Minería de Datos

**Reglas de Asociación**

**Reglas de Asociación**

El algoritmo A-priori nos permite encontrar reglas de asociación de manera automática desde los datos.

* Itemset: colección de uno o más ítems.
* Soporte: numero de veces que aparece un Itemset en la base de datos.
* Itemset frecuente: Soporte > Umbral

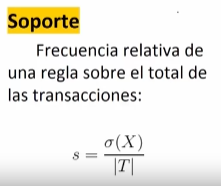
**Soporte, Confianza y Lift**

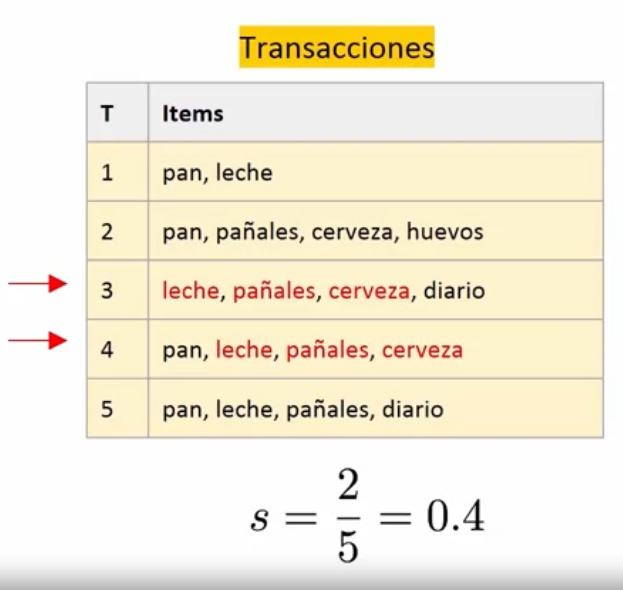
Indicadores de rendimiento en las regla de asociación para evaluar el grado de credibilidad de cualquier decisión que dependen de las reglas encontradas



**{Antecedentes} -> {Consecuentes}**

* **Soporte**: Frecuencia relativa del ItemSet dentro de una base de datos
* **Soporte de una regla:** Corresponde a la frecuencia relativa de un ItemSet que contiene todos los ítems presentes en la regla

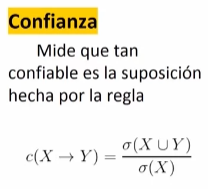


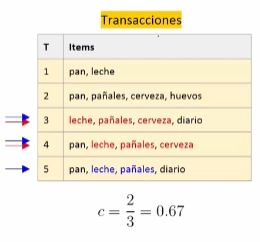


Ejemplo:

40% de las transacciones indican que leche, pañales y cerveza se compraron juntos

* **Confianza**: Probabilidad empírica de que ocurra el consecuente dado que ocurrió el antecedente de la regla



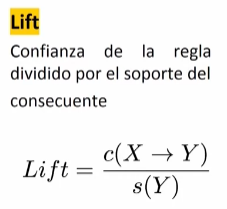


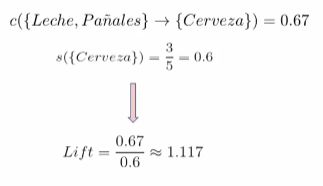
Ejemplo:

Confianza de la regla = 67%.

67% de los consumidores que compraron leche y pañales también compraron cerveza

* **Lift**: Refleja el aumento de la probabilidad de que ocurra el consecuente cuando nos enteramos de que ocurre el antecedente





s({# apariciones de cerveza / # total de transacciones})

La probabilidad de comprar cerveza aumenta de 60% a 67% cuando sabemos que el cliente compra leche y pañales.

|  |  |
| --- | --- |
| LIFT | |
| > 1 | La probabilidad del consecuente de la regla aumento una vez que sabemos que el consumidor compro los ítems del antecedente. |
| = 1 | La Probabilidad no se vio afectada, por lo tanto, el antecedente no aporta nada respecto a la compra u ocurrencia del consecuente. |
| < 1 | El antecedente tuvo un efecto negativo en la ocurrencia del consecuente, haciendo que su probabilidad baje. |

**Algoritmo A-Priori**

Algoritmo más común para encontrar reglas de asociación que cumplan con un mínimo valor de soporte y confianza

Efknonipnfpinre

Lsmerkvneknvrker

Kmwervnw4rpnvre