

Sistemas Recomendadores con *recommenderlab*

Elías Alegría P.

30 de junio, 2020

Agenda

- ❑ ¿Por qué un sistema recomendador?
- ❑ Tipos de sistemas recomendadores
- ❑ Collaborative filtering
 - User based
 - Item based
 - Matrix Factorization

¿Por qué sistemas de recomendación?





Spotify®

Objetivo

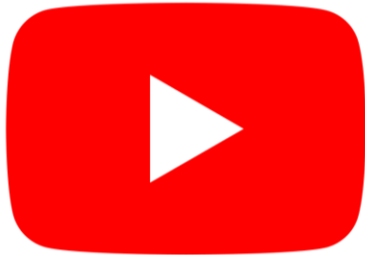
Un Sistema Recomendador trata de identificar necesidades y preferencias de los usuarios, filtran una enorme cantidad de data y presenta la opción más adecuada a los usuarios

Ejemplos

NETFLIX



Ejemplos

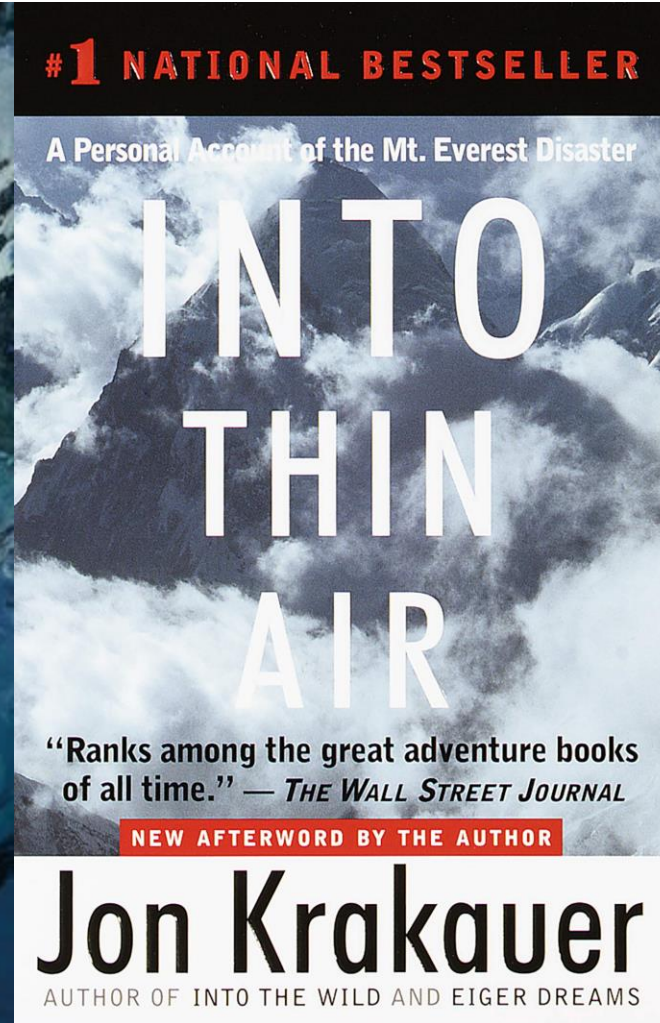


amazon

Historia de éxito



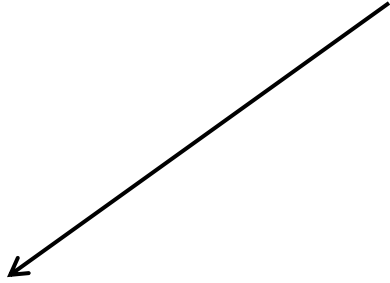
1988



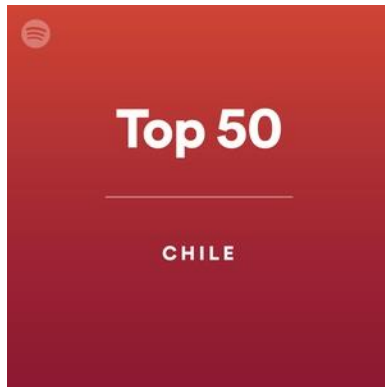
1996

Tipos de Sistemas Recomendadores

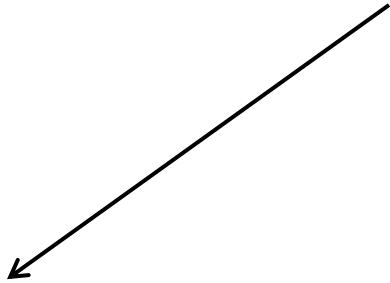
Tipos de Sistemas Recomendadores



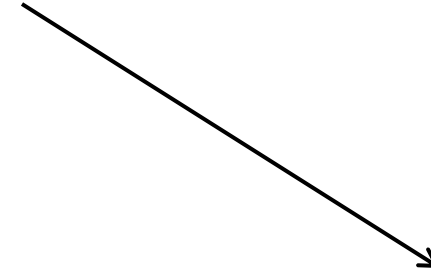
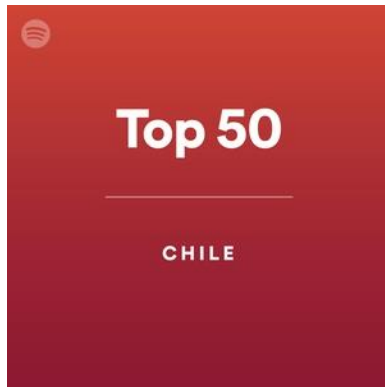
Popularity based



