**Cloud Computing for Data Analysis**

**Exercise 06 : Association Rule Mining**

Consider the data set shown in Table 6.1.

(a) Compute the support for itemsets *{e}*, *{b, d}*, and *{b, d, e}* by treating

each transaction ID as a market basket.

* *{e}*
  + *Support: 8/10 = 80%*
* *{b, d}*
  + *Support: 2/10 = 20%*
* *{b, d, e}*
  + *Support: 2/10 = 20%*

(b) Use the results in part (a) to compute the confidence for the association

rules *{b, d} → {e}* and *{e} → {b, d}*. Is confidence a symmetric

measure?

* *{b, d} → {e}*
  + *Confidence: 2/2 = 100%*
* *{e} → {b, d}*
  + *Confidence: 2/8 = 25%*

No, confidence is not a symmetric measure.

**Table 6.1.** Example of market basket transactions.

|  |  |  |
| --- | --- | --- |
| Customer ID | Transaction ID | Items Bought |
| 1 | 1 | {a, d, e} |
| 1 | 24 | {a, b, c,e} |
| 2 | 12 | {a, b, d,e} |
| 2 | 31 | {a, c, d,e} |
| 3 | 15 | {b, c, e} |
| 3 | 22 | {b, d, e} |
| 4 | 29 | {c, d} |
| 4 | 40 | {a, b, c} |
| 5 | 33 | {a, d, e} |
| 5 | 38 | {a, b, e} |