

Part A:

1) Team Jarlsberg

Siddhesh Madeshwar

Zachary Noel

Erin Dolson

2) Given a database of transactions and a min-support = 2

a) An example of an association rule that matches min-suport = 2 and min-conf = 70.

(I1, I2, I3, I4, IX -> IY)

If IX = I6 and IY = I7,

$$s = \text{sup}(I1, I2, I3, I4, I6, I7) = 2$$

$$c = \text{conf}(X \rightarrow Y) = \text{sup}(XY) / \text{sup}(X) = \text{sup}(I1, I2, I3, I4, I6, I7) / \text{sup}(I1, I2, I3, I4, I6) \\ = 2/2 = 1.$$

Therefore (I1, I2, I3, I4, I6 -> I7) is an acceptable rule.

b) For the association rule I1- > I6, compute the support, confidence, lift, and conviction.

$$\text{sup}(I1I6) = 2$$

$$\text{conf}(I1I6) = \text{sup}(I1I6) / \text{sup}(I1) = 2/3$$

$$\text{lift}(I1I6) = \text{conf}(I1I6) / \text{rsup}(I6) = (2/3) / (\text{sup}(I6)/5) = (2/3) / (2/5) = 2/5$$

$$\text{conviction}(I1I6) = (1 - \text{rsup}(I6)) / (1 - \text{conf}(xy)) = (1 - 2/5) / (1 - 2/3) = 3/5 / 1/3 = 6/5$$

3) Association Analysis II

1st scan

$C_1 =$

Itemset	sup
{M}	3
{O}	4
{N}	2
{K}	5
{E}	4
{Y}	3
{D}	1
{A}	1
{U}	1
{C}	2
{I}	1

$L_1 =$

Itemset	sup
{M}	3
{O}	4
{K}	5
{E}	4
{Y}	3

2nd scan

$C_2 =$

Itemset	sup
{M,O}	1
{M,Y}	2
{M, K}	3
{M, E}	2
{O,K}	3
{O,E}	3
{O,Y}	2
{K,E}	4
{K,Y}	3
{Y,E}	2

$L_2 =$

Itemset	sup
{M, K}	3
{O,K}	3
{O,E}	3
{K,E}	4
{K,Y}	3

3rd Scan

$C_3 =$

Itemset	sup
{O,K,E}	3

$L_3 =$

Itemset	sup
{O,K,E}	3