Tipos de Sistemas Recomendadores



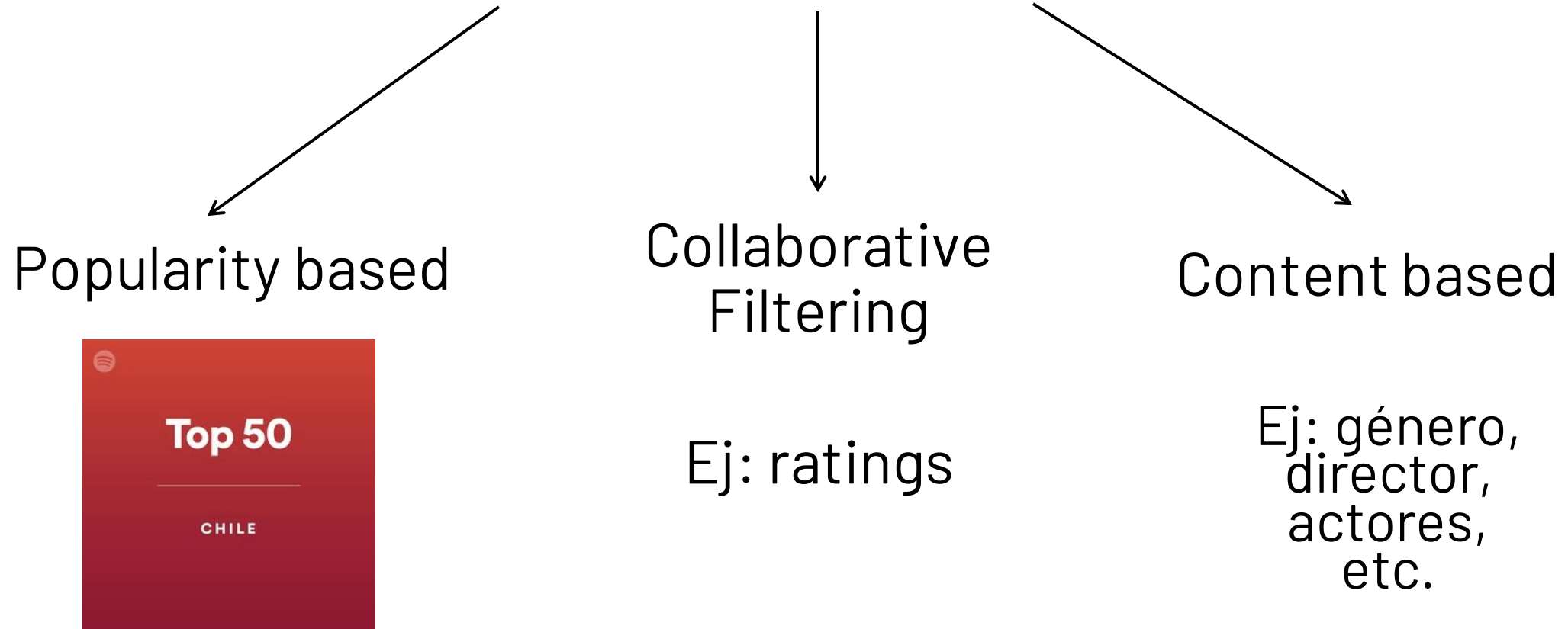
Popularity based



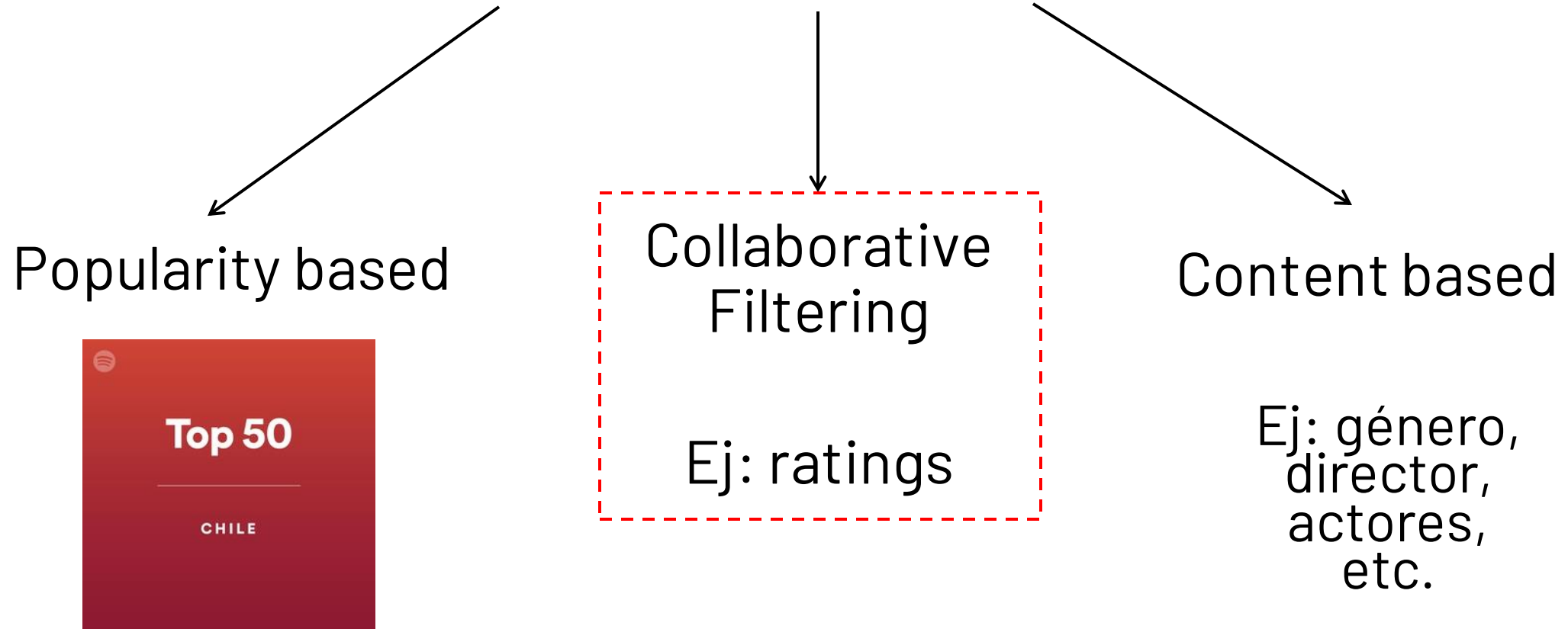
Content based

Ej: género,
director,
actores,
etc.

Tipos de Sistemas Recomendadores



Tipos de Sistemas Recomendadores



Collaborative Filtering

Recomendar película a **User 3**

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5
User 1	5	3	1	2	5
User 2	2	4	4		2
User 3	4	3	1	?	?

Collaborative Filtering

Recomendar película a **User 3**

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5
User 1	5	3	1	2	5
User 2	2	4	4		2
User 3	4	3	1	?	?

Opción 1: Encontrar un usuario similar y recomendar una película que a él le guste. **User Based CF**

