

Caso Práctico (Módulo 0)

Sistemas de Recomendación

¿Qué es un sistema de recomendación?

- Soluciones que realizan recomendaciones para emparejar usuarios con productos
- El objetivo es ayudar a los usuarios a tomar decisiones frente a grandes cantidades de información, mejorando la experiencia de usuario y reduciendo el tiempo empleado en la selección de productos a consumir.



Tipos de sistemas de recomendación

- **Basados en el Contenido:** Elaboran recomendaciones centrándose en la información que pueden extraer de los objetos. Por ejemplo, el género o los actores en una película.
- **Filtrado Colaborativo:** Utilizan a usuarios con gustos afines u objetos similares para estimar recomendaciones.
- **Basados en el Contexto:** Recomiendan según las características generales de los usuarios como la edad, género, ciudad, etc.
- **Basados en Conocimiento:** Consideran las necesidades e intereses del usuario para realizar recomendaciones.
- **Basados en Comunidades:** Recomiendan objetos en función de las preferencias de los amigos de los usuarios. Siguen el refrán, “Dime con quién andas y te diré quién eres”.
- **Híbridos:** Es la combinación de dos o más enfoques

Datos Disponibles

Dentro del archivo Excel del caso práctico, los datos que se utilizarán se encuentran en dos hojas:

1) **Películas:** Contiene información acerca de las 100 películas de nuestro catálogo

movieId	title	genres
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
2	Jumanji (1995)	Adventure Children Fantasy
3	Grumpier Old Men (1995)	Comedy Romance
4	Waiting to Exhale (1995)	Comedy Drama Romance
5	Father of the Bride Part II (1995)	Comedy

2) **Ratings:** Contiene información acerca de las puntuaciones de los usuarios de nuestra plataforma

userId	movieId	rating
1	1	4
1	3	3



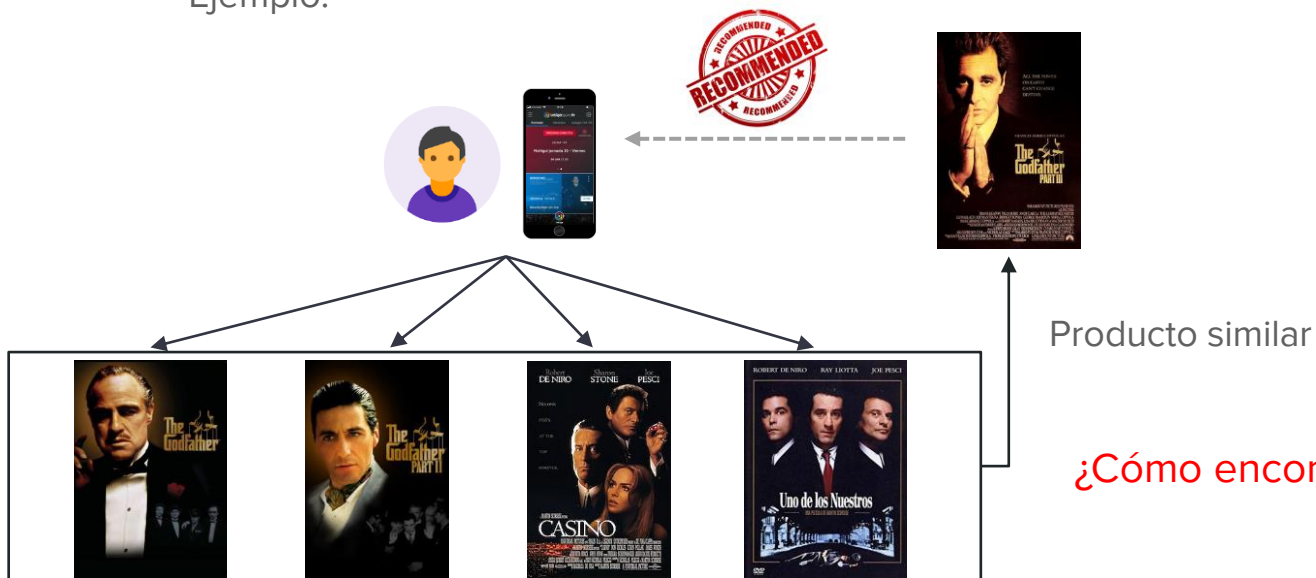
El usuario con ID=1 ha visualizado “Toy Story” y “Grumpier Old Men”, puntuándolas con 4 y 3 estrellas, respectivamente

Algoritmo

Filtrado Colaborativo Item – Item

Estos algoritmos realizan recomendaciones utilizando los consumos del usuario y la afinidad entre los items

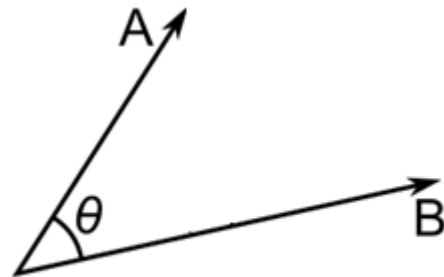
Ejemplo:



¿Cómo encontrar productos similares?

