura end-to-end



Client



Client

Tudo começa com
o **usuário!**

Definições da oferta

- Qual problema de negócio será trabalhado?
 - Churn
 - Engajamento
 - Upsell
- Onde estará a oferta?
 - Home
 - Página de consumo
 - Fim do vídeo
- Qual é a seleção de acervo?
- Qual é o tipo da oferta?
 - Top
 - Item-based
 - User-based

Application Programming Interface

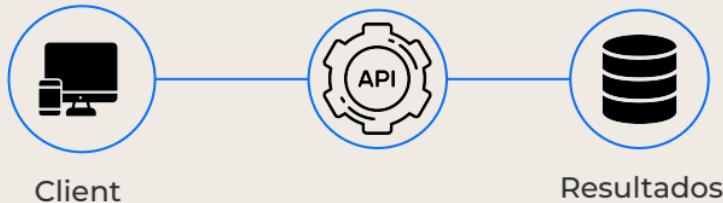


Client

Requisitando as
recomendações

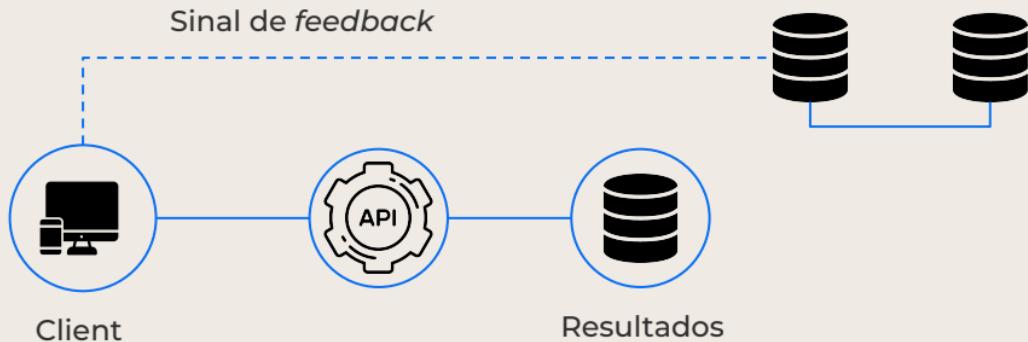
- Como será o **contrato** da API?
 - Headers
 - Cookies
 - Response
- Qual é o **tempo de resposta** máximo?
 - 300ms
 - 500ms
- O que fazer se a **API não responder**?
 - Apagar a oferta
 - Buscar um *fallback*
- Como entregar diferentes **alternativas**?
 - AB
 - MAB

Armazenamento de resultados



- **Batch** ou **Streaming**?
- Arquitetura do banco
 - Relacional
 - No-sql
 - In-memory
- O banco terá **replicação**?

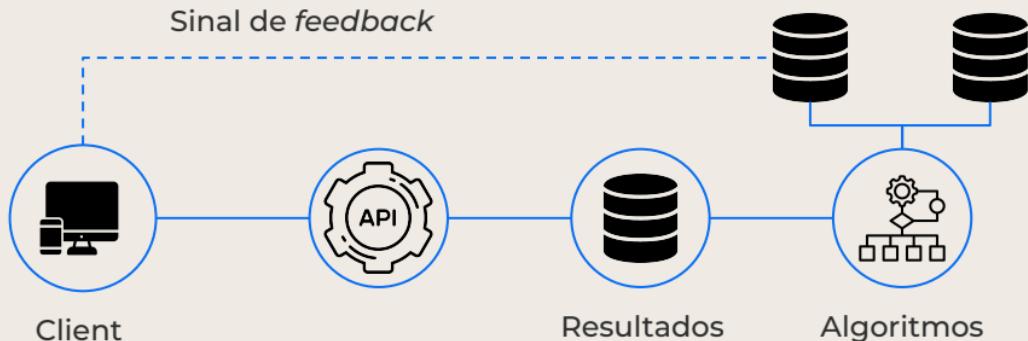
Metadados de itens e usuários



A **implementação de sinais** é fundamental para uma boa recomendação

- Sinais de **feedback**
 - Impressão
 - Click
 - Consumo
 - Compartilhamento
- Dados **Near Real Time**?
- Como **enriquecer** os metadados dos itens?

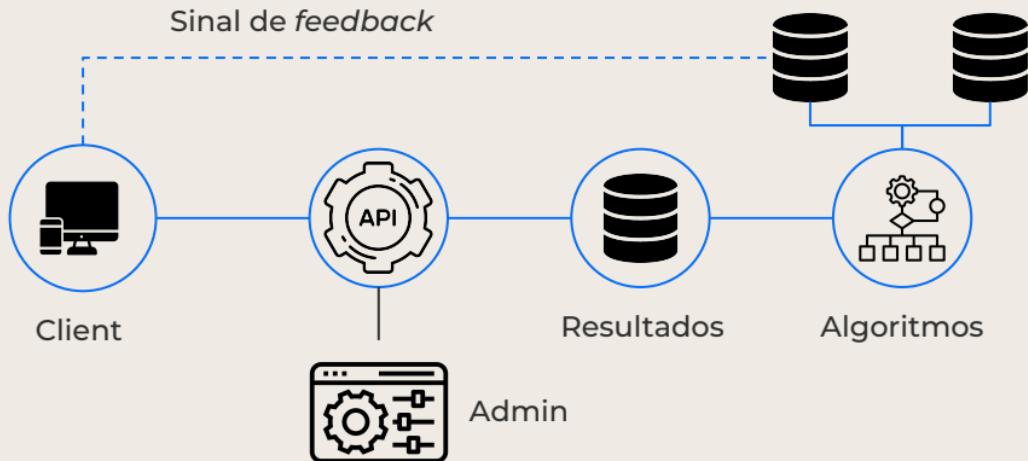
Estratégias de algoritmos



Diferentes algoritmos devem ser testados constantemente na cultura de experimentação

- Quais são as alternativas de algoritmos?
- Onde será executado?
- Como monitorar a execução dos jobs? (MLOps)

Administração das ofertas



Configuração e monitoramento de todas as ofertas recomendadas

- Quais ofertas estão **ativas**?
- Quais algoritmos estão em cada oferta?
- Quem tem **permissão** para configurar uma oferta?



Obrigado.