Collaborative Filtering

Recomendar película a **User 3**

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5
User 1	5	3	1	2	5
User 2	2	4	4		2
User 3	4	3	1	?	?

Opción 1: Encontrar un usuario similar y recomendar una película que a él le guste. **User Based CF**

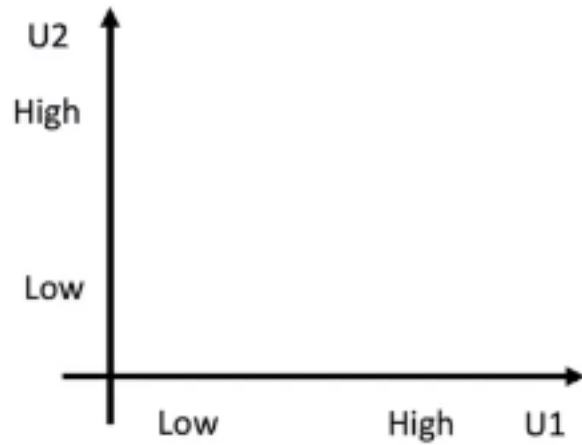
Opción 2: Encontrar películas similares a las que le gustan al usuario, basado en sus ratings. **Item Based CF**

Collaborative Filtering

Recomendar película a **User 4**

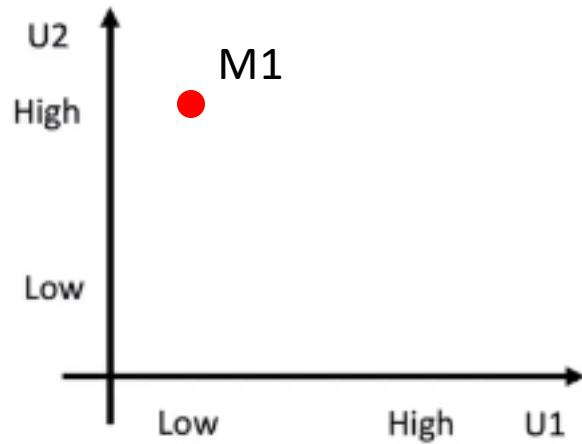
	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5	Movie 6
User 1	5	3	1	2	5	3
User 2	2	4	4	1	2	5
User 3	4	3	1		4	3
User 4	1	5	4			

