

Sistemas de Recomendação

Elaine C. R. Cândido

Pontifícia Universidade Católica de Minas Gerais

Março 2021

Programa do Curso

1 Introdução

- O que é?
- Definições do problema: predição e ranqueamento
- Tipos de sistemas de recomendação

2 Fundamentos matemáticos

- Métricas de similaridade
- Fatorização de matrizes

3 Praticando Sistemas de Recomendação

- Simples
- Baseado em conteúdo (Content-based)
- Baseado em filtragem colaborativa (Collaborative Filters)

4 Learning to rank

- Tipos de algoritmos

Pré-requisitos

- Lógica de programação
- Programação em Python
- Nível do curso: intermediário
- Conhecimento em aprendizado de máquina

Visão Geral

1 Introdução

- Filtragem Colaborativa
- Tipos de Sistemas de Recomendação
- Content-based - Baseado em Conteúdo
- Knowledge-based - Baseado em Conhecimento

2 Prática

- SR simples
- Knowledge-based
- Content-based
- Filtragem Colaborativa

3 Learning to Rank (LTR)

- Livros de sistemas de recomendação de Kim Falk e Rounak Banik

Introdução

- Sistemas de recomendação estão por todos os lugares...
- O que ver...



Introdução

- O que comprar junto...

Frequently bought together



+



+

Total price: **\$25.78**[Add all three to Cart](#)[Add all three to List](#)

- ✓ **This item:** AmazonBasics Pre-sharpened Wood Cased #2 HB Pencils, 150 Pack **\$12.49**
- ✓ Elmer's All Purpose School Glue Sticks, Washable, 7 Gram, 30 Count **\$9.00**
- ✓ Paper Mate Pink Pearl Erasers, Large, 12 Count **\$4.29**

Introdução

- O que ouvir...

MADE FOR SOPHIA

Discover Weekly

Your weekly mixtape of fresh music. Enjoy new discoveries and deep cuts chosen just for you. Updated every Monday, so save your favourites!

Made for Sophia Ciocca by Spotify · 30 songs, 2 hr 3 min

PLAY **FOLLOWING** ...

FOLLOWER 1








Filter Download

TITLE	ARTIST	ALBUM	
To Hugo	Clogs	The Crestures In Th...	2 days ago
Little Worlds	Mandolin Orange	Such Jubilee	2 days ago
Quiet Voices	Mike Vass	In the Wake of Neil ...	2 days ago
Sometimes	Goldmund	Sometimes	2 days ago
Sileo	Rhian Sheehan	Stories From Elsewh...	2 days ago
Hollow Home Rd	Brolly	Hollow Home Rd	2 days ago
Marigold	Mother Falcon	You Knew	2 days ago
Things Happen	Dawes	All Your Favorite Ba...	2 days ago
Sliding Down	Edgar Meyer, Mike ...	The Best of Edgar M...	2 days ago
Celeste	Pete Kuzma	Equilibrium	2 days ago

3:43 5:49

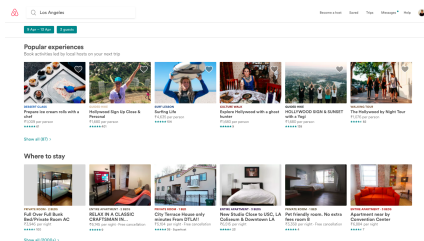
Predição

- Matriz esparsa: todos os usuários terem visto/ouvido/etc apenas uma pequena fração de todos os itens
- Objetivo: prever os valores faltantes da matriz com informação que se tem

	i_1	i_2	i_3	i_4	i_5	i_6
 U1	4	?	3	?	5	?
 U2	?	2	?	?	4	1
 U3	?	?	1	?	2	5
 U4	?	?	3	?	?	1
 U5	1	4	?	?	2	5
 U6	5	?	2	1	?	4
 U7	?	2	3	?	4	5

Ranqueamento

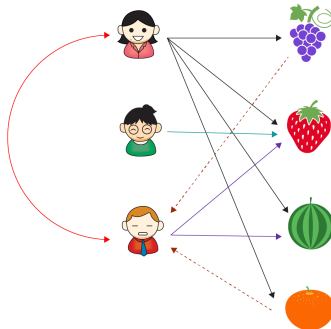
- Tenta ranquear os top k itens para recomendar a um usuário
- Mostrar os 10 resultados que satisfazem busca de usuário



- O problema de predição sempre chega ao de ranqueamento. Se os valores faltantes são preditos, podemos extrair os top valores e mostrar os resultados

SR baseado em Filtragem Colaborativa

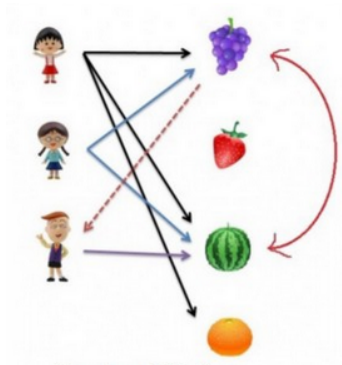
- Baseado em usuário
 - Usuários que compram ou gostaram dos mesmos itens no passado são mais prováveis de gostar das mesmas coisas no futuro também



SR baseado em Filtragem Colaborativa

- Baseado em item

- Se um grupo de usuários avaliou dois itens de forma similar, então estes dois itens devem ser similares. Assim, se uma pessoa gosta de um item em particular, é provável que ela esteja interessada no outro item também



