



# Sistemas de Recomendação

No domínio da música



UNIVALI



VNIVERSIDAD  
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL



**ESJALB**

Expert Systems and Applications Laboratory

# Sobre mim

## Álvaro Lozano Murciego

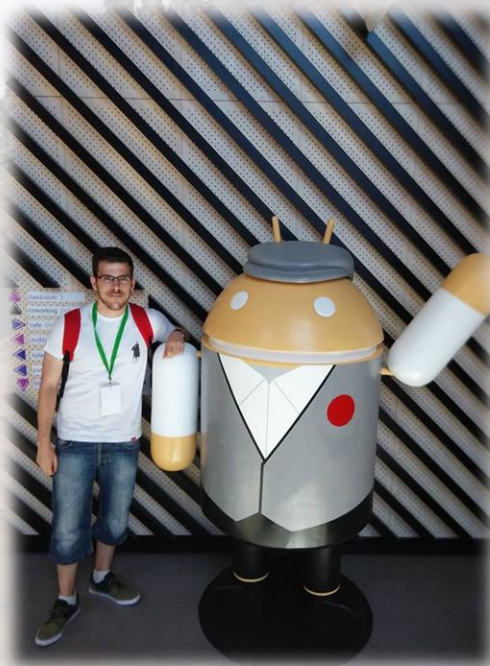
Professor Associado  
Departamento de Informática e Automatização  
Área: Linguagens e sistemas de computação  
Universidade de Salamanca  
Espanha



Grupo de investigación  
Universidad de Salamanca

Linhas de pesquisa e áreas de interesse:

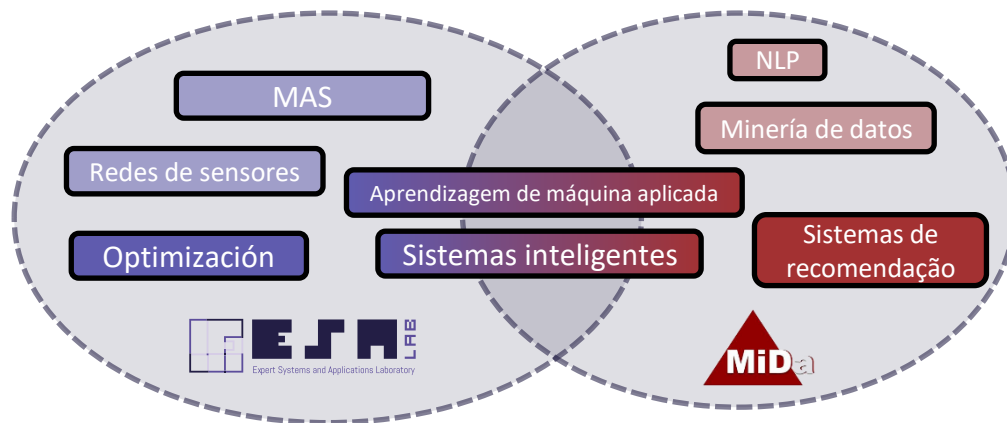
- ✓ *Aprendizagem da máquina*
- ✓ *Sistemas de recomendação*
- ✓ *Computação de bordas*
- ✓ *Desenvolvimento Android*
- ✓ *IoT, sensores e MCU*



# Research interest



Grupo de investigación  
Universidad de Salamanca



## Conteúdo

1. Introdução
2. Tipos de sistemas de recomendação clássicos
3. Sistemas de recomendação sensível ao contexto (CARS)
4. Características do domínio da música
5. Trabalho recente
6. Por onde eu começo? Recursos!

# Sistemas de recomendação

Uma breve introdução

# O que é um sistema de recomendação (RecSys)?

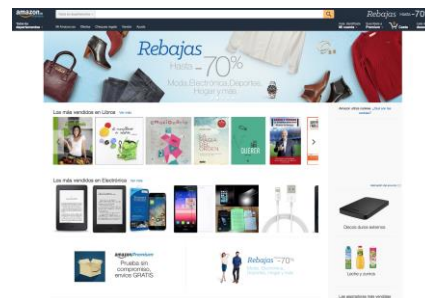
## Definição

Os sistemas recomendados são ferramentas que ajudam um usuário (ou grupo de usuários) a interagir com um sistema para escolher itens de um vasto catálogo ou espaço de informação (MacNee et. al 2006).

Robin Burke, acrescenta: Procure fornecer itens que **sejam relevantes para o usuário de forma personalizada\***.



> 180 M  
Amazon.es



## Por que o RecSys é útil?

### ❑ Por que muitos Big Tech Comanies os usam?

- ▶ Com a explosão da internet e da web nas últimas décadas, há muitos casos em que os usuários têm de tomar uma decisão diante de um enorme catálogo: YouTube, Spotify, Amazon, etc.
- ▶ Melhorar a **retenção de usuários**.
- ▶ É uma maneira simples **de analisar o mercado**.
- ▶ **Aumento das vendas:**
  - 35% of Amazon.com, 75% Netflix
- ▶ Melhorar a **experiência do usuário (UX)**.





# Onde posso encontrá-los?

## Amazon

### Recomendado en función de tus tendencias de compras

### Cómpralo con

1 Estos productos los envían y venden distintos vendedores. Mostrar detalles

- Este producto: MAYZERO Zapatos de tenis para niños, para correr, deportivos, transpirables, ligeros, para caminar, t... \$557.46
- Greenco Mini contenedores de almacenamiento de alimentos, condimentos y salsas, contenedores de alimentos par... \$230.99

### Los clientes que vieron este producto también vieron

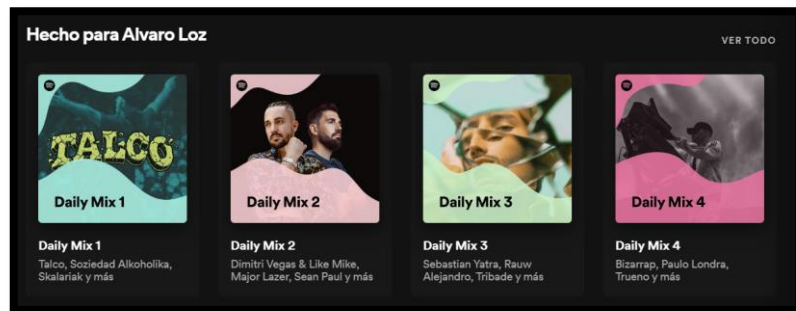
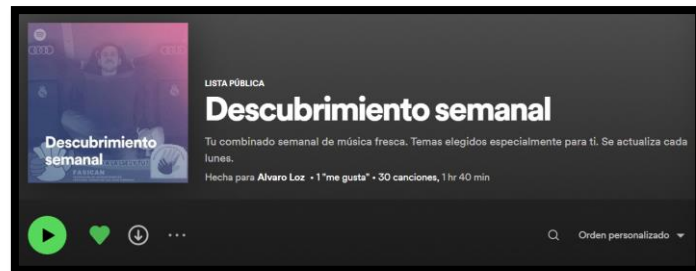
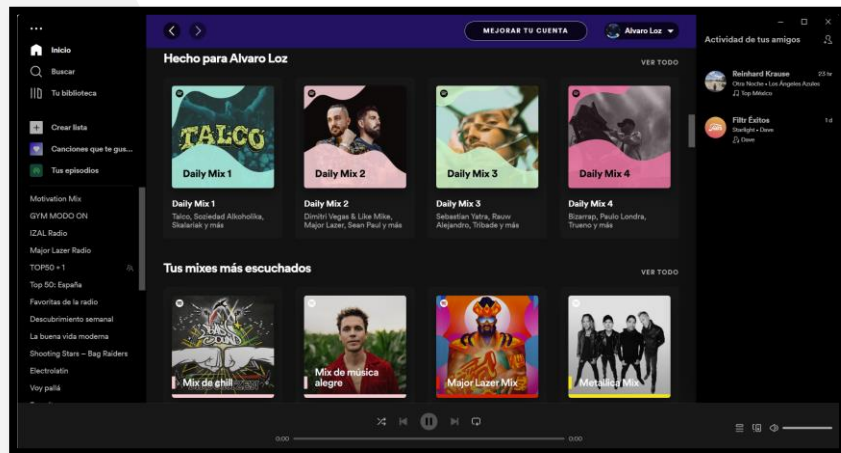
### Artículos destacados que te pueden gustar