Cuantificando la similitud



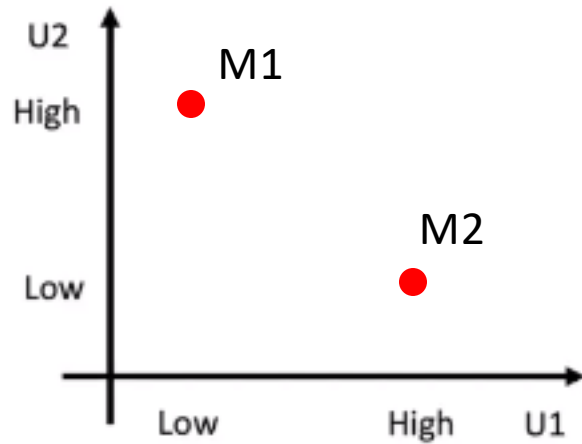
Usuarios diferentes

Cuantificando la similitud



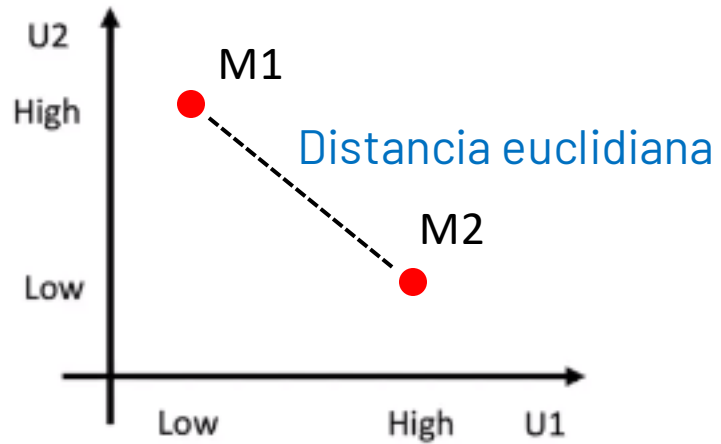
Usuarios diferentes

Cuantificando la similitud



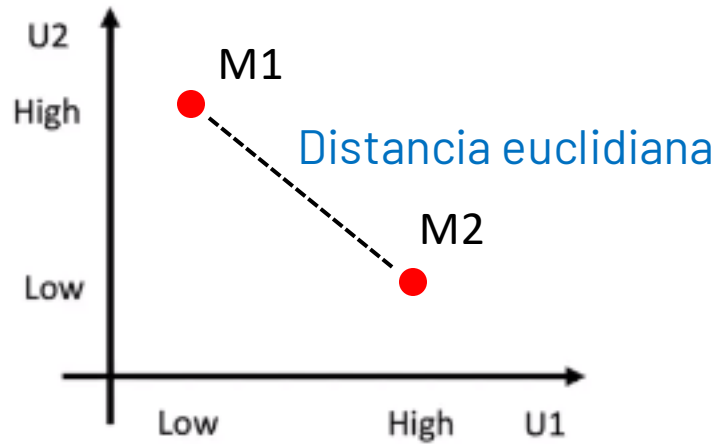
Usuarios diferentes

Cuantificando la similitud

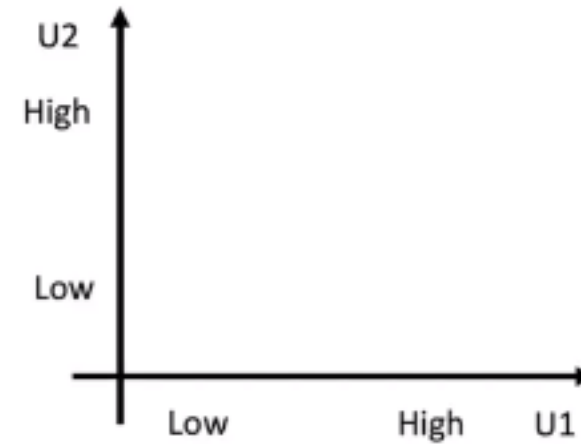


Usuarios diferentes

