



## APRIORI ALGORITHM

- Es un algoritmo que se usa en Reglas de Asociacion
- Cada row del dataset se lo considera una “transaction”

Procedimiento:

- Identifica de manera individual el item mas frecuente
- A partir de ahi lo asocia con items menos frecuentes creando reglas.



Ejemplo:

Set de Transacciones

{A,B,C,D}

{A,B,D}

{A,B}

{B,C,D}

{B,C}

{C,D}

{B,D}

Item	Support
{A}	3
{B}	6
{C}	4
{D}	5

Se puede setear el “Support Therseshold” para determinar cual es frecuente.

Si por ejemplo, el Support Therseshold = 3. Todos los items son frecuentes, si es 5, solo lo son B y D.



Item	Support
{A,B}	3
{A,C}	1
{A,D}	2
{B,C}	3
{B,D}	4
{C,D}	3

$\{A,B\}, \{B,C\}, \{B,D\}, \{B,D\}$

Son frecuentes si el Support Thereshold es de 3

No es frecuente  $\{A,C\}$ , por ejemplo



CONFIDENCE:

$X \rightarrow Y$  Es la proporción en la transacción que contiene  $X \rightarrow Y$

transaction ID	milk	bread	butter	beer	diapers
1	1	1	0	0	0
2	0	0	1	0	0
3	0	0	0	1	1
4	1	1	1	0	0
5	0	1	0	0	0

$\{\text{bread}, \text{butter}\} \rightarrow \{\text{milk}\}$

Confidence =  $\text{supp}(X \cup Y) / \text{supp}(X)$

$\text{supp}(X \cup Y) = \text{supp } X \text{ y } Y$

Confidence =  $0.2/0.2 = 1$

LIFT:  $\text{lift}(X \Rightarrow Y) = \frac{\text{supp}(X \cup Y)}{\text{supp}(X) \times \text{supp}(Y)}$

Si una regla tiene Lift 1 implica que la probabilidad de ocurrencia del lhs y rhs son independientes. Entonces, no hay regla.