



REGLAS DE ASOCIACIÓN

Equipo 10:

Andrea Isabel Cortez Sánchez - 1736472

Sergio Oviedo Martínez - 1750532

Andrés Hernandez de Luna - 1757180

Johan Antonio Sustaita Martinez - 1723476

Ejemplo: Si + el cliente compra leche + entonces + también comprara café



REGLAS DE ASOCIACIÓN

SI + ANTECEDENTES + ➔ + CONSECUENCIA

Las Reglas de asociación nos permiten contar las combinaciones de artículo o ítems que ocurren con mayor frecuencia en una base de datos.

¿Cuál es el objetivo?

- Dado un conjunto de transacciones T , el objetivo de la minería de reglas de asociación es encontrar todas las reglas teniendo:
 1. Umbral mínimo de soporte
 2. Umbral mínimo de confianza



Ejemplo común: Amazon



Pasa el mouse encima de la imagen para aplicar zoom

LonEasy Auriculares para Juegos, Auriculares para Juegos con Aislamiento de Ruido con Cable de, Control del Volumen audifonos Gaming

Marca: LonEasy



3,002 calificaciones | 169 preguntas respondidas

Amazon's Choice de "audifonos"

Precio: **\$499.00**

Solicita tu tarjeta Amazon Recargable y obtén \$100 de descuento en tu primera compra mayor a \$500

Marca LonEasy

Tipo de conector Alámbrico

Factor de forma de los auriculares On Ear

Tipo de conector USB, Jack de 3,5 mm

Control de ruido Isolation acoustique

Esto es el ejemplo mas común Amazon cuando vas a comprar un articulo muestra un apartado llamado “Comprados juntos habitualmente”, pues su base de datos ah observado que los clientes que muestran interés por cierto producto también se interesan por otro similar o complementario.

Comprados juntos habitualmente



Precio total: **\$1,007.00**

Agregar los tres al carrito



Estos productos los envían y venden distintos vendedores. [Mostrar detalles](#)



También Permite medir la fuerza e importancia de estas combinaciones o relaciones dentro de los artículos en la asociación.

Ejemplo: Cuando se desea comprar un articulo, el sistema internamente asocia mi intención de compra con la de otros clientes en la base de datos.

Base de datos Transaccional



Una Lista.

Una
Representacion
Vertical.

Una
Representacion
Horizontal

Lista de Items

	A	B	C	D
1				
2	tomates	lechuga	mostaza	jamon
3	tomates	pepinos	salad dressing	
4	agua	periodico		
5	agua	coca-cola		
6				
7				

LISTA

Cada Fila es una transacción, por lo que cada fila puede tener un numero diferente de columnas

Representación Vertical

- Es la forma mas eficiente de guardar los datos de tamaño industrial o comercial. Ocupa solo 2 columnas:
- *1 columna indica el # o id de la transacción.
- *2 columna indica un ítem presente (artículo).

T_id	Item
1	Tomates
1	lechuga
1	Mostaza
1	Jamon
2	Tomates
2	Pepinos
2	Queso
3	Agua
3	Periodico
4	Agua
4	Coca cola

Representación Horizontal

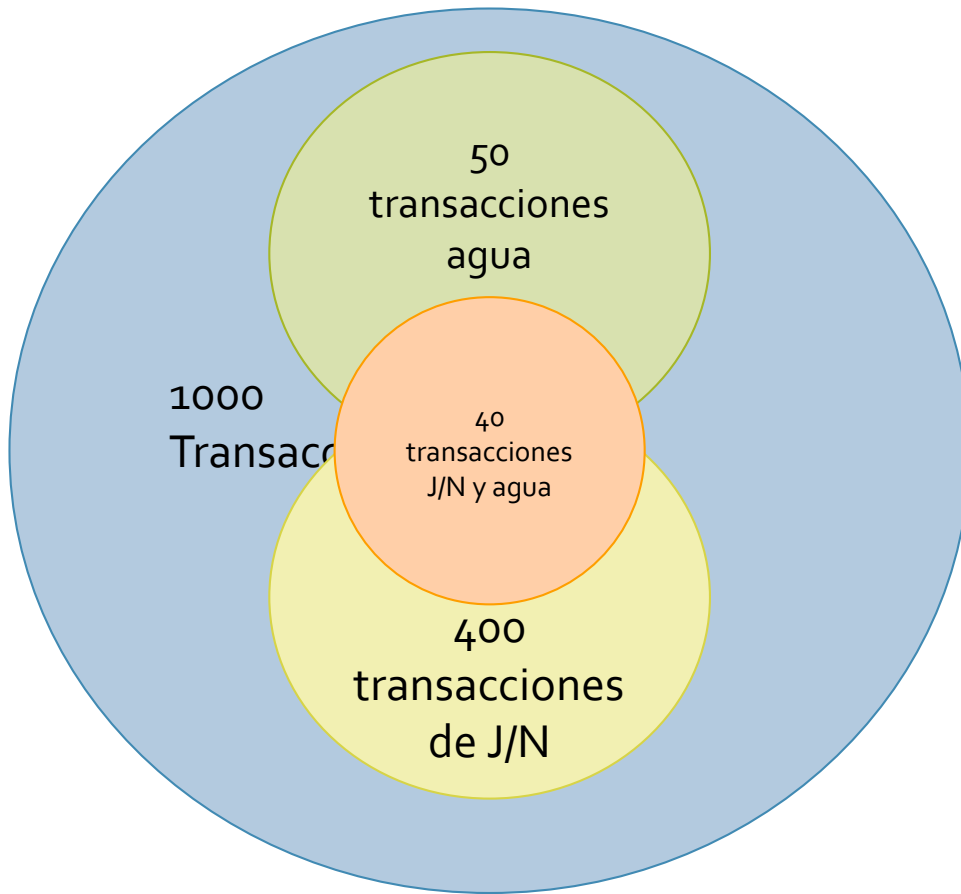
items T_id	Tomate	Lechuga	Mostaza	Jamon	Pepinos	Queso	Agua	Periodico	Coca-Cola
1	1	1	1	1	0	0	0	0	0
2	1	0	0	0	1	1	0	0	0
3	0	0	0	0	0	0	1	1	0
4	0	0	0	0	0	0	1	0	1