Cuantificando la similitud

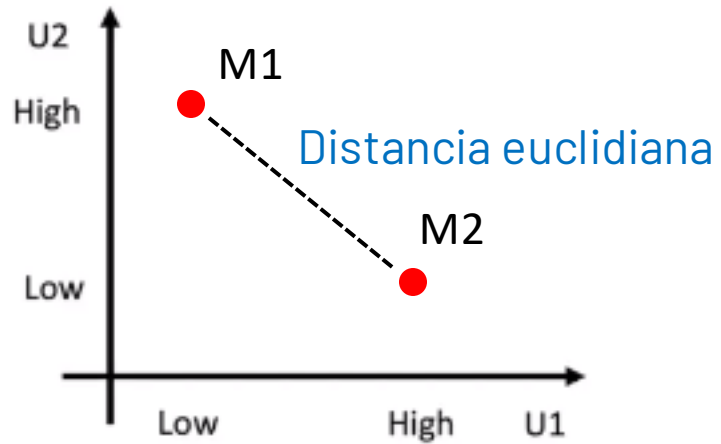


Usuarios diferentes

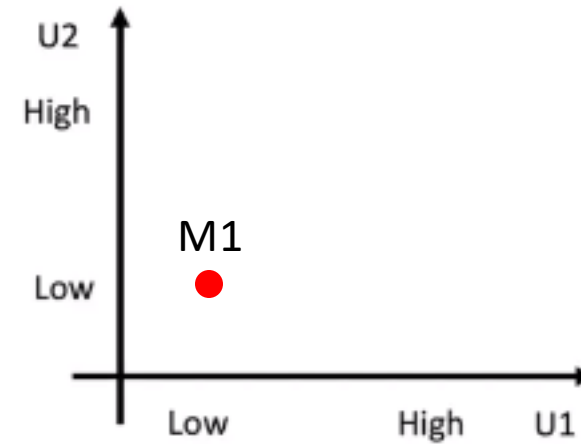


Usuarios similares

Cuantificando la similitud

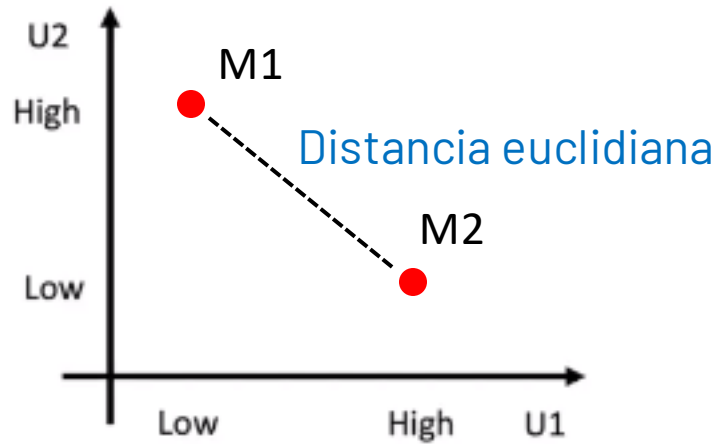


Usuarios diferentes

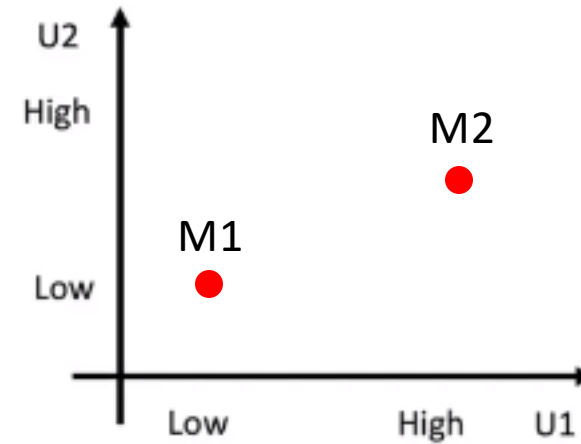


Usuarios similares

Cuantificando la similitud

