|  |  |
| --- | --- |
| **Transaction ID** | **Items Purchased** |
| 1 | Bread, Cheese, Egg, Juice |
| 2 | Bread, Cheese, Juice |
| 3 | Bread, Milk, Yogurt |
| 4 | Bread, Juice, Milk |
| 5 | Cheese, Juice, Milk |

**Association Rule Mining--Apriori Algorithm Solved Problems** Q.1) For the following given Transaction Data-set, Generate Rules using Apriori Algorithm. Consider the values as Support=50% and Confidence=75%

Answer:

Given Support=50% and Confidence=75%

**Step 1) Find Frequent Item Set and their support**

|  |  |  |
| --- | --- | --- |
| **Item** | **Frequency** | **Support (in %)** |
| Bread | 4 | 4/5=80% |
| Cheese | 3 | 3/5=60% |
| Egg | 1 | 1/5=20% |
| Juice | 4 | 4/5=80% |
| Milk | 3 | 3/5=60% |
| Yogurt | 1 | 1/5=20% |

**Support (item) = Frequency of item/Number of transactions**

**Step 2) Remove all the items whose support is below given minimum support.**

|  |  |  |
| --- | --- | --- |
| **Item** | **Frequency** | **Support (in %)** |

|  |  |  |
| --- | --- | --- |
| Bread | 4 | 4/5=80% |
| Cheese | 3 | 3/5=60% |
| Juice | 4 | 4/5=80% |
| Milk | 3 | 3/5=60% |

**Step 3) Now form the two items candidate set and write their frequencies.**

|  |  |  |
| --- | --- | --- |
| **Items Pair** | **Frequency** | **Support (in %)** |
| Bread, Cheese | 2 | 2/5=40% |
| Bread, Juice | 3 | 3/5=60% |
| Bread, Milk | 2 | 2/5=40% |
| Cheese, Juice | 3 | 3/5=60% |
| Cheese, Milk | 1 | 1/5=20% |
| Juice, Milk | 2 | 2/5=40% |

**Step 4) Remove all the items whose support is below given minimum support.**

|  |  |  |
| --- | --- | --- |
| **Items Pair** | **Frequency** | **Support (in %)** |
| Bread, Juice | 3 | 3/5=60% |
| Cheese, Juice | 3 | 3/5=60% |

**Step 5) Generate rules**

For Rules we consider item pairs:

* (Bread, Juice)

Bread->Juice and Juice->Bread

* (Cheese, Juice)

Cheese->Juice and Juice->Cheese

**Confidence (A->B) = support (AUB)/support (A)**

**Therefore,**

* Confidence (Bread->Juice) = support (Bread U Juice)/support (Bread)

= 3/5 \* 5/4=3/4= 75%

* Confidence (Juice->Bread) = support (Juice U Bread)/support (Juice)

= 3/5\*5/4=3/4=75%

* Confidence (Cheese->Juice) = support (Cheese U Juice)/support (Cheese)

=3/5\*5/3=1=100%

* Confidence (Juice->Cheese) = support (Juice U Cheese)/support (Juice)

= 3/5\*5/4=3/4=75%

All the above rules are good because the confidence of each rule is greater than or equal to the minimum confidence given in the problem.

* **For the following given transaction data set, generate rules using Apriori Algorithm. Consider the values as Support=22% and Confidence= 70%**

|  |  |
| --- | --- |
| **Transaction ID** | **Items Purchased** |
| **1** | **I1,I2,I5** |
| **2** | **I2,I4** |
| **3** | **I2,I3** |
| **4** | **I1,I2,I4** |
| **5** | **I1,I3** |
| **6** | **I2,I3** |
| **7** | **I1,I3** |
| **8** | **I1,I2,I3,I5** |
| **9** | **I1,I2,I3** |