### Relacionado con los artículos que viste



# Onde posso encontrá-los?

## □ Spotify



# Onde posso encontrá-los?

## □ Netflix

**Everything is a Recommendation**

The diagram illustrates the Netflix recommendation system. It features a screenshot of the Netflix homepage with various rows of content. A horizontal blue double-headed arrow labeled 'Title Ranking' spans the width of the content rows. A vertical blue double-headed arrow labeled 'Row Selection & Ordering' spans the height of the content rows. A blue arrow labeled 'Image' points to a specific content tile. The content rows include 'Continue Watching for Family' with titles like NARCOS, Super Hero Squad, and JONESTOWN; 'My List' with titles like STRANGER THINGS, THE DARK KNIGHT TRILOGY, and CARTEL LAND; 'NETFLIX ORIGINALS' with titles like NETFLIX ORIGINAL, HOUSE of CARDS, KUFFY ROCK, GO!, and GUMBALL; and 'Trending Now' with titles like NARCOS, SUPER HERO SQUAD, and JONESTOWN.

Recommendations are driven by machine learning algorithms

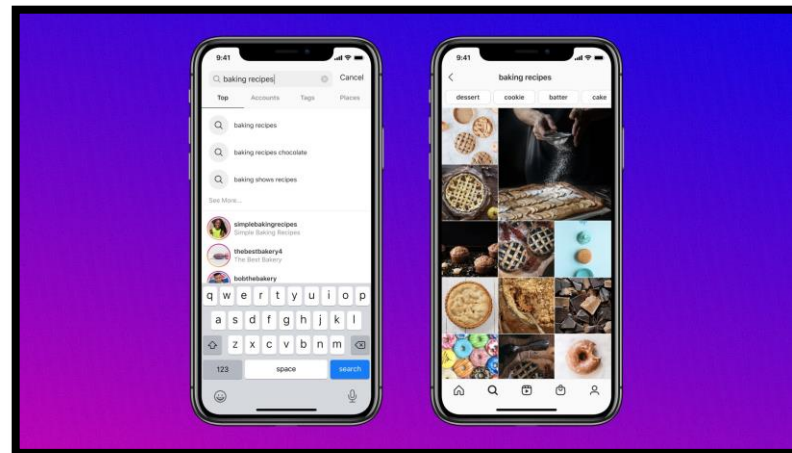
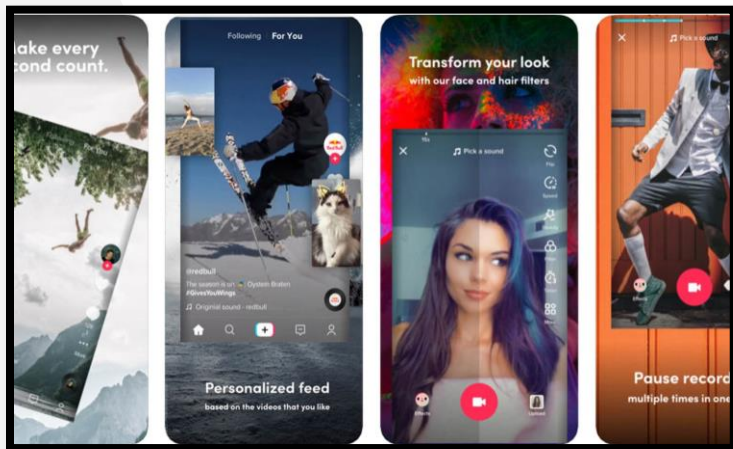
**Over 80%** of what members watch comes from our recommendations

7 **N**

[Fuente: Balancing Discovery and Continuation in Recommendations Hossein Taghavi](#)

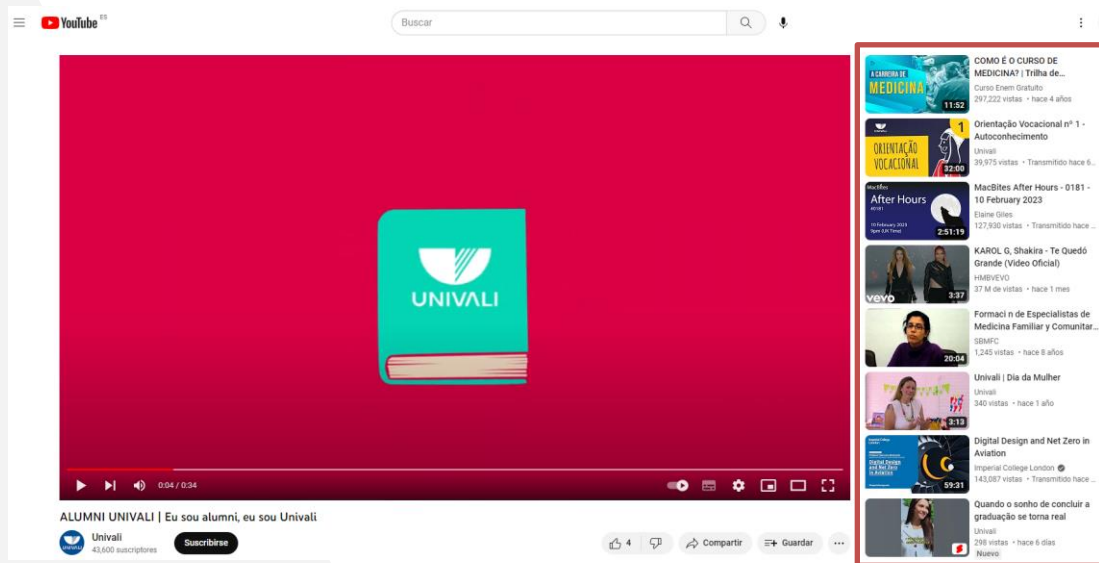
# Onde posso encontrá-los?

- ❑ TikTok e Instagram: *Feed* personalizado



# Onde posso encontrá-los?

## ▣ Youtube



Recomendações

# Onde posso encontrá-los?

## □ Twitter?



Our “algorithm” is overly complex & not fully understood internally. People will discover many silly things , but we’ll patch issues as soon as they’re found!

