

**INDIAN INSTITUTE OF INFORMATION TECHNOLOGY,  
ALLAHABAD  
(Deemed University)**

(A centre of excellence in IT, established by Ministry of HRD, Govt. of India)

---

**Object Oriented Modeling Lab**

Prof. O.P. Vyas

**OOM Mini Project:** Market Basket Analysis

**Abstract**

Develop software which performs the association rule mining on given transitional data of super market sales transition

The software makes use of two data mining algorithms namely APRIORI algorithm. The software will graphically represent generation of frequent items sets for given value of minimum support and confidence.

**Specific Technology**

Java, Swing

**Project Tasks**

Use case analysis, Design, Implementation, User Interface

**Functional components of the project**

Following is a list of functionalities of the system. More functionality that you find appropriate can be added to this list. And, in places where the description of functionality is not adequate, you can make appropriate assumptions and proceed.

Existing information in the database contains the transactions for the items in a supermarket.

1. The GUI screen will take as input a transaction data base (file). And minimum value of support and confidence provided by user. The result page for apriori algorithm should be divided in three parts.

The first part shows the tracing of code of apriori algorithm

The second part shows the generation of frequent items sets with respect to the execution of the code.

Third part will show the generated association rule .

\*\*\* **The steps of algorithms are given below.**

### **Apriori Algorithm:**

Find frequent itemsets using an iterative level-wise approach based on candidate generation.

Input : Database , D, of transactions; minimum support threshold, min\_sup,  
Association rule

Output : L, frequent itemsets in D.

#### **Steps:**

- Scan the transaction database to get the support S of each 1 itemsets . where support is the number of transaction containing that item. Compare S with min sup. And get the frequent 1 itemset L1, by taking only those whose support value is more than the min sup.
- Use L(k-1) join L(k-1) to generate a set of candidate k itemsets. And use the Apriori property(i.e to discard those whose subset are already been discarded in earlier steps due to less support) to prune the unfrequented k itemsets from this set.
- Scan the transaction database to get the support S of each candidate k-itemset in the final set, compare S with min\_sup, and get a set of frequent k-itemsets, L k.
- If the candidate set is null then goto stop otherwise goto second step .

#### **Association rules generation :**

Once the frequent itemsets are found ,the strong association rules can be generated that satisfies both minimum support and minimum confidence.

$\text{Confidence}(A \Rightarrow B) = P(B|A) = \frac{\text{support\_count}(A \cup B)}{\text{support\_count}(A)}.$

Where  $\text{support\_count}(A \cup B)$  = no.of transactions containing the itemset A U B and  $\text{support\_count}(A)$  = no. of transactions contains A.

- For each frequent itemset I, generate all nonempty subsets of I.
- For every nonempty subset s of I, output the rule " s => (I-s)" if  $(\frac{\text{support S of I}}{\text{support S of s}}) \geq \text{min\_conf}.$