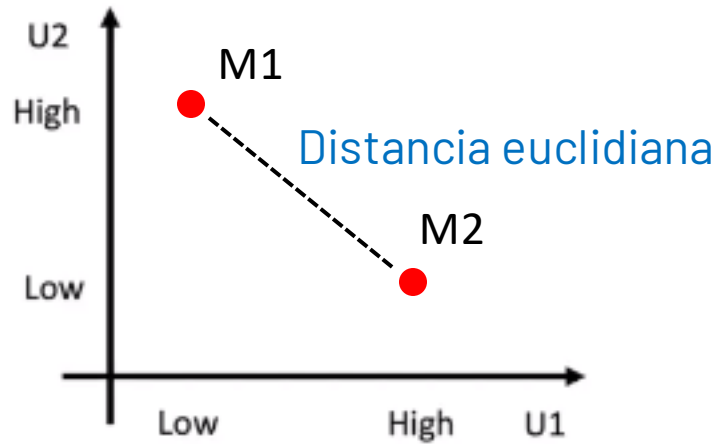


Usuarios diferentes

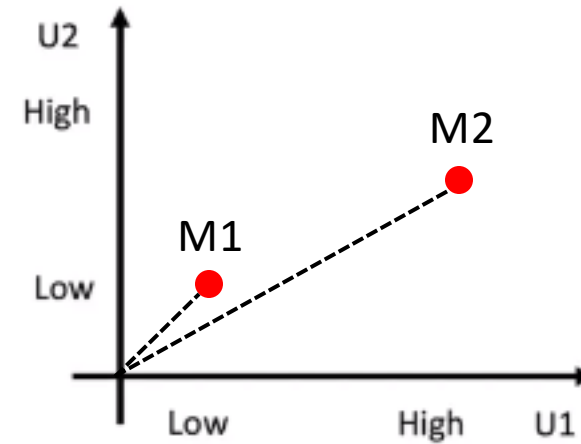


Usuarios similares

Cuantificando la similitud



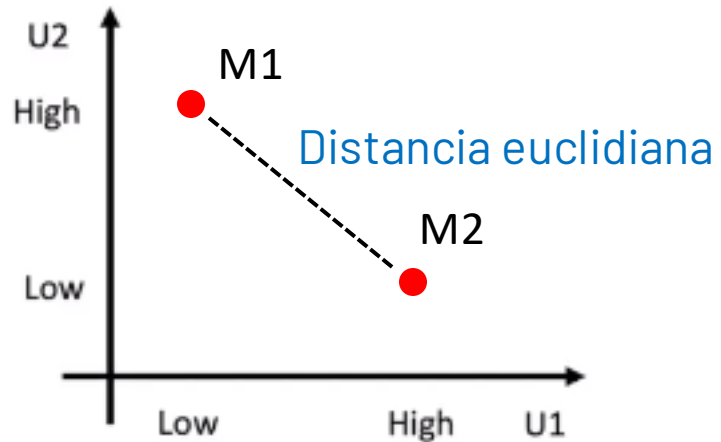
Usuarios diferentes



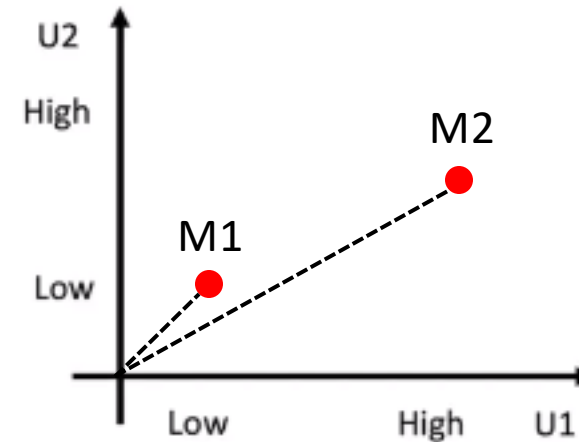
Usuarios similares

Cuantificando la similitud

Opción 1: Distancia coseno

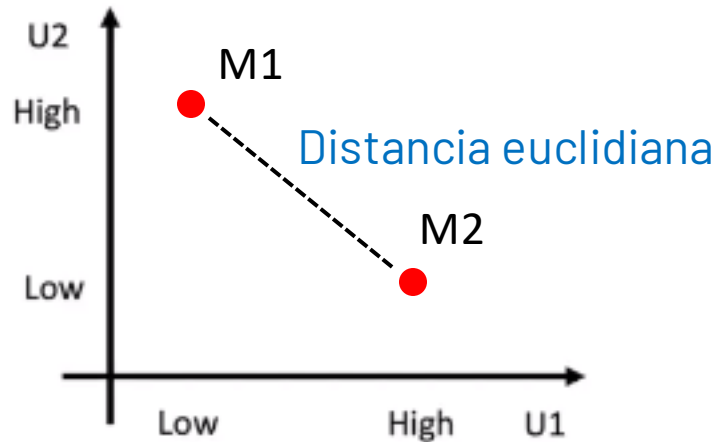