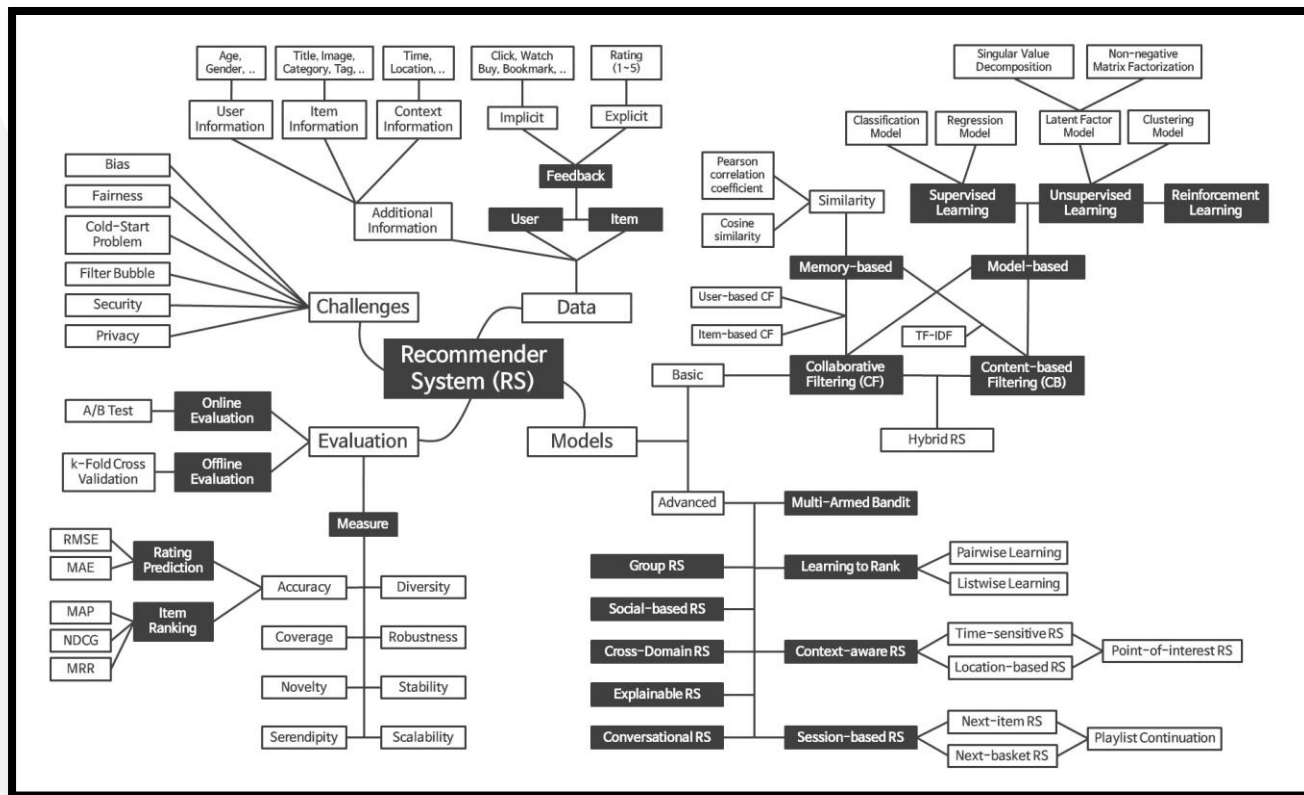
We’re developing a simplified approach to serve more compelling tweets, but it’s still a work in progress. That’ll also be open source.

Providing code transparency will be incredibly embarrassing at first, but it should lead to rapid improvement in recommendation quality. Most importantly, we hope to earn your trust.

[Traducir Tweet](#)

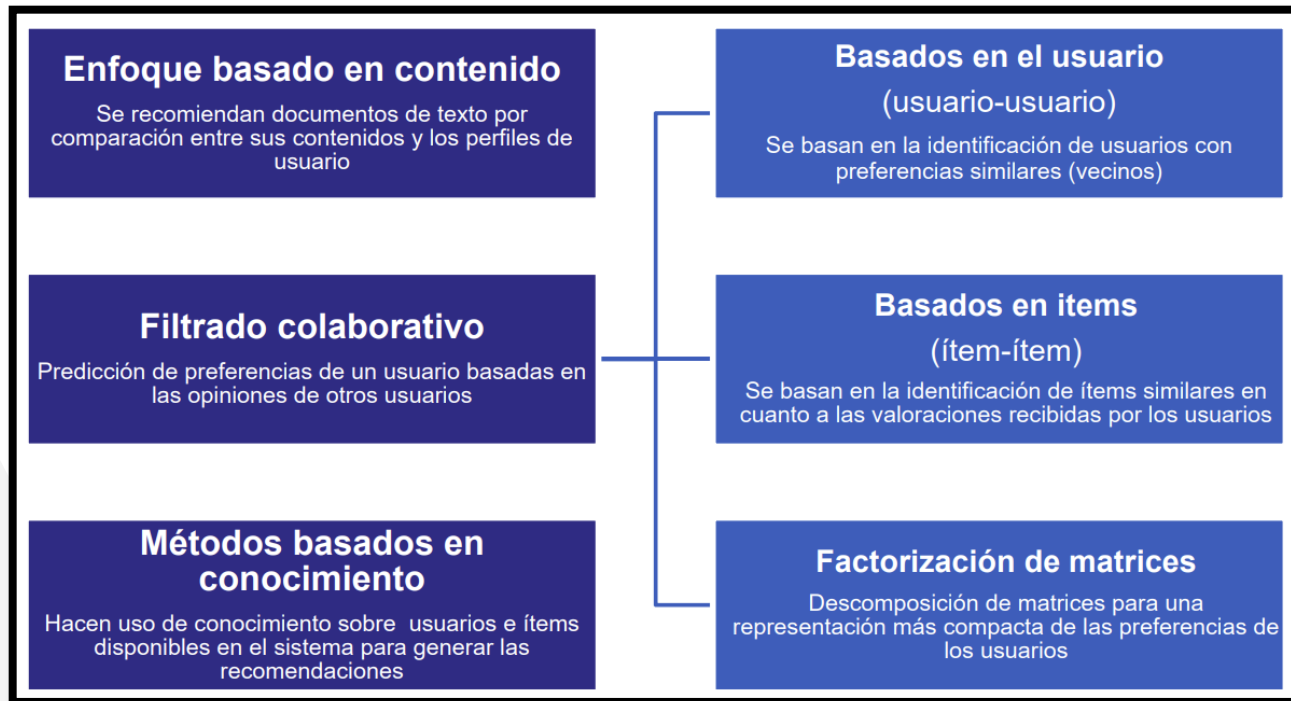
10:16 p. m. · 17 mar. 2023 · **9,8 M** Reproducciones

# Mapa da área de sistemas de recomendação



# Tipos de sistemas de recomendação (classificações)

CONSIDERANDO  
OS DADOS  
UTILIZADOS





# Tipos de sistemas de recomendação (classificações)

## CONSIDERANDO O MODELO

### Métodos basado en memoria

Los cálculos para realizar las recomendaciones se realizan en tiempo de recomendación (*on-line*)

Algunos métodos basados en contenido

Métodos de filtrado colaborativo basados en vecindad (k-NN)

### Métodos basados en modelos

Generan un modelo con anterioridad al momento de la recomendación (*off-line*) en el que se basan las recomendaciones que se proporcionan *on-line*

