|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Transactions** | **Items** | | | |
| 1 | Milk | Egg | Bread | Butter |
| 2 | Milk | Butter | Egg | Ketchup |
| 3 | Bread | Butter | Ketchup |  |
| 4 | Milk | Bread | Butter |  |
| 5 | Bread | Butter | Cookies |  |
| 6 | Milk | Bread | Butter | Cookies |
| 7 | Milk | Cookies |  |  |
| 8 | Milk | Bread | Butter |  |
| 9 | Bread | Butter | Egg | Cookies |
| 10 | Milk | Butter | Bread |  |
| 11 | Milk | Bread | Butter |  |
| 12 | Milk | Bread | Cookies | Ketchup |

The **minimum support level** is 33% of all the transactions.

12 transactions.

33% of 12 = **4 transactions (prune the itemsets)**

**Confidence level: 50%**

Frequent Itemsets Mining

1-item sets

**Candidates (Ci)**

|  |  |
| --- | --- |
| **1-item Sets** | **Frequency** |
| Milk | 9 |
| Bread | 10 |
| Butter | 10 |
| Egg | 3 |
| Ketchup | 3 |
| Cookies | 5 |

List of Frequent Items (Li)

|  |  |
| --- | --- |
| **Frequent 1-item Sets** | **Frequency** |
| Milk | 9 |
| Bread | 10 |
| Butter | 10 |
| Cookies | 5 |

Candidates

|  |  |
| --- | --- |
| **2-item Sets** | **Frequency** |
| Milk, Bread | 7 |
| Milk, Butter | 7 |
| Milk, Cookies | 3 |
| Bread, Butter | 9 |
| Bread, Cookies | 4 |
| Butter, Cookies | 3 |

|  |  |
| --- | --- |
| **Frequent 2-item Sets** | **Frequency** |
| Milk, Bread | 7 |
| Milk, Butter | 7 |
| Bread, Butter | 9 |
| Bread, Cookies | 4 |

Milk, Bread, Butter, Cookies

Candidates

|  |  |
| --- | --- |
| **3-item sets** | **Frequency** |
| Milk, Bread, Butter | 6 |
| Milk, Bread, Cookies | 2 |
| Milk, Butter, Cookies | 1 |
| Bread, Butter, Cookies | 3 |

|  |  |
| --- | --- |
| **Frequent 3-item sets** | **Frequency** |
| Milk, Bread, Butter | 6 |

**Association Rule Mining**

Frequent 3-item set = I =>

{Milk, Bread, Butter}

ind out all the non-empty subsets of I.

{{Milk} {Bread} {Butter} {Milk, Bread}

{Milk, Butter} {Bread, Butter}}

How to form the association rules?

For every non-empty subset S of I, the association rule is,

S -> (I-S)

If support(I)/support(S) >= minimum confidence threshold

I = {Milk, Bread, Butter}

S = {Milk}

6/12/9/12 = 6/9 = 0.66 > 50%

**Milk -> Bread, Butter**

Bread - > Milk, Butter

**FP Growth**

|  |  |
| --- | --- |
| **Transaction ID** | **Items** |
| T1 | {E, K, M, N, O , Y} |
| T2 | {D, E, K, N, O, Y} |
| T3 | (A, E, K, M) |
| T4 | {C, K, M, U, Y} |
| T5 | {C, E, I, K, O} |

Minimum support threshold: 3

|  |  |
| --- | --- |
| A | 1 |
| C | 2 |
| D | 1 |
| E | 4 |
| I | 1 |
| K | 5 |
| M | 3 |
| N | 2 |
| O | 3 |
| U | 1 |
| Y | 3 |

Frequent patter set:

**L = {K:5, E:4, M:3, O:3, Y:3}**

|  |  |  |
| --- | --- | --- |
| **Transaction ID** | **Items** | **Ordered-Item Set** |
| T1 | {E, K, M, N, O , Y} | {K, E, M, O, Y} |
| T2 | {D, E, K, N, O, Y} | {K, E, O, Y} |
| T3 | (A, E, K, M) | {K, E, M} |
| T4 | {C, K, M, U, Y} | {K, M, Y} |
| T5 | {C, E, I, K, O} | {K, E, O} |

|  |  |
| --- | --- |
| **Items** | **Conditional Pattern Base** |
| Y | {{K, E, M, O: 1}, {K, E, O: 1}, {K, M: 1}} |
| O | { {K, E, M: 1}, {K, E: 2}} |
| M | {{K, E: 2}. {K:1}} |
| E | {K:4} |
| K |  |

|  |  |  |
| --- | --- | --- |
| **Items** | **Conditional Pattern Base** | **Conditional Frequent Pattern Tree** |
| Y | {{K, E, M, O: 1}, {K, E, O: 1}, {K, M: 1}} | {K: 3} |
| O | { {K, E, M: 1}, {K, E: 2}} | {K, E: 3} |
| M | {{K, E: 2}. {K:1}} | {K:3} |
| E | {K:4} | {K:4} |
| K |  |  |

|  |  |
| --- | --- |
| **Items** | **Conditional Pattern Base** |
| Y | {<K, Y: 3>} |
| O | {<K, O: 3>, <E, O: 3>, <K, E, O: 3>} |
| M | {<K, M:3>} |
| E | {<K, E: 4>} |
| K |  |

Considering the **first frequent pattern** we can generate two association rules as:

Y -> K

and

K -> Y

We need to check if they are **satisfying the minimum confidence level**.

NULL

Y: 1

M: 1

Y: 1

O: 2

Y: 1

O: 1

M: 2

E: 4

K: 5