

Reglas de Asociación: (Análisis de Canasta)

→ Sirve para identificar asociaciones entre items (productos).

• Aplicaciones:

- Sistemas de recomendación de películas (Netflix, Prime, HBO+)
- Sistemas de recomendaciones de comidas (Pappi, Uber Eats).

• Métricas de Asociación: (Son 3)

- Dataset

Factura	Artículo 1	Artículo 2	Artículo 3
1	Manzana	Cerveza	Huevos
2	Piña	Manzana	
3	Cerveza	Manzana	
4	Ketchup	Piña	

1.) Soporte: (Métrica de Asociación No. 1)

- Frecuencia relativa de una regla respecto al total de las transacciones. Mide la popularidad de un item.

Ej: Soporte de q se compre una manzana es de $\frac{3}{4}$. i.e. en 3 de las cuatro facturas se compro manzana.

Ej: Soporte de q se compre la manzana y la cerveza es de $\frac{2}{4}$. i.e. en 2 de las cuatro facturas se compro manzana y Cerveza.

2.) Confianza: (Métrica de Asociación No.2)

Indica cuán probable el ítem b sea comprado cuando el ítem a es comprado. Mide q tan confiable es la suposición hecha por la regla.

Ej: La confianza de q la cerveza sea comprada dado q se compro la manzana es la siguiente:

$$\Rightarrow \text{Confianza (manzana} \leftarrow \text{Cerveza)} = \frac{\text{Soporte (manzana y Cerveza)}}{\text{Soporte (manzana)}}$$

$$\Rightarrow \text{Confianza (b | A)} = \frac{\text{Soporte (a \cap b)}}{\text{Soporte (a)}}$$

$$\Rightarrow \text{Confianza (manzana} \leftarrow \text{Cerveza)} = \frac{2/4}{3/4} = 2/3 = 0.75$$

3.) Sustento (lift): (Métrica de Asociación No.3)

Indica la probabilidad de q el ítem y sea comprado cuando el ítem x es comprado, controlando cuán popular es el ítem y. El control de y es importante ya que si el ítem y es muy popular que aparece en todas las facturas, es muy probable que se asocie con el ítem x.

$$\Rightarrow \text{Sustento (y | x)} = \frac{\text{Soporte (y \cap x)}}{\text{Soporte (y)} * \text{Soporte (x)}}$$

$$\Rightarrow \text{Sustento (manzana | Cerveza)} = \frac{\text{Soporte (manzana y Cerveza)}}{\text{Soporte (manzana)} * \text{Soporte (Cerveza)}}$$

$$\Rightarrow \text{Sustento (manzana | Cerveza)} = \frac{\text{Support}(M \text{ y } C)}{\text{Support}(M) \cdot \text{Support}(C)} = \frac{2/4}{3/4 \cdot 2/4}$$

$$\Rightarrow \text{Sustento \& lift} = 1.33$$

Parámetros para Sustento ó lift:

- Sustento = 1 \Rightarrow No hay Asociación
- Sustento < 1 \Rightarrow poca probabilidad de comprar y si se compra X.
- Sustento > 1 \Rightarrow Alta probabilidad de comprar y si se compra X.

Ver ejemplo en Python