Part-2

Problem-2:

Multi-Dimensional Association Rules.

Data Base F = {F1, F2, F3, F4, F5}

F1 = {Bread, Endive}

F2 = {Bread, Milk}

F3 = {Bread, Milk, Endive}

F4 = {Bread, Endive}

F5 = {Endive}

Additional information:

F1: livesin = SBrook, income = high

F2: income = high

F3: livesin = SBrook

F4: livesin = SBrook, income = high

F5: income = high

1. Combine the information to one transactional database.

T = {T1, T2, T3, T4, T5}

Combined Transactional Data is represented as below

	I1 Bread	I2 Milk	I3 Endive	I4 livesin SBrook	I5 income high
T1	+	-	+	+	+
T2	+	+	-	-	+
T3	+	+	+	+	-
T4	+	-	+	+	+
T5	-	-	+	-	+
Sc	4	2	4	3	4

Writing transactions using numbers gives

$T1 = \{1,3,4,5\}$

$T2 = \{1,2,5\}$

$T3 = \{1,2,3,4\}$

$T4 = \{1,3,4,5\}$

$T5 = \{3,5\}$

2. Minimum Support = 3.

∴ I2 cannot be included in Frequent 1-item set

Frequent 1-item set L1:

$L1 = \{\{1\}, \{3\}, \{4\}, \{5\}\}$

Join step:

Joining Lk-1 with itself will generate candidate set Ck.

∴ 2-item Candidate set C2 for the given transactions with support count is

$\{1,3\}(3), \{1,4\}(3), \{1,5\}(3), \{3,4\}(3), \{3,5\}(3), \{4,5\}(2)$

Prune step:

Any (k-1) itemset that is not frequent cannot be a subset of a frequent k-itemset.

All the sets in C2 are valid, so pruning is not required.

{4,5} has 2 occurences and min support is 3. so, it cannot be included in Frequent 2-itemset.

Frequent 2-itemset L2:

$L2 = \{\{1,3\}, \{1,4\}, \{1,5\}, \{3,4\}, \{3,5\}\}$

Join Step:

3-item Candidate set C3 with support count is

{1,3,4}, {1,3,5}, {1,4,5}, {3,4,5}

Prune step:

sets {1,4,5} and {3,4,5} cannot be included for Frequent 3-itemset because all their corresponding possible subsets donot exist in Frequent 2-itemset.

Now we have, {1,3,4} (3), {1,3,5} (2)

{1,3,5} has 2 occurences and min support is 3. so, it cannot be included in Frequent 3-itemset.

Frequent 3-itemset L3:

$L3 = \{\{1,3,4\}\}$

END of Apriori

∴ Frequent itemset for Association rules is

FR = {{1}, {3}, {4}, {5}, {1,3}, {1,4}, {1,5}, {3,4}, {3,5}, {1,3,4}}

Single-Dimensional rules:

The rules formed involving only 1,2,3 are single dimensional.

From frequent itemset the rules formed out of {1,3} are single dimensional.

Two single-dimensional rules can be formed:

buys (x, "bread") → buys (x, "endive")

buys (x, "endive") → buys (x, "bread")

Multi-Dimensional rules are formed involving 1,2,3 and 4,5.

From frequent itemset the rules formed out of {1,3,4} are multi dimensional.

Inter-Dimensional rules:

All the rules formed out of {1,3,4} will include 1 and 3, and we see the predicate

"buys" repeats in rules formed.

So, no inter-dimensional rules can be formed from {1,3,4}.

Considering 2-item Frequent sets {1,4}, {1,5}, {3,4}, {3,5}.

Inter-Dimensional rules formed from these sets are:

buys (x, "bread") → livesin (x, "SBrook")

livesin (x, "SBrook") → buys (x, "bread")

buys (x, "endive") → livesin (x, "SBrook")

livesin (x, "SBrook") → buys (x, "endive")

buys (x, "bread") → income (x, "high")

income (x, "high") → buys (x, "bread")

buys (x, "endive") → income (x, "high")

income (x, "high") → buys (x, "endive")

These inter-dimensional rules are formed by taking different combinations of elements from the applicable frequent set & forming two disjoint sets in

which one set of elements implies the other.

Hybrid-Dimensional rules:

All the rules formed out of {1,3,4} are Hybrid-Dimensional rules, because they must involve 1,3 that represent the same predicate "buys".

Some of the hybrid dimensional rules are:

$lives(x, \text{"SBrook"}) \wedge buys(x, \text{"bread"}) \rightarrow buys(x, \text{"endive"})$

$lives(x, \text{"SBrook"}) \wedge buys(x, \text{"endive"}) \rightarrow buys(x, \text{"bread"})$

$buys(x, \text{"endive"}) \rightarrow lives(x, \text{"SBrook"}) \wedge buys(x, \text{"bread"})$

$buys(x, \text{"bread"}) \rightarrow lives(x, \text{"SBrook"}) \wedge buys(x, \text{"endive"})$

$lives(x, \text{"SBrook"}) \rightarrow buys(x, \text{"endive"}) \wedge buys(x, \text{"bread"})$

$buys(x, \text{"endive"}) \wedge buys(x, \text{"bread"}) \rightarrow lives(x, \text{"SBrook"})$

These hybrid dimensional rules are formed by

taking different combinations of 1, 3, 5 and forming two disjoint sets in which one set of elements implies the other.