Usuarios diferentes



Usuarios similares

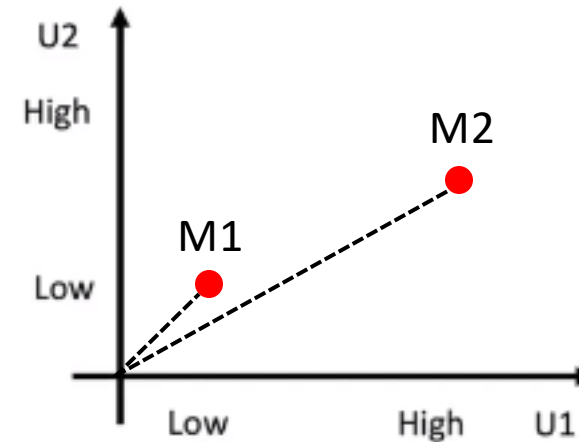
Cuantificando la similitud



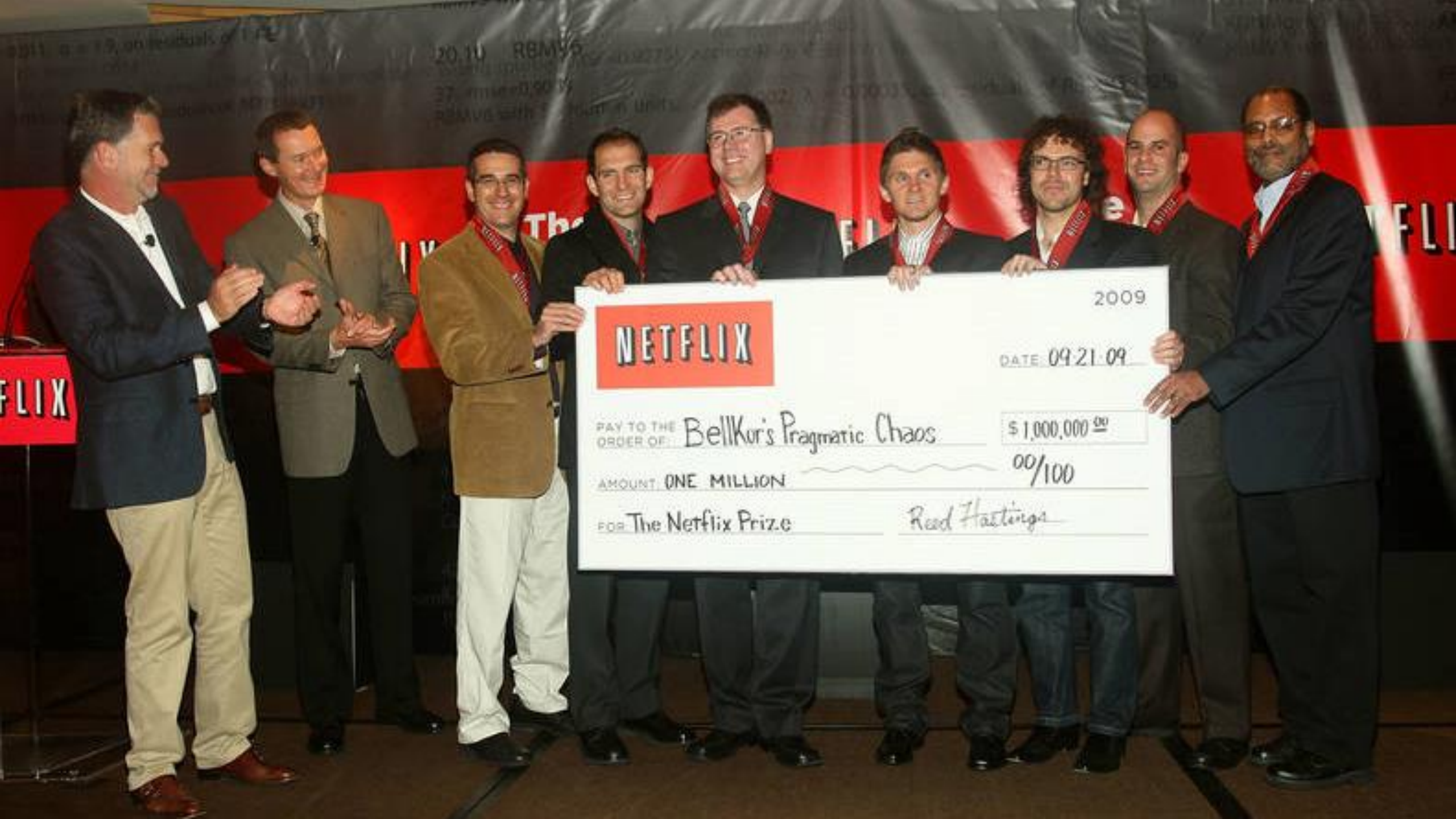
Usuarios diferentes

Opción 1: Distancia coseno

Opción 2: Correlación de Pearson



Usuarios similares



2011, 0.19, on residuals of 1.72
20.10 RBMVS
37.71 RBMVS with 1.72 on residuals of 1.72