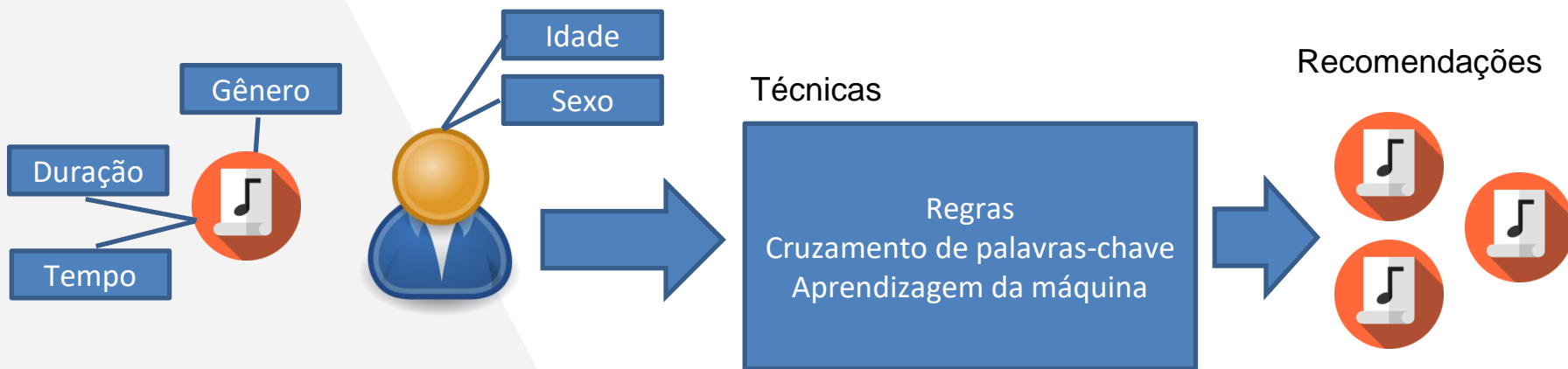
Métodos de filtrado colaborativo basados en ítems

Métodos de filtrado colaborativo basados en factorización de matrices

La mayoría de los métodos basados en conocimiento

## Com base no conteúdo

- ❑ O usuário recebe recomendações de itens similares a itens nos quais ele/ela expressou interesse no passado. Os perfis dos usuários são construídos utilizando:
  - **Propriedades do item**
  - **Informações para o usuário**



## Com base no conteúdo

Componentes principais:

1. Analisador de conteúdo
2. Aprendizagem do perfil do usuário
3. Filtragem de conteúdo

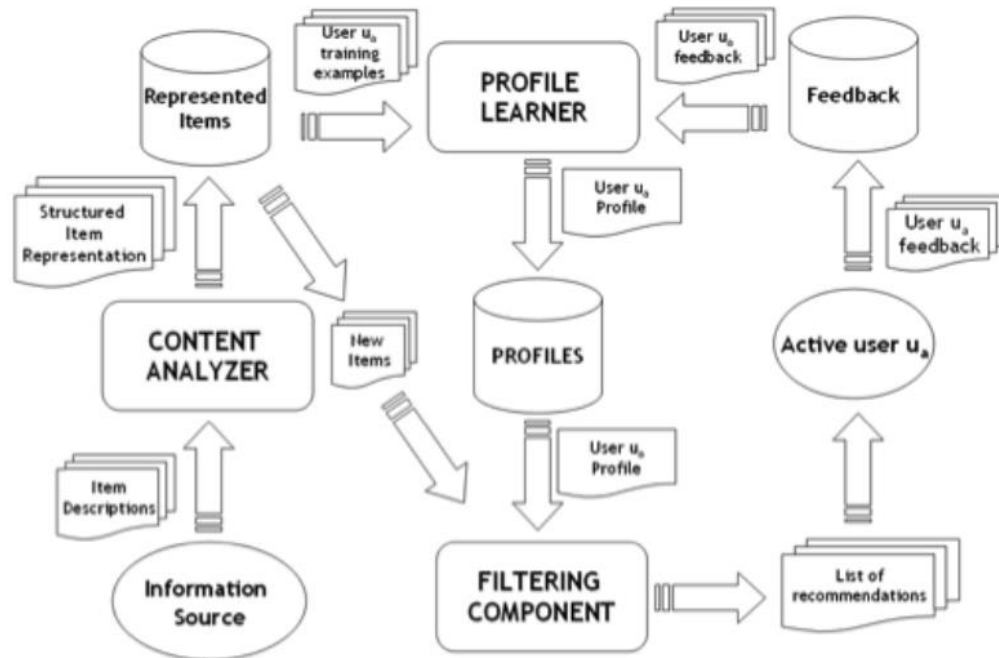


Fig. 3.1: High level architecture of a Content-based Recommender

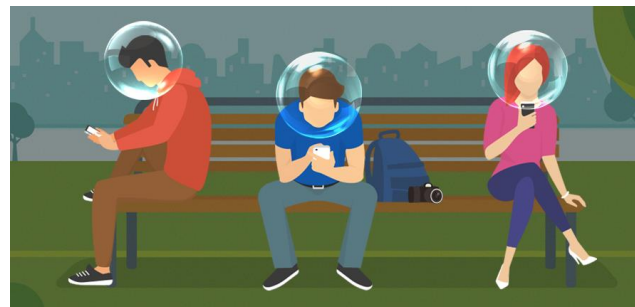
## Com base no conteúdo

### ❑ Prós:

- Nenhuma informação de outros usuários é necessária para fazer a recomendação.
- Escalas facilmente para um grande número de usuários
- O modelo captura os interesses específicos do usuário e pode recomendar informações sobre itens em um nicho muito específico.
- É fácil fazer um sistema transparente, onde as recomendações podem ser explicadas.

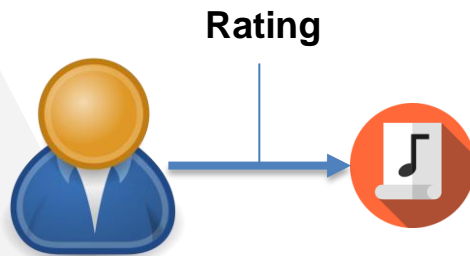
### ❑ Contras:

- É necessário conhecer as propriedades dos itens e a qualidade do modelo dependerá dessas propriedades.
- É necessário conhecer o interesse do usuário, ou seja, o modelo tem pouca capacidade para expandir os interesses do usuário.
- **Recomendação sobre especialização, tendência a “filter bubble”**



# Filtragem colaborativa

- ❑ As recomendações para um usuário são baseadas na opinião de **outros usuários com preferências similares**.
- O usuário tem que expressar suas classificações (rating) em diferentes produtos para ser recomendado.
- A semelhança entre os usuários é calculada com base nas classificações. A partir de las valoraciones se calcula la similitud entre usuarios



						
	4	3			5	
	5		4		4	
	4		5	3	4	
		3				5
		4				4
			2	4		5

# Filtragem colaborativa.

## Classificações do usuário:

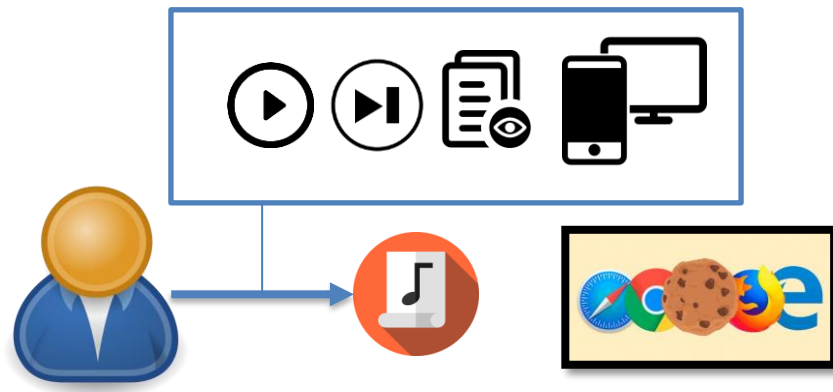
### Explícito (feedback ativo)

Classificações, preferência exibida diretamente pelo usuário.



