

HW03 Apriori Algorithms

C109152304 許智程

下表為某商店的五筆交易紀錄，試回答下列問題。

設最小支持度(min_support)=2，也就是出現次數以下時，不列入選擇

交易紀錄	商品項目-代碼
601	麵包-A、果醬-B、花生醬-C
602	麵包-A、花生醬-C
603	麵包-A、花生醬-C、牛奶-D
604	麵包-A、啤酒-E
605	牛奶-D、啤酒-E

1. 利用Apriori演算法求所有2-項目集的可能規則

itemset	sup	->	A	4	B	1	C	3	D	2	E	2
A	4											
B	1											
C	3											
D	2											
E	2											

itemset	sup	->
{A,C}	3	
{A,D}	1	
{A,E}	1	
{C,D}	1	
{C,E}	0	
{D,E}	1	

itemset	sup
{A,C}	3

∴ 所有2-項目集的可能規則為 {A,C}

2. 試求所有2-項目極可能規則的支持度(Confidence)與信賴度(Support)

itemset	sup
{A,C}	3

支持度(Confidence)與信賴度(Support)的算法如下

$$\begin{cases} Support\{X \Rightarrow Y\} = P(X \cap Y) \\ Confidence\{X \Rightarrow Y\} = P(Y|X) = \frac{P(X \cap Y)}{P(X)} \end{cases}$$

套用上面的式子，得到

$$\begin{cases} Support\{A \Rightarrow C\} = P(A \cap C) = \frac{3}{5} \\ Confidence\{A \Rightarrow C\} = P(C|A) = \frac{P(A \cap C)}{P(A)} = \frac{\frac{3}{5}}{\frac{4}{5}} = \frac{15}{20} = \frac{3}{4} \end{cases}$$

3. 令支持度門檻為20%，信賴度門檻為20%，試求被列入候選項目集的規則
 $20\% = \frac{20}{100} = 0.2$

$$\therefore \{A, C\} \begin{cases} \frac{3}{5} = \frac{6}{10} > \frac{2}{10} \\ \frac{3}{4} = \frac{15}{20} > \frac{2}{10} = \frac{4}{20} \end{cases}$$

\therefore 被選入的是 $\{A, C\}$

4. 利用Apriori演算法求所有3-項目集的可能規則，求算其支持度(Confidence)與信賴度(Support)，最後列出被選入候選項目集的規則

itemset	sup
{A,C}	3

->

itemset	sup
{A,C,D}	1
{A,C,E}	0

\therefore 沒有任何規則被選入候選項目