

Reglas de Asociación

Apriori

Como se puede observar hay una serie de asociación entre las películas

User ID	Películas que le han gustado
46578	Película1, Película2, Película3, Película4
98989	Película1, Película2
71527	Película1, Película2, Película4
78981	Película1, Película2
89192	Película2, Película4
61557	Película1, Película3

Reglas Significativas:	Película1	→	Película2
	Película2	→	Película4
	Película1	→	Película3

El algoritmo Apriori sirve para ver cuales asociaciones son más fuertes y hacer un buen sistema de recomendaciones y se divide en 3 pasos:

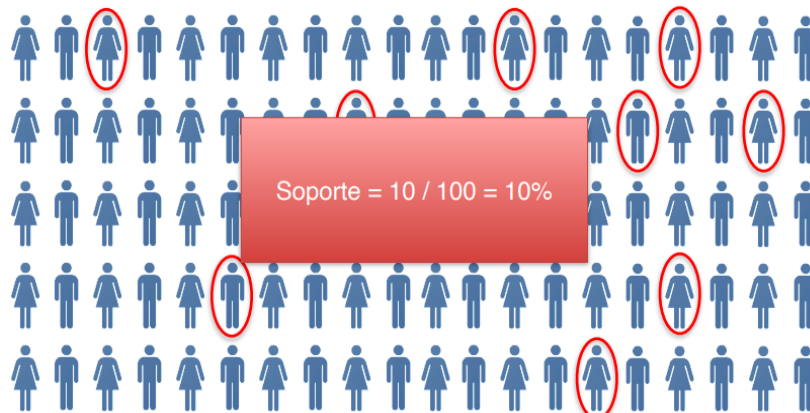
1 Soporte

El soporte es la cantidad de usuarios que vieron una película dividido la cantidad de usuarios, es un porcentaje

Recomendación de Películas:

$$sop(\mathbf{M}) = \frac{|\text{usuarios que vieron } \mathbf{M}|}{|\text{usuarios}|}$$

Por ejemplo, hay 10 personas que le gusta la película de Batman en un dataset de 100 personas



2 Confianza

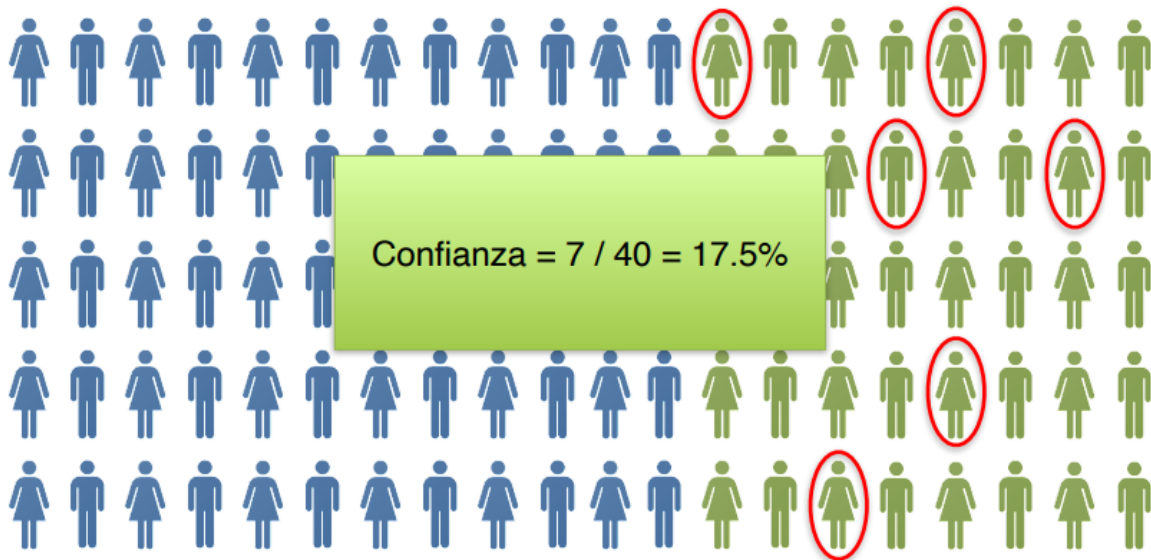
Nos preguntamos cual es la confianza de cuando vemos una película ver otra, sería como que tan seguro estamos de que ocurran 2 cosas a la vez