### Implícito (feedback passivo)

Sinal indireto Log de interação do usuário: visualização, clique, compra, etc.



# Filtragem colaborativa.



## Tipos de técnicas de Filtragem Colaborativa:

### Basados en el usuario

(usuario-usuario)

Se basan en la identificación de usuarios con preferencias similares (vecinos)

### Basados en items

(ítem-ítem)

Se basan en la identificación de ítems similares en cuanto a las valoraciones recibidas por los usuarios

### Factorización de matrices

Descomposición de matrices para una representación más compacta de las preferencias de los usuarios



## Filtragem colaborativa.

### ❑ Com base no usuario (UBCF)

- ❖ Com base nas preferências do grupo de usuários mais parecidas com as do usuário ativo.
- ❖ Use todos os registros da matriz de classificação para procurar usuários K com preferências similares (KNN).
- ❖ É necessária uma função de similaridade, por exemplo:
  - Coeficiente de **correlação Pearson**
  - **Similaridade cosseno**.
- ❖ Uma previsão da classificação de um item j é feita para o usuário ativo.

$$\omega(u_a, u_i) = \frac{\sum_j (r_{aj} - \bar{r}_a)(r_{ij} - \bar{r}_i)}{\sqrt{\sum_j (r_{aj} - \bar{r}_a)^2 (r_{ij} - \bar{r}_i)^2}}$$

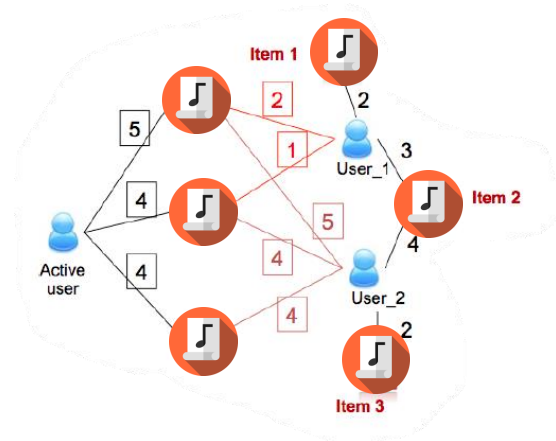
$$pr_{aj} = \bar{r}_a + \frac{\sum_{i=1}^k \omega(u_a, u_i) (r_{ij} - \bar{r}_i)}{\sum_{i=1}^k |\omega(u_a, u_i)|}$$

# Filtragem colaborativa.



## Com base no usuario (UBCF) Exemplo:

1. Encontre os vizinhos K mais próximos (KNN) do usuário, usando uma medida de semelhança (por exemplo, Corr. Pearson). **Este é, portanto, um método baseado na memória..**
2. Prever a **classificação que um usuário dará a um item.**
3. Fornecer ao usuário uma lista top-N de recomendações.



K=2

	active user
user_1	0.4472136
user_2	0.49236596
user_3	-0.91520863

	User_1	User_2	predicted interest
item_1	2	-	3.262013982
item_2	3	4	3.56331003
item_3	-	2	2.277268083

Lista top-N

	item_2
	item_1
	item_3

# Filtragem colaborativa.

## Baseado no item (IBCF):

Também considera as semelhanças do item a serem recomendadas com outros itens que o usuário já tenha classificado.

1. Primeiro, as semelhanças entre os itens são pré-calculadas.
2. A previsão de classificação é feita usando uma soma ponderada ou regressão.

Dentro desta família de algoritmos está SLOPE ONE

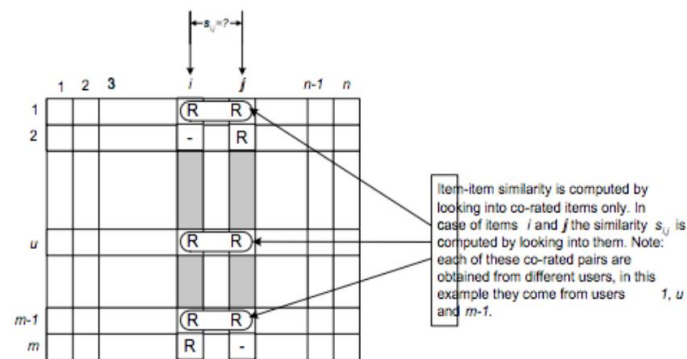


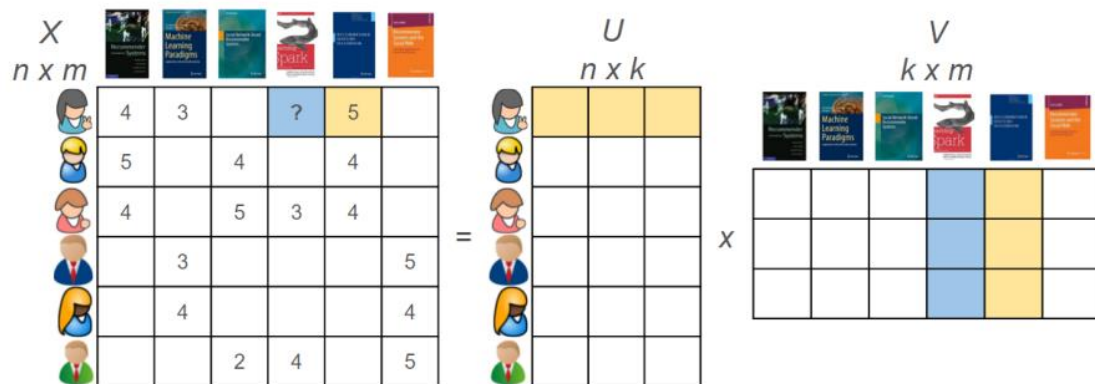
Figure 2: Isolation of the co-rated items and similarity computation



# Filtragem colaborativa.

## Factorização de matrizes:

Consiste em decompor a matriz de classificações contendo muitos elementos vazios em duas matrizes **menos esparsas**.



# Filtragem colaborativa.

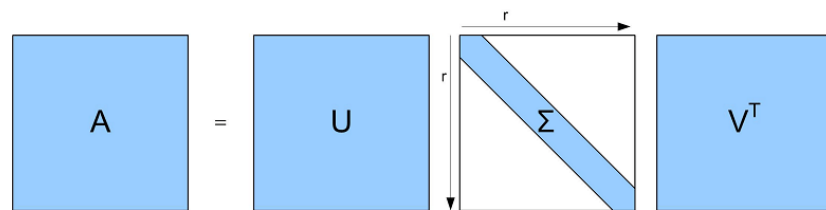
## Factorização de matrizes

**Motivação:** Enormes matrizes de classificação! Amazon, Netflix, etc.

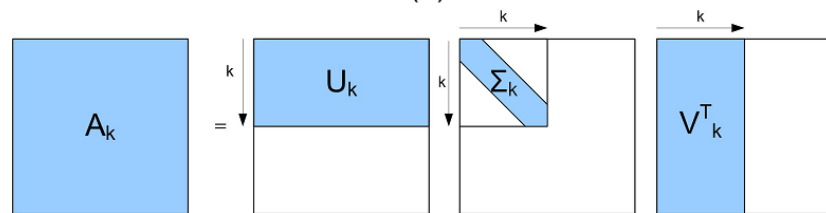
**Solução:** Redução da dimensionalidade.

Métodos:

- ▶ SVD
- ▶ Funk's SVD
- ▶ NMF
- ▶ Etc.