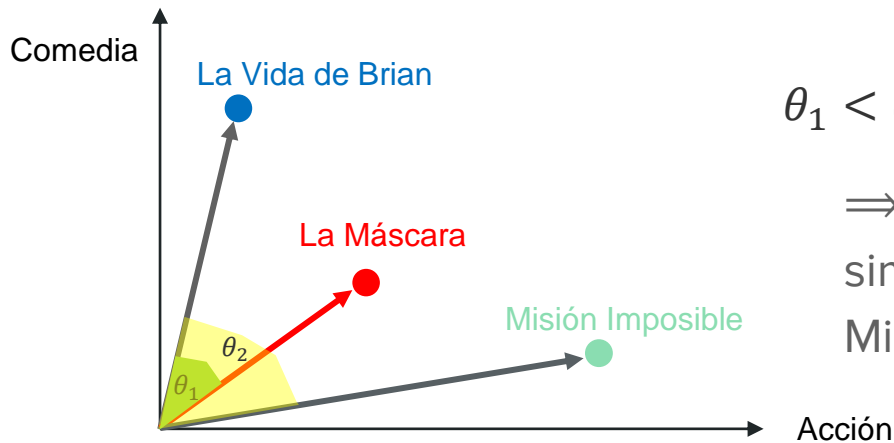
Similitud entre vídeos

$$\text{simil}(X, Y) = \cos(\theta) = \frac{v_A v_B}{|v_A| \cdot |v_B|}$$



A y B son los objetos

v_A y v_B son los vectores que representan los objetos A y B



$$\theta_1 < \theta_2 \Rightarrow \cos(\theta_1) > \cos(\theta_2)$$

\Rightarrow La Vida de Brian es más similar a La Máscara que a Misión Imposible

Algoritmo

PASO 1. Se representan las películas con vectores binarios

movielf	title	genres
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
2	Jumanji (1995)	Adventure Children Fantasy
3	Grumpier Old Men (1995)	Comedy Romance Drama
4	Waiting to Exhale (1995)	Comedy Drama Romance
5	Father of the Bride Part II (1995)	Comedy



movielf	Adventure	Animation	Children	Comedy	Fantasy	Romance	Drama
1	1	1	1	1	1	0	0
2	1	0	1	0	1	0	0
3	0	0	0	1	0	1	1

Algoritmo

$$\text{simil}(A, B) = \cos(\theta) = \frac{v_A v_B}{|v_A| |v_B|}$$

PASO 2. Se normalizan los vectores

movied	Adventure	Animation	Children	Comedy	Fantasy	Romance	Drama
1	1	1	1	1	1	0	0
2	1	0	1	0	1	0	0
3	0	0	0	1	0	1	1



movied	Adventure	Animation	Children	Comedy	Fantasy	Romance	Drama
1	0.45	0.45	0.45	0.45	0.45	0.00	0.00
2	0.58	0.00	0.58	0.00	0.58	0.00	0.00
3	0.00	0.00	0.00	0.58	0.00	0.58	0.58

Algoritmo

$$\text{simil}(A, B) = \cos(\theta) = \frac{v_A v_B}{|v_A||v_B|}$$

PASO 3. Se crea la matriz de similitudes para todas las películas

movied	Adventure	Animation	Children	Comedy	Fantasy	Romance	Drama
1	0.45	0.45	0.45	0.45	0.45	0.00	0.00
2	0.58	0.00	0.58	0.00	0.58	0.00	0.00
3	0.00	0.00	0.00	0.58	0.00	0.58	0.58



movied	1	2	3
1	1	0.77	0.26
2	0.77	1	0
3	0.26	0	1

EXCEL: función
SUMAPRODUCTO

Algoritmo

PASO 4. Realizar recomendaciones en base a las similitudes entre películas y las puntuaciones de los usuarios



	movieID	Rating
$similitud(1,3)$	1	4
	2	3
$similitud(2,3)$	3	???

$$\begin{aligned} Score(user, 3) &= rating_1 \times similitud(1,3) \\ &+ rating_2 \times similitud(2,3) \\ &= 4 \times 0.26 + 3 \times 0 = 1.04 \end{aligned}$$

Haciendo este cálculo para cada película no visualizada, podemos establecer un ránking de los vídeos con mayor score



userId=1

movieID	Score
54	7.56
21	6.45
13	3.21
3	1.04

¡Gracias!

Contacto: Rafael Zambrano

rafazamb@gmail.com