Fundamentals of Data Science

Semester B 20-21

Tutorial 9

1a. $s(\{e\}) = \frac{8}{10} = 0.8$

$s(\{b,d\}) = \frac{2}{10} = 0.2$

$s(\{b,d,e\}) = \frac{2}{10} = 0.2$

b. $c(\{b,d\} \rightarrow \{e\}) = \frac{\sigma(\{b,d,e\})}{\sigma(\{b,d\})} = \frac{s(\{b,d,e\})}{s(\{b,d\})} = \frac{0.2}{0.2} = 1$

$c(\{e\} \rightarrow \{b,d\}) = \frac{\sigma(\{b,d,e\})}{\sigma(\{e\})} = \frac{s(\{b,d,e\})}{s(\{e\})} = \frac{0.2}{0.8} = 0.25$

c. $s(\{e\}) = \frac{4}{5} = 0.8$

$s(\{b,d\}) = \frac{5}{5} = 1$

$s(\{b,d,e\}) = \frac{4}{5} = 0.8$

d. $c(\{b,d\} \rightarrow \{e\}) = \frac{\sigma(\{b,d,e\})}{\sigma(\{b,d\})} = \frac{s(\{b,d,e\})}{s(\{b,d\})} = \frac{0.8}{1} = 0.8$

$c(\{e\} \rightarrow \{b,d\}) = \frac{\sigma(\{b,d,e\})}{\sigma(\{e\})} = \frac{s(\{b,d,e\})}{s(\{e\})} = \frac{0.8}{0.8} = 1$

2. $c_1 = \frac{\sigma(\{p,q\})}{\sigma(\{p\})}$

$c_2 = \frac{\sigma(\{p,q,r\})}{\sigma(\{p\})}$

$c_3 = \frac{\sigma(\{p,q,r\})}{\sigma(\{p,r\})}$

Since $\sigma(\{p,q\}) \geq \sigma(\{p,q,r\})$ and $\sigma(\{p\}) \geq \sigma(\{p,r\})$, $c_1 \geq c_2$ and $c_3 \geq c_2$

Since c_1, c_2 and c_3 all have different values, $c_1 > c_2$ and $c_3 > c_2$

As a result, c_2 is the lowest confidence value.