(a)

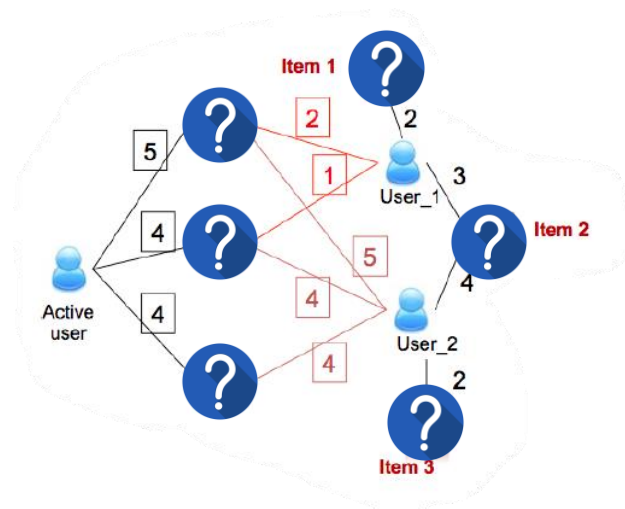


(b)

# Filtragem colaborativa.

## ❑ Vantagens:

- Fácil de implementar.
- Incentiva a serendipidade.
- Aplicação agnóstico do domínio



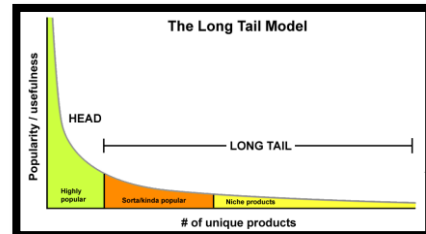
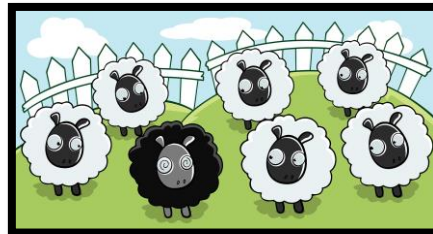
# Filtragem colaborativa.



## Limitações:

- Dispersão (Sparsity)
- Escalabilidade
- Primeira avaliação (*early rater*) e arranque a frio (*cold start*)
- Oveja negra o Grey Sheep
- Long Tail

	product rated by user							
	1	2	3	4	5	6	7	8
user id	1.0	0	5.0	0	0	0	0	0
	0	3.0	0	0	0	0	11.0	0
	0	0	0	0	9.0	0	0	0
	0	0	6.0	0	0	0	0	0
	0	0	0	7.0	0	0	0	0
	2.0	0	0	0	0	10.0	0	0
	0	0	0	8.0	0	0	0	0
	0	4.0	0	0	0	0	0	12.0





# Sistemas baseados no conhecimento



## **Knowledge-Based Systems (métodos baseados no conhecimento):**

Eles utilizam informações dependentes de domínio sobre itens e usuários..



## **Fases:**

1. Construção de um banco de dados descrevendo as propriedades dos itens disponíveis no sistema.
2. Obtenção das preferências do usuário com base nos itens selecionados pelo usuário
3. Cálculo das semelhanças de outros itens usando alguma medida de distância entre itens que leva em conta os valores de seus atributos (ou propriedades).
4. Obtenção de uma lista de itens de interesse para o usuário
5. Recomendação

# Context-Aware

RecSys sensível ao contexto (CARS)

# CARS

- ❑ Os sistemas CARS (Context-aware recommendations systems) fornecem recomendações baseadas tanto em informações de rotina como em **informações contextuais**.
- ❑ Dependendo do domínio, estas informações podem ter **um impacto direto nas preferências dos usuários**.
- **Informações contextuais:**
  - Clima atmosférico
  - Hora: mês, semana, dia, hora, etc.
  - Atividade realizada.
  - Localização física.
  - Estado de espírito.
  - Etc.



# CARS

- ❑ O problema da recomendação tradicional (2D)

$$R: User \times Item \rightarrow Rating$$

- ❑ Muda para (MD):

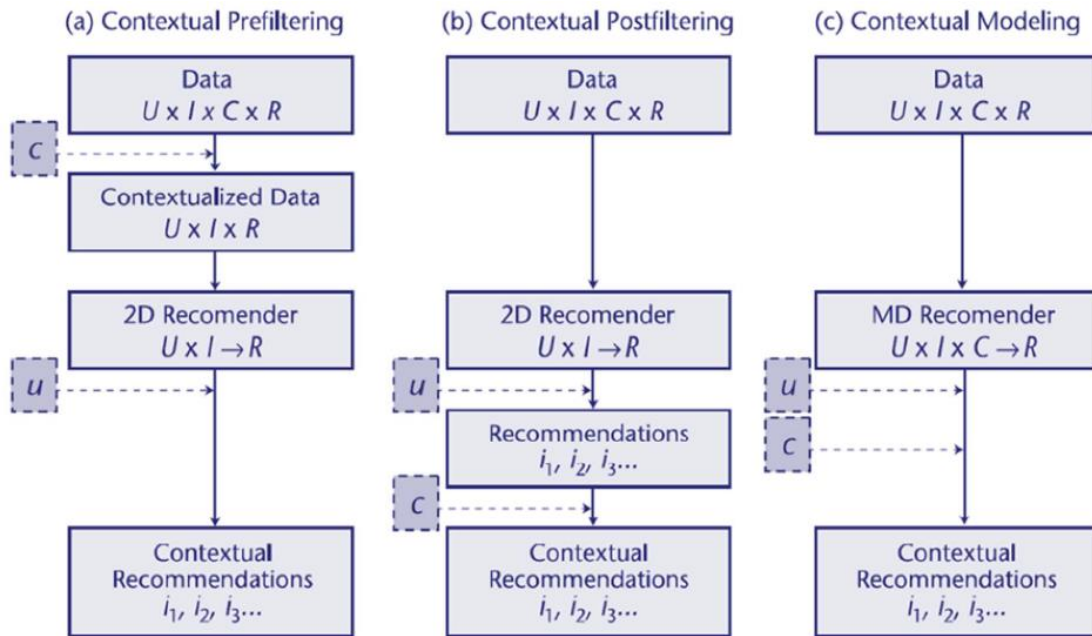
$$R: User \times Item \times context \rightarrow Rating$$

Usuario	Película	Tiempo	Localización	Compañía	Rating
U1	Titanic	Fin de semana	Casa	Novia	4
U2	Titanic	Fin de semana	Casa	Hermano	5
U3	Titanic	Diario	Cine	Amigo	4

