2011021764
KangByungWook
201. 3. 30.

Assignment 1
Apriori algorithm

1. Algorithm summary

First, scan all transaction and generate all set of size k and frequency. Then, prune element of which support is under minimal support. After that, join element of the result and generate new list of set of which size is k+1. Then prune element that is not satisfy minimal support. Repeat those process from k=1, until further set is no longer generated. Then, Generate all subsets of each result in set list and make pair of item and associative set by combine subsets of each result. Finally calculate support and confidence of each pair. All the item and associative item set can have higher support than minimal support because their superset's support is also higher than minimal support.

2. File and class structure

A. program.py : file for execution

B. set_module.py : module for set information and operation

- set_stat(_file, _set) : method for getting statistics information in file such as frequency, support and confidence
- associative_set_stat(_file, item_set, associative_item_set) : getting statistics of item set and associative item set in file such as support and confidence. Similar to set_stat.
- join_and_prune(_file, set_list, min_sup, step) : generating list of set that has size of step by joining and pruning on the basis of min_sup in file.
- generate_item_associative(_set) : generate all (item set, associative item set) pair among within parameter set. This is used when the final result of joining and pruning process is handled.
- generate_all_subsets(_set) : generate all subset of parameter set.

3. Execution -> **python program.py 5 input.txt output.txt**

```
(untitled) KangByungui-MacBook-Pro:Programming_Assignment_1 byungwook$ python program.py 5 input.txt output.txt
1 step proceeding...
2 step proceeding...
3 step proceeding...
4 step proceeding...
done
```

4. Output

{16}	{3}	25.2	59.4
{3}	{16}	25.2	84.0
{16}	{7}	9.8	23.1
{7}	{16}	9.8	40.8
{16}	{8}	30.2	71.2
{8}	{16}	30.2	66.8
{3}	{7}	7.4	24.7
{7}	{3}	7.4	30.8
{8}	{3}	25.8	57.1
{3}	{8}	25.8	86.0
{8}	{7}	11.6	25.7
{7}	{8}	11.6	48.3
{16}	{3,7}	6.4	15.1
{3}	{16,7}	6.4	21.3
{7}	{16,3}	6.4	26.7
{16,3}	{7}	6.4	25.4
{16,7}	{3}	6.4	65.3
{3,7}	{16}	6.4	86.5
{16}	{8,3}	24.0	56.6
{8}	{16,3}	24.0	53.1
{3}	{16,8}	24.0	80.0
{16,8}	{3}	24.0	79.5
{16,3}	{8}	24.0	95.2
{8,3}	{16}	24.0	93.0
{16}	{8,7}	7.4	17.5
{8}	{16,7}	7.4	16.4
{7}	{16,8}	7.4	30.8
{16,8}	{7}	7.4	24.5
{16,7}	{8}	7.4	75.5
{8,7}	{16}	7.4	63.8