

关联规则 <  $\begin{matrix} \text{Apriori} \\ \text{FP} \end{matrix}$  改善無效率

Itemset (项目集), 以大括号  $\{ \}$  表示

$\Rightarrow A = \{ 2 \}$ , 1 itemset

$\Rightarrow B = \{ 1, 3 \}$ , 2 itemset

$\Rightarrow C = \{ 2, 4, 5 \}$ , 3 itemset

若在 Dataset 出现  $k$  次, support count (支持数量) =  $k$ , 用  $\sigma(A) = k$  表示

$\hookrightarrow$  support count = 3,  $\sigma(B) = 3$

$$\text{支持度 } S(A) = \frac{\sigma(A)}{|D|} = \frac{5}{10} = 50\%$$

support

$$S(B) = \frac{3}{10} = 30\%$$

門檻值: 最低支持介數 = 3

ITD	Items
1	2, 5, 7
2	1, 3, 4, 6
3	2, 6, 7
4	2, 4, 5
5	3, 6
6	2, 4, 6
7	1, 4, 5
8	1, 3, 5
9	2, 3, 5
10	1, 3, 5

G1

1-itemset	support count	1-itemset	support count
{1}	4	{2}	5
{2}	5	{3}	5
{3}	5	{4}	4
{4}	4	{5}	6
{5}	6	{6}	3
{6}	3	{7}	2

G2

2-itemset	support count	2-itemset	support count	2-itemset	support count
1, 2	0	2, 5	3	5, 6	0
1, 3	3	2, 6	2		
1, 4	2	3, 4	1		
1, 5	3	3, 5	3		
1, 6	1	3, 6	2		
2, 3	1	4, 5	2		
2, 4	2	4, 6	2		



L <sub>2</sub>		support	confidence
2-itemset	support count		
{1, 3}	3	d, e, f	1
{1, 5}	3	f, d, e	1
{2, 5}	3	e, f, d	1
{3, 5}	3	d, e	1
L <sub>3</sub>		e, f, d	1
3-itemset	support count	no item fit	
{1, 3, 5}	2 (不符合)	f, e, d	1
{2, 3, 5}	2 (不符合)	f, d	1

找出 frequent itemset (G<sub>1</sub> → L<sub>1</sub> → G<sub>2</sub> → L<sub>2</sub> → G<sub>3</sub> → L<sub>3</sub>)

$$\text{信賴度} = \frac{s(x \cap y)}{s(x)}, \text{ 買 } x \text{ 同時也會買 } y \text{ 的機率}$$

$$ex: \{1, 3\} \Rightarrow \{1\} \rightarrow \{3\} = \frac{30\%}{40\%} = \frac{3}{4} = 75\% \text{ (confidence)}$$

$$\{3\} \rightarrow \{1\} = \frac{30\%}{50\%} = \frac{3}{5} = 60\% \text{ (confidence)}$$

\*  $X \rightarrow Y [ \Delta, \square ]$ ,  $\Delta$  支持度 support

$\square$  信賴度 confidence

x, y 為 2 項 items

$$\{2, 5\} \Rightarrow \{2\} \rightarrow \{5\} [ 30\%, 60\% ]$$

$$2, 5 \text{ 交集 } \frac{30\%}{50\%} = \frac{3}{5} = 60\%$$

$$\{5\} \rightarrow \{2\} [ 30\%, 50\% ]$$

$$\frac{30\%}{60\%} = \frac{1}{2} = 50\%$$

$$\{3, 5\} \Rightarrow \{3\} \rightarrow \{5\} [ 30\%, 60\% ]$$

$$\{5\} \rightarrow \{3\} [ 30\%, 50\% ]$$

提升度 =  $\frac{S(X \cap Y)}{S(X)S(Y)}$  , 類似於相關係數,  $<$  大於1, 正相關  
lift 小於1, 負相關