U1	Titanic	Diario	Casa	Hermano	?
----	---------	--------	------	---------	---



## Abordagens:



# Domínio da música.

Aspectos específicos a este domínio

## Aspectos chave do domínio da música

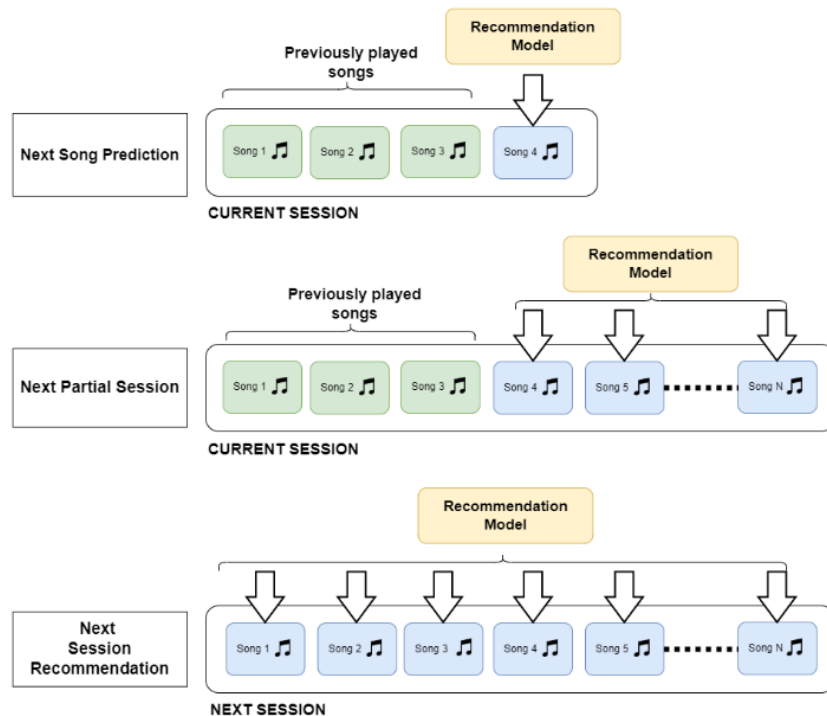
- ☐ Um item **pode ser consumido várias vezes no mesmo dia**, ao contrário de outros domínios.
- ☐ O tempo de consumo de um **item é muito limitado**.
- ☐ Os usuários são mais tolerantes **a uma má recomendação**
- ☐ Recomendação em **diferentes níveis**: artista, álbum, canção.
- ☐ O feedback é, em sua maioria, **implícito**.
- ☐ Conjuntos de dados com **baixa densidade** (alta esparsidade) devido ao tamanho dos catálogos.
- ☐ Muitas abordagens baseadas no conteúdo da área MIR, porém o usuário às vezes procura **descobrir novas músicas (Avaliação da recomendação)**.
- ☐ O contexto do usuário tem um papel vital nas preferências do usuário.





# Tarefas de recomendação em sistemas de streaming

- ❑ Próxima canção
- ❑ Próxima lista parcial
- ❑ Próxima lista de reprodução (*playlist*)

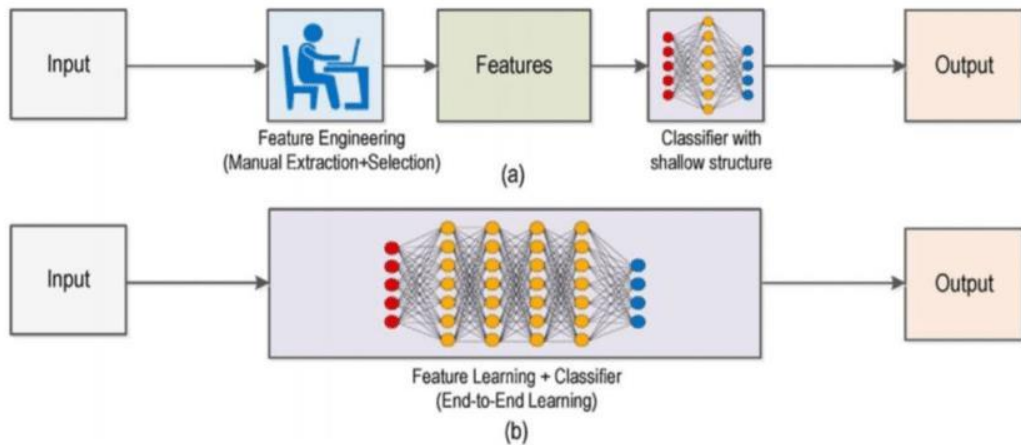


# Tendências atuais RecSys

Trabalhos atuais na Music RecSys

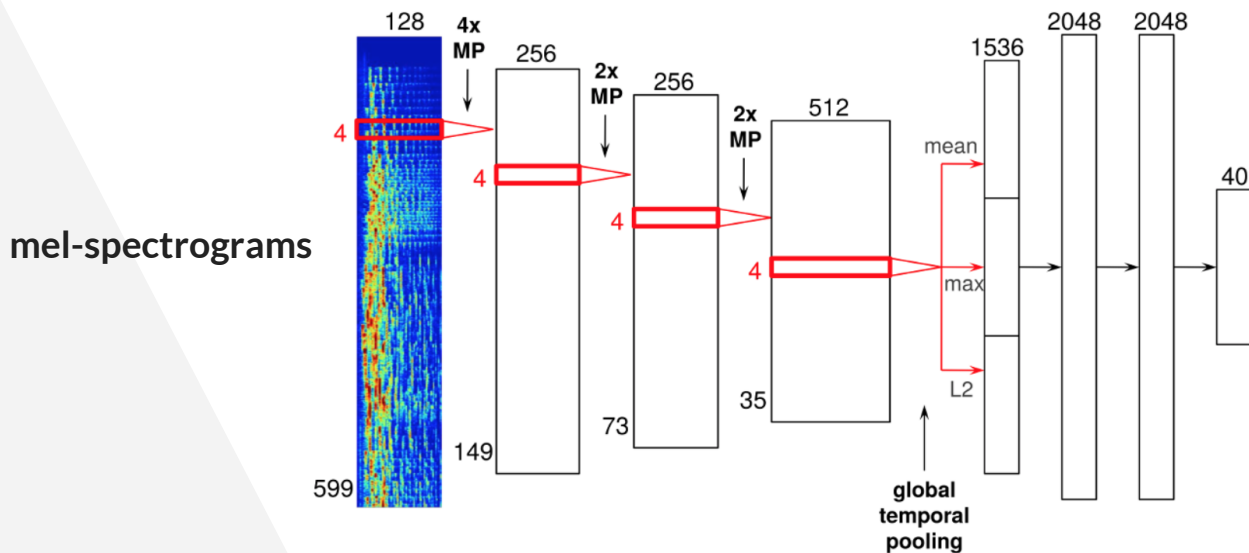
# Deep Learning

## □ Deep Learning:



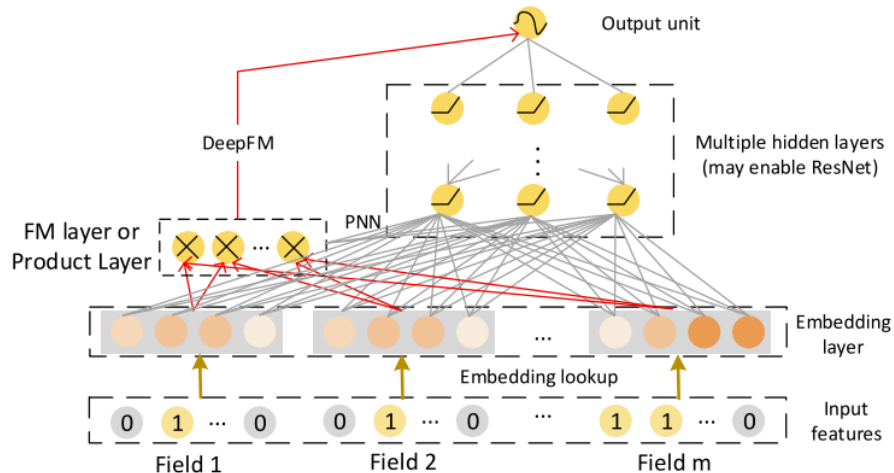
# Deep Learning

## Exatção das características do sinal bruto



# Aprendizagem profunda e Filtragem colaborativa

- Introduzir estas representações latentes em modelos de filtragem colaborativa.

