**Data mining**

**Assignment 1**

**2011003865 박준형**

* **Apriori algorithm**

1. Make L1 set (L1 is sets that contains one element larger than minimum support)
2. Make set , by using
3. Check each item set of if subsets of the item set are in frequent item set
4. Check each item set of if support of the item set is larger than minimum support
5. After step 4, remains of are
6. Saving to frequent item set, then do 2-5 until would be empty set
7. After step 6, we have frequent item sets of the given transactions.

* **Making association rules**

1. By using apriori algorithm, we have frequent item sets.
2. Every element item set of frequent item sets, separate the sets to two subsets using this rule
3. Making rule that , calculate support and confidence of the rule.

* **Functions that used in codes**

**(details are in comment written in my c++ file)**

* + void find\_frequent\_item\_set(set< set<int> > item\_set);
    - Function that use to find frequent item set. It is trigger function that begins apriori algorithm.
  + bool prunning(set<int> item\_set);
    - Function that used to if all subsets of item\_set(parameter of function) are frequent item set or not.
  + int calculate\_support(set<int> item\_set);
    - Calculate support value of item\_set(parameter of function).
  + void split\_input\_string(char \* \_input\_string);
    - This function is for transforming input text file properly.
  + set< set<int> > cut\_smaller\_than\_support\_set(set< set<int> > item\_set);
    - Every element set in item\_set(parameter of function), calculate support and erase the element set if the support of element set is smaller than minimum set
  + set< set<int> > make\_next\_candidate\_set(set< set<int> > item\_set);
    - Using item\_set like L(x-1) set, find (x)th of candidate set.(find C(x))
  + void make\_rules(set<int> first\_set, set<int> second\_set);
    - This function implement using recursive call.
  + int\* check\_set\_k\_1\_same(set<int> set\_a, set<int> set\_b)
    - This function check if K-1(K = set size) elements of two sets (parameters of function) are same and return integer array that contains the result of checking, two different elements.
* **Instructions for compiling your source codes at TA`s computer**
  + - Using Visual studio, put my source code in a project and compile it(using Ctrl + F7).