**Cloud Computing for Data Analysis**

**Exercise 07 : Association Rules using Apriori algorithm**

6.Consider the market basket transactions shown in Table 6.2.

(a) What is the maximum number of association rules that can be extracted

from this data (including rules that have zero support)?

* R = {3 ^ d} – {2 ^ (d + 1)} + 1
* In the question, there are 6 unique items, i.e. d = 6
* R = {3 ^ 6} – {2 ^ 7} + 1 = 602

(b) What is the maximum size of frequent itemsets that can be extracted

(assuming *minsup >* 0)?

* The transaction id 6 and 9 are the longest transactions, containing 4 items each. The maximum size of frequent itemsets that can be extracted is 4.

(c) Write an expression for the maximum number of size-3 itemsets that

can be derived from this data set.

* C(6, 3) = 20

(d) Find an itemset (of size 2 or larger) that has the largest support.

* { Bread, Butter } => Support: 5

(e) Find a pair of items, a and b, such that the rules {a} → {b} and {b} → {a} have the same confidence.

* { Bread, Butter }
  + Confidence(Bread → Butter) = 5 / 5
  + Confidence(Butter → Bread) = 5 / 5
* { Beer, Cookies }
  + Confidence(Beer → Cookies) = 2 / 4
  + Confidence(Cookies → Beer) = 2 / 4

|  |  |
| --- | --- |
| Transaction ID | Items Bought |
| 1 | {Milk, Beer, Diapers} |
| 2 | {Bread, Butter, Milk} |
| 3 | {Milk, Diapers, Cookies} |
| 4 | {Bread, Butter, Cookies} |
| 5 | {Beer, Cookies, Diapers} |
| 6 | {Milk, Diapers, Bread, Butter} |
| 7 | {Bread, Butter, Diapers} |
| 8 | {Beer, Diapers} |
| 9 | {Milk, Diapers, Bread, Butter} |
| 10 | {Beer, Cookies} |