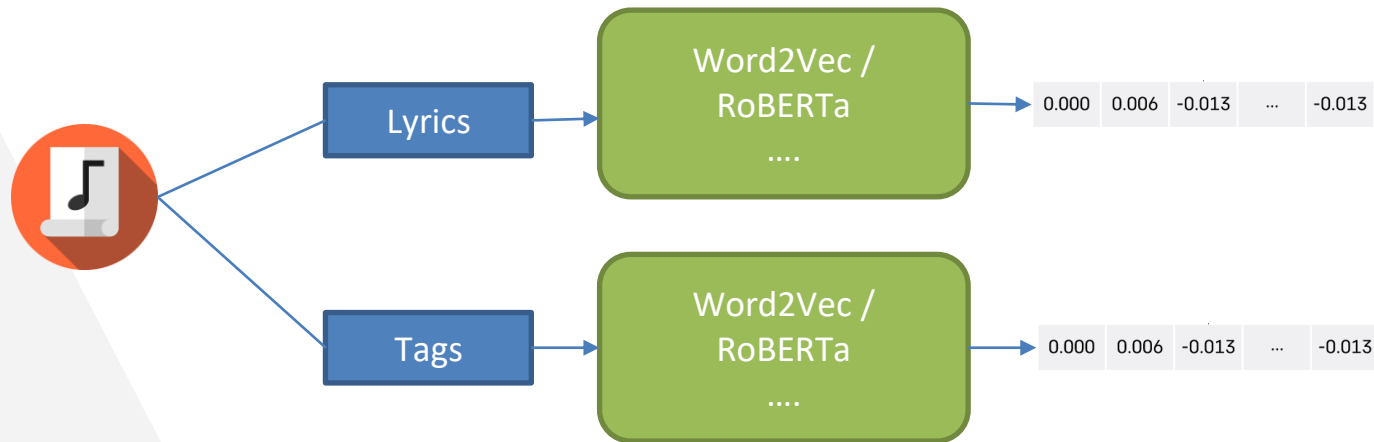


**xDeepFM:** Combining Explicit and Implicit Feature Interactions for Recommender Systems

# NLP (Embeddings)

## ❏ NLP (Natural Language Processing)

- Obtenção de representações sob a forma de *embeddings* de letras de músicas utilizando modelos como **Word2Vec**, **BERT**, **RoBERTa** (entre outros).



# Transformers

- Usando arquiteturas baseadas em Transformers para obter essas *embeddings*

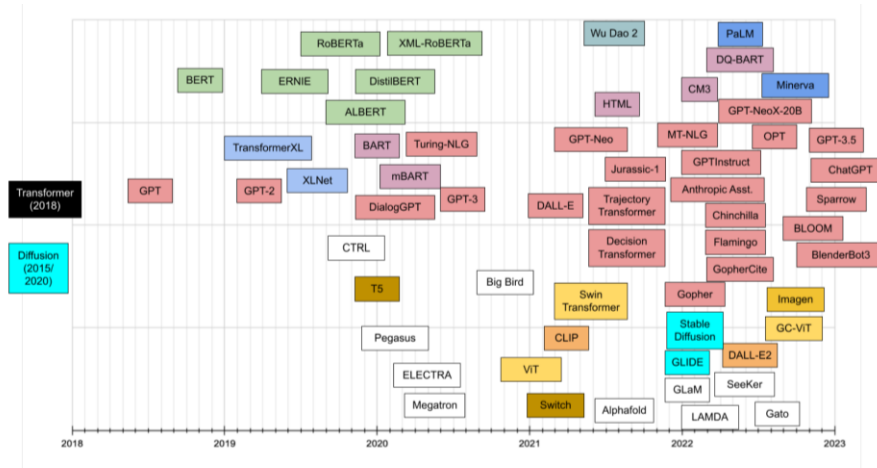


Figure 7: Transformer timeline. Colors describe Transformer family.

Origem:: [Xavier Amatriain TRANSFORMER MODELS: AN INTRODUCTION AND CATALOG](#)

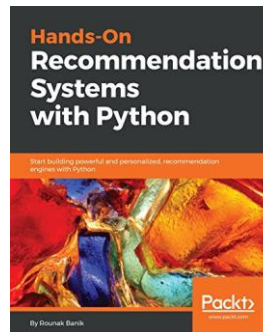
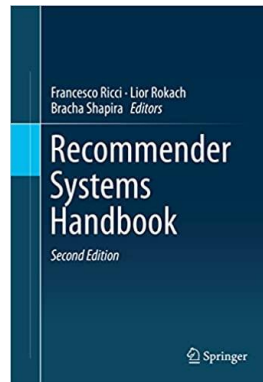
# Por onde começar?

Lista de recursos úteis



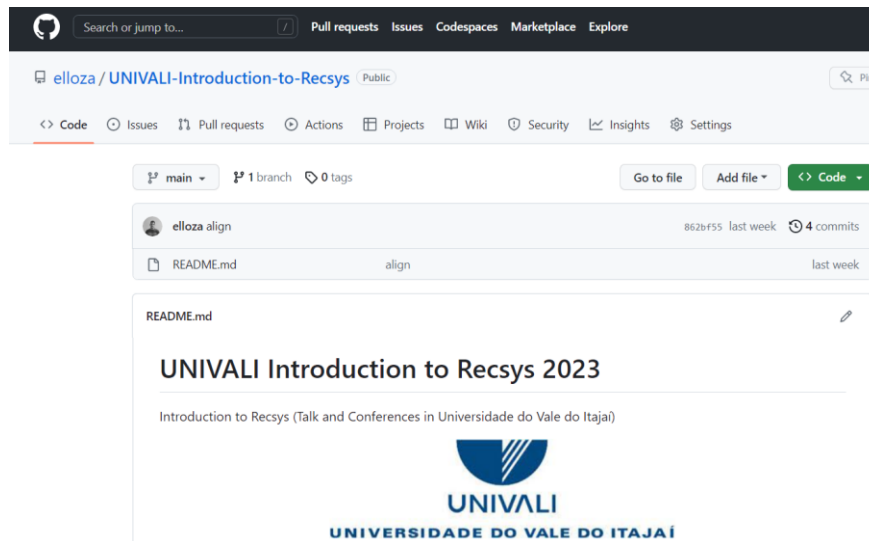
# Recursos

- Cursos (do menos ao mais difícil):
  - ❖ [Recommender Systems Google](#)
  - ❖ [Coursera Specialization](#)
  - ❖ [Recommender systems behind de screen](#)
  - ❖ [Denis Parra Course UPC](#)
- Livros:
  - [Recommender Systems Handbook](#) (Livro de referência teórica)
  - [Hands-On Recommendation Systems with Python](#) (Exemplos práticos)
- Repositórios Github:
  - [Best Practices Recsys \(Microsoft\)](#)
  - [Awesome Recsys](#)



# Recursos

- Apresentação de slides:  
<https://github.com/elloza/UNIVALI-Introduction-to-Recsys>



# Muito obrigado!



## Contato:

Email: [loza@usal.es](mailto:loza@usal.es)

Twitter: [@\\_lozanillo\\_](https://twitter.com/_lozanillo_)

Website: <https://elloza.com/>

Github, recursos e apresentação:

<https://github.com/elloza/UNIVALI-Introduction-to-Recsys>



VNIVERSIDAD  
D SALAMANCA  
CAMPUS DE EXCELENCIA INTERNACIONAL



**ESALAB**  
Expert Systems and Applications Laboratory



UNIVALI