Name Group

Intelligent Systems

20 June 2020

2 Association Analysis

Consider the data set shown in the table to be used as scenario in this chapter,

TID	Bread	Butter	Cheese	Milk	Orange Juice	Yogurt
T_1	1	1	1	1	1	1
T_2	1	1	1	0	0	0
T_3	1	1	0	1	1	1
T_4	1	1	1	0	1	0
T_5	1	1	1	0	0	1
T_6	1	0	0	0	1	0
T_7	1	1	1	1	1	1
T_8	0	1	1	1	0	0
T_9	1	1	0	0	1	0
T_{10}	1	1	1	1	1	1

which deals with $X \rightarrow Y|(c, s)$. (That is, at least c fraction of the transactions that contain X also contain Y (confidence condition), and at least a fraction s of all transactions contain both X and Y (support condition). Use $minsup = 20\%$ and $minconf = 50\%$.

1. Explain the Itemset and Support on this scenario with the definition provided in the chapter.
2. Explain fully the Association Rule mining problem on this set.
3. Explain the frequent itemset generation using the brute-force approach for finding frequent itemsets in the lattice structure.
4. Explain and show the Apriori principle on this lattice structure.
5. Explain the support-based pruning. What property is used?
6. Apply the Apriori algorithm on the data set by explaining all the steps.
7. Apply the brute-force method for candidate generation while explaining.
8. Show and explain the $F_{k-1} \times F_1$ method for candidate generation.
9. Show and explain the $F_{k-1} \times F_{k-1}$ method for candidate generation.