Recomendación de Películas:

$$conf(M_1 \Rightarrow M_2) = \frac{|\text{usuarios que vieron } M_1 \text{ y } M_2|}{|\text{usuarios que vieron } M_1|}$$

Calculamos cuantos usuarios han visto la primera película y luego la otra

Por ejemplo, de verde están las personas que vieron El Joker y luego vieron Batman



3 Lift

Es parecido a Naïves Bayes, sería la confianza/soporte

Recomendación de Películas:

$$lift(M_1 \Rightarrow M_2) = \frac{conf(M_1 \Rightarrow M_2)}{sup(M_2)}$$

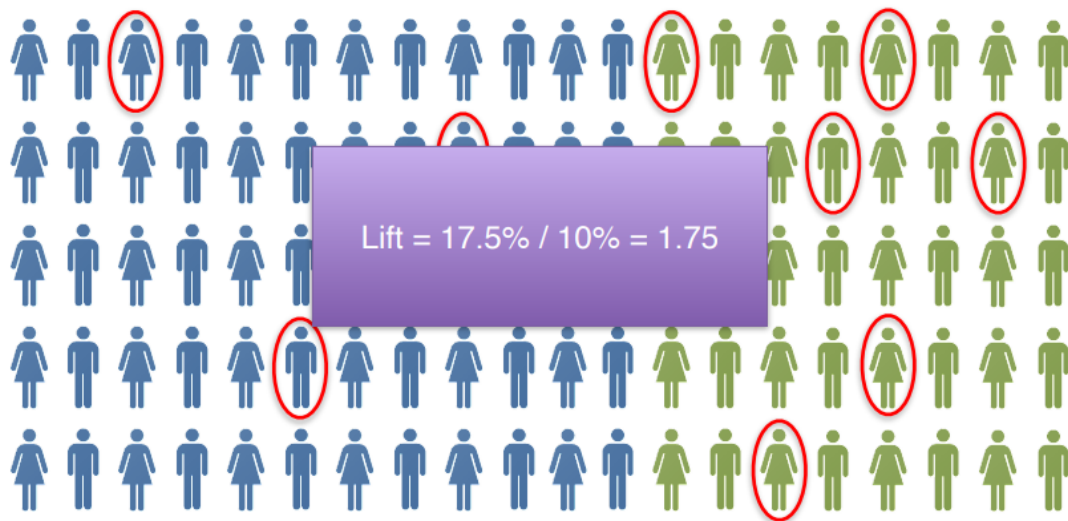
El Lift es una forma de mejorar una respuesta aleatoria

Es como recomendar a 100 personas la película de Batman, la probabilidad de que le guste es de 10%, que son los redondeados de rojo

Y habíamos visto que los de verde eran los que les gustaba el Joker, de ellos, el 17% veía luego Batman

Hay una subida en el porcentaje para los que vieron el Joker y eso es lo que mide el Lift

Entonces el cálculo sería la probabilidad de que le guste si vio el Joker dividido la probabilidad de que le guste sin verla



Los pasos para realizar las recomendaciones utilizando Apriori

Paso 1: Decidir un soporte y nivel de confianza mínimo



Paso 2: Elegir todos los subconjuntos de transacciones con soporte superior que el mínimo elegido



Paso 3: Elegir todas las reglas de estos subconjuntos con nivel de confianza superior al mínimo elegido



Paso 4: Ordenar todas las reglas anteriores por lift descendiente