NETFLIX

2009

NETFLIX

DATE 09.21.09

PAY TO THE ORDER OF: BellKor's Pragmatic Chaos

AMOUNT ONE MILLION

FOR The Netflix Prize

\$1,000,000.00

00/100

Reed Hastings

Collaborative Filtering

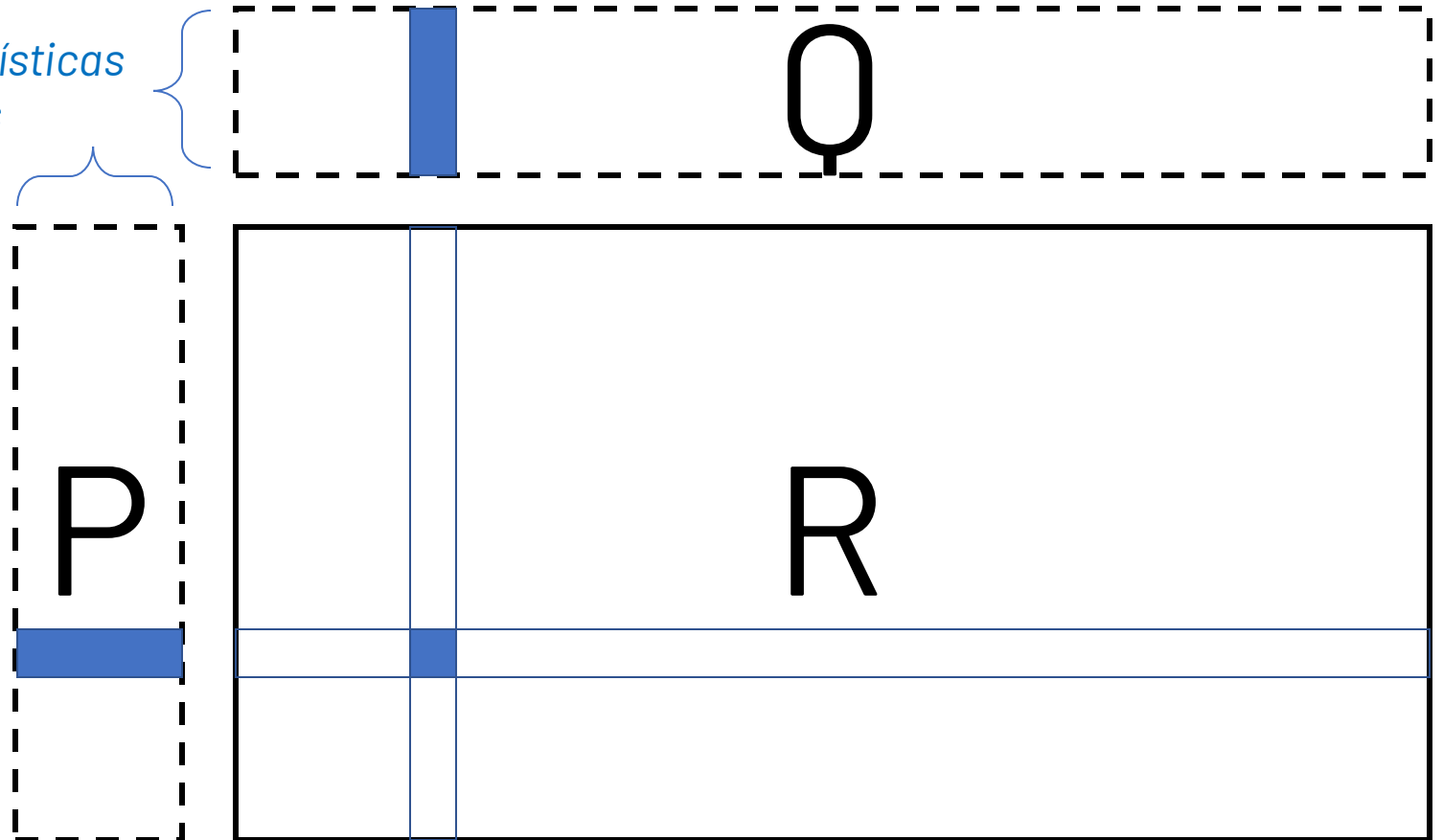
Matrix Factorization

Algoritmos de Matrix Factorization:

- *Single Value Decomposition*
- *Alternating Last Squares*

$$R \approx P^T Q$$

*Catacterísticas
Latentes*



Collaborative Filtering

Matrix Factorization

	M1	M2	M3	M4	M5
Comedy	3	1	1	3	1
Action	1	2	4	1	3

	Comedy	Action
U1	1	0
U2	0	1
U3	1	0
U4	1	1

	M1	M2	M3	M4	M5
U1	3	1	1	3	1
U2	1	2	4	1	3
U3	3	1	1	3	1
U4	4	3	5	4	4

Métricas de desempeño

```
graph TD; A[Métricas de desempeño] --> B[Predicción de rating]; A --> C[Predicción de elección];
```

Predicción de rating

- RMSE
- MSE
- MAE

Predicción de elección

Se predice una cantidad de ítems para c/ usuario. Luego, se calculan métricas: **Presicion, recall, TPR, FPR**, etc.

recommenderlab

Sistemas Recomendadores con *recommenderlab*

Elías Alegría P.

30 de junio, 2020

Anexos

Fórmulas de similitud

$$\text{sim}_{\text{Pearson}}(\vec{x}, \vec{y}) = \frac{\sum_{i \in I} (\vec{x}_i \bar{\vec{x}})(\vec{y}_i \bar{\vec{y}})}{(|I| - 1) \text{sd}(\vec{x}) \text{sd}(\vec{y})}$$

$$\text{sim}_{\text{Cosine}}(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|},$$

$$\text{sim}_{\text{Jaccard}}(\mathcal{X}, \mathcal{Y}) = \frac{|\mathcal{X} \cap \mathcal{Y}|}{|\mathcal{X} \cup \mathcal{Y}|},$$

Matrix factorization problem

$$\arg \min_{H, W} \|R - \tilde{R}\|_F + \alpha \|H\| + \beta \|W\|$$