- Se representa como una matriz binaria. Cada fila de la matriz representa una transacción y cada columna representa un artículo o ítem.
- Item presente → 1
- Item Ausente → 0

- Agrawal en 1994 descubre un método eficiente de hallar reglas, pues uno de los grandes problemas cuando queremos procesar reglas de asociación es que podemos encontrar muchas reglas.
- **Soporte:** dado una regla "Si $A \rightarrow B$ ", el soporte de esta regla se define como el número de veces o la frecuencia relativa con que A y B aparecen juntos en una B.D de transacciones.
- **Confianza:** dada una regla "Si $A \rightarrow B$ ", la confianza de esta regla es el cociente de soporte de la regla y el soporte del antecedente solamente. Mide la fortaleza de la regla.
$$\text{Confianza}(A \rightarrow B) = \text{Soporte}(A \rightarrow B) / \text{Soporte}(A)$$
- **Lift:** refleja el aumento de la probabilidad de que ocurra el consecuente, cuando no enteramos de que ocurrió el antecedente.

Métricas de interés

Soporte



$$* \text{Soporte (J/N)} = 400/1000 = 0.4 = 40\%$$

$$* \text{Soporte (A)} = 50/1000 = 0.05 = 5\%$$

$$* \text{Soporte (J/N y Agua)} = 40/1000 = 0.04 = 4\%$$

Confianza

$$* \text{Confianza de Agua} \rightarrow \text{J/N}$$

$$= \text{Soporte(Agua} \rightarrow \text{J/N)} / \text{Soporte(Agua)} = 40/50 = 0.8 = 80\%$$

$$* \text{Confianza de J/N} \rightarrow \text{Agua}$$

$$= \text{Soporte(J/N} \rightarrow \text{Agua)} / \text{Soporte(J/N)} = 40/400 = 0.1 = 10\%$$

LIFT

	COMPRO PAN	NO COMPRO PAN	
COMPRO JUGO DE NARANJA	280	120	400
NO COMPRO JUGO DE NARANJA	420	180	600
	700	300	1000

$$\text{Lift}(A \rightarrow B) = \text{Soporte}(A \rightarrow B) / \{\text{Soporte}(A) * \text{Soporte}(B)\}$$

*Lift = 1 o cercana a 1, indica que la relación es producto del azar.

*Lift > 1 indica que la relación es realmente fuerte (controlada por la frecuencia con la que ambos ocurren) (Complementos)

*Lift < 1 Indica una relación débil. Probablemente uno es sustituto de otro.

$$*\text{Soporte}(\text{Pan}) = 700/1000 = 0.70$$

$$*\text{Soporte}(\text{J/N}) = 400/1000 = 0.4$$

$$*\text{Soporte}(\text{Pan} \rightarrow \text{J/N}) = 280/1000 = 0.28$$

$$*\text{Soporte}(\text{J/N} \rightarrow \text{Pan}) = 280/1000 = 0.28$$

$$*\text{Confianza}(\text{Pan} \rightarrow \text{J/N}) = 0.28/0.70 = 0.4$$

$$*\text{Confianza}(\text{J/N} \rightarrow \text{Pan}) = 0.28/0.4 = 0.7$$

$$\text{Lift}(\text{Pan} \rightarrow \text{J/N}) = \text{Lift}(\text{J/N} \rightarrow \text{Pan}) = 0.28 / (0.7 * 0.4) = 1$$



Referencias

- <https://jaywrkr.medium.com/miner%C3%ADa-de-datos-3-f75d15f90c46>
- <https://www.youtube.com/watch?v=YRhu6yEseh8>
- <https://www.youtube.com/watch?v=CcdgSZ2pdW4>
- <https://www.amazon.com.mx/>