

## Homework 2 Solutions

1) Read Chapter 6 (only sections 6.1 and 6.7).

2) Do Chapter 6 textbook problem #2 (parts a,b,c,d only) on page 404.

a)

$$\begin{aligned}s(\{e\}) &= \frac{8}{10} = 0.8 \\s(\{b, d\}) &= \frac{2}{10} = 0.2 \\s(\{b, d, e\}) &= \frac{2}{10} = 0.2\end{aligned}$$

b)

$$\begin{aligned}c(bd \longrightarrow e) &= \frac{0.2}{0.2} = 100\% \\c(e \longrightarrow bd) &= \frac{0.2}{0.8} = 25\%\end{aligned}$$

No, confidence is not a symmetric measure.

c)

$$\begin{aligned}s(\{e\}) &= \frac{4}{5} = 0.8 \\s(\{b, d\}) &= \frac{5}{5} = 1 \\s(\{b, d, e\}) &= \frac{4}{5} = 0.8\end{aligned}$$

d)

$$\begin{aligned}c(bd \longrightarrow e) &= \frac{0.8}{1} = 80\% \\c(e \longrightarrow bd) &= \frac{0.8}{0.8} = 100\%\end{aligned}$$

3) Do Chapter 6 textbook problem #6 (parts d,e only) on page 406.

d)  $s(\{\text{Bread}, \text{Butter}\}) = 5/10$  is the largest support using for an itemset of size 2 or larger.

e) from the ICE we saw  $c(\{\text{Bread}\} \rightarrow \{\text{Butter}\}) = 1 = c(\{\text{Butter}\} \rightarrow \{\text{Bread}\})$

also, the following all work

$c(\{\text{Milk}\} \rightarrow \{\text{Bread}\}) = .6 = c(\{\text{Bread}\} \rightarrow \{\text{Milk}\})$

$c(\{\text{Milk}\} \rightarrow \{\text{Butter}\}) = .6 = c(\{\text{Butter}\} \rightarrow \{\text{Milk}\})$

$c(\{\text{Beer}\} \rightarrow \{\text{Cookies}\}) = .5 = c(\{\text{Cookies}\} \rightarrow \{\text{Beer}\})$

$c(\{\text{Beer}\} \rightarrow \{\text{Bread}\}) = 0 = c(\{\text{Bread}\} \rightarrow \{\text{Beer}\})$

$c(\{\text{Beer}\} \rightarrow \{\text{Butter}\}) = 0 = c(\{\text{Butter}\} \rightarrow \{\text{Beer}\})$

4)

$s(\text{ip}=65.57.245.11 \rightarrow \text{Mozilla/5.0 (X11; U; Linux i686 (x86_64); en-US; rv:1.8.1.3) Gecko/20070309 Firefox/2.0.0.3}) = 1385/14803 = 0.094$

$c(\text{ip}=65.57.245.11 \rightarrow \text{Mozilla/5.0 (X11; U; Linux i686 (x86_64); en-US; rv:1.8.1.3) Gecko/20070309 Firefox/2.0.0.3}) = 1385/5018 = 0.276$

9.4% of all the rows list that particular ip address and that particular user agent.

27.6% of the rows which have that particular ip address also have that particular user agent.