Imaduddin Sheikh

# CS 422 - Homework 6

## 1. Exercises

### 1.1 Tan, Ch. 5 (Association Analysis)

**Q.15)**

**a)** e
**b)** d
**c)** e
**d)** b
**e)** e

### 1.2 Zaki, Chapter 8 (Frequent Pattern Mining)

**Q.1)**

**a)** In Apriori Algorithm Scanning, 1-itemset will be generated and scanned from the transaction table/data set.

| Itemset | Sup. count |
|---|---|
| {A} | 5 |
| {B} | 4 |
| {C} | 5 |
| {D} | 6 |
| {E} | 1 |
| {F} | 4 |
| {G} | 5 |

Then the 1-itemset will be pruned by the minsup value.

| Itemset | Sup. count |
|---|---|
| {A} | 5 |
| {B} | 4 |
| {C} | 5 |
| {D} | 6 |
| {F} | 4 |
| {G} | 5 |

The pruned 1-itemset will be joined with itself to form a 2-itemset.

| Itemset |
|---|
| {A,B} |
| {A,C} |
| {A,D} |
| {A,F} |
| {A,G} |
| {B,C} |
| {B,D} |
| {B,F} |
| {B,G} |
| {C,D} |
| {C,F} |
| {C,G} |
| {D,F} |
| {D,G} |
| {F,G} |

Database is scanned to determine frequent 2-item sets.

| Itemset | Sup. count |
|---|---|
| {A,B} | 3 |
| {A,C} | 3 |
| {A,D} | 4 |
| {A,F} | 2 |
| {A,G} | 2 |
| {B,C} | 2 |
| {B,D} | 2 |
| {B,F} | 1 |
| {B,G} | 2 |
| {C,D} | 4 |
| {C,F} | 2 |
| {C,G} | 3 |
| {D,F} | 4 |
| {D,G} | 3 |
| {F,G} | 2 |

The 2-itemset data set will be pruned with respect to the minimum support.

| Itemset | Sup. count |
|---|---|
| {A,B} | 3 |
| {A,C} | 3 |
| {A,D} | 4 |
| {C,D} | 4 |
| {C,G} | 3 |
| {D,F} | 4 |
| {D,G} | 3 |

The pruned 2-itemset will be joined with itself to form a 3-itemset.

| Itemset |
|---|
| {A,B,C} |
| {A,B,D} |
| {A,C,D} |
| {A,C,G} |
| {A,D,F} |
| {A,D,G} |
| {C,D,G} |
| {C,D,F} |
| {D,G,F} |

Database scanned again for frequent 3-itemsets.

| Itemset | Sup. count |
|---|---|
| {A,B,C} | 1 |
| {A,B,D} | 2 |
| {A,C,D} | 3 |
| {A,C,G} | 1 |
| {A,D,F} | 2 |
| {A,D,G} | 1 |
| {C,D,G} | 2 |
| {C,D,F} | 2 |
| {D,G,F} | 2 |

The 3-itemsets will be pruned with respect to minimum support value.

| Itemset | Sup. count |
|---|---|
| {A,C,D} | 3 |

Further 4-itemsets are not possible, since $L_3$ has only one set.

**Since $L_3$ has set {ACD} with support count 3, {ACD} is the frequent item set.**

**Q.4)**

1-item set

| Itemset | Sup. count |
|---|---|
| {A} | 4 |
| {B} | 5 |
| {E} | 4 |

2-item set

| Itemset | Sup. count |
|---|---|
| {A,B} | 3 |
| {A,E} | 2 |
| {B,E} | 4 |

3-item set

| Itemset | Sup. count |
|---|---|
| {A,B,E} | 2 |

**{A} => {B,E}**

Confidence = 2/4 = 0.5

**{B} => {A,E}**

Confidence = 2/5 = 0.4

**{E} => {A,B}**

Confidence = 2/4 = 0.5

**{A,B} => {E}**

Confidence = 2/3 = 0.66

**{A,E} => {B}**

Confidence = 2/2 = 1.0

**{B,E} => {A}**

Confidence = 2/4 = 0.5