

Definições

Busca a relação entre itens

Aplicações:

- Cesta de compras
- Detecção de intrusão
- Analise de avaliações de consumidores
- Diagnóstico médico

Conceitos



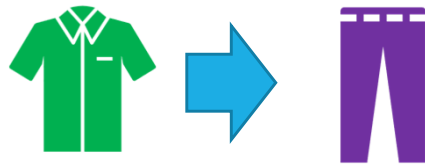
Suporte: numero de transações que contem todos os itens da transação.



Confiança: indica a proporção de vezes que, em uma transação contendo o elemento A, também tem B.



Força da regra (lift) mostra se a associação entre os itens é positiva ou negativa



Suporte

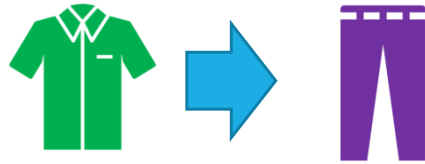
O suporte nos ajuda a entender a frequência com que esses dois itens são encontrados juntos.

$\text{Suporte}(A) = (\text{Número de transações contendo } A) / (\text{Total de transações})$

$\text{Suporte}(A, B) = (\text{Número de transações contendo } A \text{ e } B) / (\text{Total de transações})$

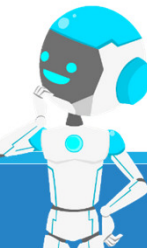


Confiança

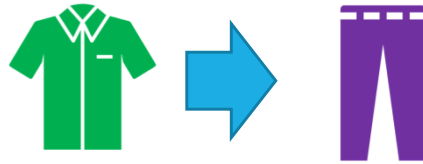


A confiança indica a probabilidade de que havendo item A, também haverá o item B

$$\text{Confiança}(A \rightarrow B) = \text{Suporte}(A, B) / \text{Suporte}(A)$$
















Lift

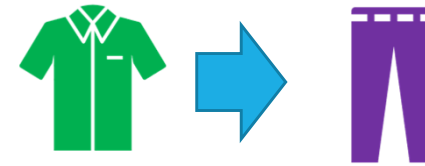


Indica a força de uma regra de associação

- Lift = 1: Indica que A e B são independentes, ou seja, a presença de A não afeta a probabilidade de B ser comprado.
 - Lift > 1: Indica uma associação positiva entre A e B, ou seja, a presença de A aumenta a probabilidade de B ser comprado.
 - Lift < 1: Indica uma associação negativa entre A e B, ou seja, a presença de A diminui a probabilidade de B ser comprado.
-
- $\text{Lift}(A \rightarrow B) = \text{Confiância}(A \rightarrow B) / \text{Suporte}(B)$
















1			
2			
3			
4			
5			
6			

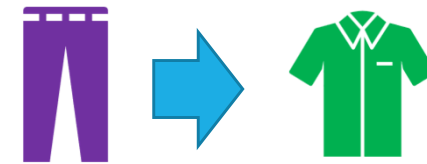


→ Quem compra camisa também compra calça?

Confiança	
Suporte	(Número de transações contendo A e B) / (Total de transações) $= 3/6 = 0,5$
Confiança	$\text{Suporte}(A, B) / \text{Suporte}(A)$ $= 0,5/0,5 = 1$
Lift	$\text{Confiança}(A \rightarrow B) / \text{Suporte}(B)$ $1/1 = 1$



1			
2			
3			
4			
5			
6			
















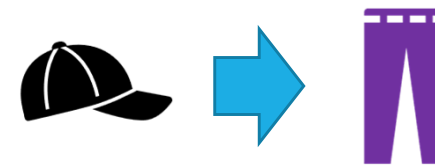
→ Quem compra calça também compra camisa?

Confiança

Suporte	(Número de transações contendo A e B) / (Total de transações) = $3/6 = 0,5$
Confiança	$\text{Suporte}(A, B) / \text{Suporte}(A)$ = $0,5/1 = 0,5$
Lift	$\text{Confiança}(A \rightarrow B) / \text{Suporte}(B)$ = $0,5/0,5 = 1$



1			
2			
3			
4			
5			
6			



→ Quem compra boné também compra calça?

Confiança

Suporte	(Número de transações contendo A e B) / (Total de transações) = $4/6 = 0,67$
Confiança	$\text{Suporte}(A, B) / \text{Suporte}(A)$ = $0,67/0,67 = 1$
Lift	$\text{Confiança}(A \rightarrow B) / \text{Suporte}(B)$ = $1/1 = 1$



Apriori e FP-Grow



Apriori e FP-Grow são os algoritmos mais comuns na mineração de regras de associações.



Apriori é baseado no princípio de que se um conjunto de itens é frequente, um subconjunto destes itens também será frequente. O princípio contrário também é válido



FP-Grow induz árvores, e busca sobreposição destas árvores, onde os itens são frequentes.

