Apriori Algorithm

Suhas P
1GA19CS159

Apriori Algorithm

- Apriori algorithm is used for frequent itemset mining and association rule mining.
- The algorithm simply follows a basis that any subset of a large itemset must be a large itemset.
- The Apriori principle can reduce the number of itemsets needed to be examined.
- Apriori principle suggests if an itemset is frequent, then all of its subsets must also be frequent.

Assume X and Y are two itemsets. Apriori principle holds due to the following property of support measure:

$$\forall\ X, Y: (X \subseteq Y) \rightarrow s\,(X) \geq s(Y)$$

Apriori Algorithm

Apriori algorithm evaluates candidates for association as follows:

C_k: Set of candidate-itemsets of size k

F_l: Set of frequent itemsets of size k

F_1 = {large items}

for (k=1; F_k != 0; k++) *do* {

C_{k+1} = New candidates generated from F_k

for each transaction t in the database *do*

Increment the count of all candidates in C_{k+i} that are contained in F_{k+i} = Candidates in C_{k+i} with minimum support }

Apriori Example