SR baseado em Filtragem Colaborativa

- Feedback explícito - usuário mostra suas preferências através de likes, listas de favoritos, etc.
- Feedback implícito - baseia-se no comportamento do usuário na plataforma, cliques, visitas a itens, cookies, etc.
- Problemas:
 - *Cold start* - Histórico de dados
 - Escalabilidade - Calcular métricas de similaridade pode ser tornar uma tarefa muito custosa (tempo e recursos)
 - Esparsidade - Baixa porcentagem de feedbacks (muitos itens poucos comprados/vistos/ouvidos)

SR baseado em Conteúdo

- Não requerem dados de histórico
 - Baseado no usuário e seus metadados
 - Geralmente, o sistema pede ao usuário para classificar itens de acordo com seu gosto
 - Baseado nessa informação e nos metadados dos itens o sistema cria uma recomendação
- Problemas:
 - Recomendações óbvias

SR baseado em Conhecimento

- Usado para itens que são raramente comprados
 - Imobiliárias, por exemplo
 - Não há histórico

Prática de Sistemas de Recomendação

- Sistema de Recomendação simples
- Baseado em conteúdo
- Filtragem colaborativa
- Filtragem colaborativa com redes neurais

IMDB Top 250 - Simples

- 1 Escolher uma métrica para classificar
 - 2 Decidir pré-requisitos para mostrar filmes
 - 3 Calcula score para cada filme que satisfaz condições
 - 4 Retornar lista final
- Métrica

$$\textit{Weighted Rating (WR)} = \left(\frac{v}{v+m} \times R \right) + \left(\frac{m}{v+m} \times C \right)$$

IMDB Top 250 - Baseado em Conhecimento

- 1 Pergunta ao usuário suas preferências
- 2 Recomenda itens com alto score e selecionando preferências do usuário

SR Baseado em Conteúdo

- 1 Baseado na descrição dos filmes
- 2 Baseado nos metadados dos filmes

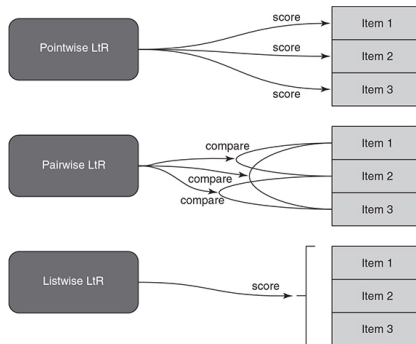
SR Filtragem Colaborativa

- Movielens dataset
- Baseado em usuário
- Baseado em itens

LTR

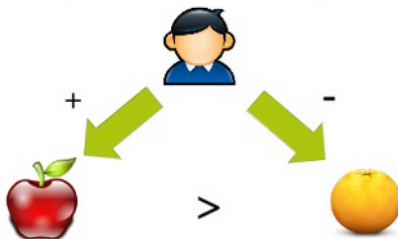
- Pointwise: simples ordenação dos scores.
- Pairwise: dado um par de resultados/queries otimiza a ordenação do par e compara com o ground truth.
 - Meta: minimizar casos onde o par de resultados estão na ordenação errada com relação ao ground truth
 - Exemplos: RankNet, LambdaRank e LambdaMART
- Listwise: observa a lista de resultados completa e otimiza a ordenação
 - SoftRank, AdaRank

LTR

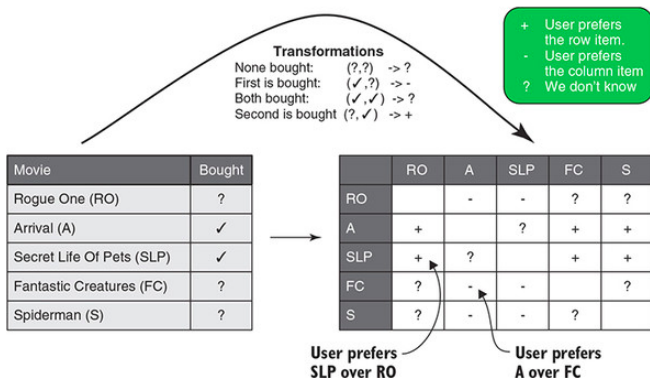


LTR

- BPR - Bayesian Personalized Ranking: Aprende pela comparação de itens que usuário gosta com itens que ele não gosta



LTR



Referências

- Livros

- Banik, Rounak. Hands-On Recommendation Systems with Python: Start building powerful and personalized, recommendation engines with Python. Packt Publishing Ltd, 2018.
- Falk, Kim. Practical recommender systems. Manning Publications, 2019.
- Yalçın, Orhan Gazi. Applied Neural Networks with TensorFlow 2.

- Códigos:

- <https://github.com/PacktPublishing/Hands-On-Recommendation-Systems-with-Python>

Frame Title

• Artigos:

- Li, Hang. "A short introduction to learning to rank." IEICE TRANSACTIONS on Information and Systems 94, no. 10 (2011): 1854-1862.
- Saikishore Kalloori, Francesco Ricci, and Marko Tkalcić. 2016. Pairwise Preferences Based Matrix Factorization and Nearest Neighbor Recommendation Techniques. In Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16). Association for Computing Machinery, New York, NY, USA, 143–146. DOI:<https://doi.org/10.1145/2959100.2959142>

• Vídeos:

- RecSys 2016: Paper Session 4 - Pairwise Preferences Based Matrix Factorization. <https://www.youtube.com/watch?v=EajR97VK8hI>
- RecSys 2016: Paper Session 11 - Bayesian Personalized Ranking with Multi-Channel User Feedback. <https://www.youtube.com/watch?v=aKHLf4P3N08>