

ABSTRACT

Mining association rules is a data mining process to find rule and pattern from a large database. The pattern can be frequent itemset from the transaction of databases. Frequent itemset generation is most time-consuming process, so we need an algorithm that can be efficient a time consuming.

A most popular algorithm is Apriori which use support base pruning to prune a vast amount of non-candidate itemsets. This algorithm has disadvantages when the cardinality of longest itemset is k , apriori needs k passes of database scan, and it has. In addition, the apriori algorithm is computation-intensive in generating the candidate itemsets and counting the support values, especially for applications with very low support threshold and/or a vast amount of items.

Cut Both Ways (CBW) combine a various technic and use cutting level (α) to divide a search space into two different part. Top-down strategy combined with breadth first search and horizontal counting, are used to find frequent itemset at below of the cutting level. In the other hand, bottom-up strategy combined with depth first search and vertical intersection, are used to find frequent itemset at upper of the cutting level. Cutting level is an average cardinality of frequent itemsets, expecting that most of the frequent itemsets will appear in this level.

In this final project will implement frequent itemset generation using Apriori and CBW algorithm. Then, compare its performance by using different parameter of minimum support.

Keywords : mining association rules, itemset, frequent itemset, support, support base pruning, longest frequent itemset, computation-intensive, cutting level, top-down, bottom-up, breadth first search, dept first search, vertical intersection.