St. Francis Institute of Technology, Mumbai-400 103

**Department Of Information Technology**

A.Y. 2024-2025

Class: TE-ITA/B, Semester: VI

Subject: **Business Intelligence Lab**

**Experiment – 9 To implement Apriori association mining using any one language (JAVA/Python)**

1. **Aim: :** Implementation of Apriori Association in Data Mining using any one Language
2. **Objectives:** After study of this experiment, the students will be able toimplement Apriori Algorithm in (JAVA/R/Python)
3. **Outcomes:** After study of this experiment, the students will be able to

**CO 5:** Design and Implement various frequent data mining techniques and formulate association rules   on large data sets

1. **Prerequisite:** Introduction to algorithms of Associativity
2. **Requirements:** Personal Computer, Windows XP operating system/Windows 7, Internet Connection, Microsoft Word, WEKA tool.
3. **Theory:**
4. **What are Association Rules in Data Mining?**

**ANS:**

1. Association rules are data mining techniques used to find relationships between variables in large datasets.  
2. They are typically expressed in “if-then” format (e.g., *if A then B*).  
3. These rules are derived from frequent itemsets found in transactional data.  
4. Commonly used in market basket analysis to find items bought together.  
5. Each rule has associated metrics like support, confidence, and lift.  
6. Support tells how frequently the itemset appears in the dataset.  
7. Confidence measures how often the rule has been found to be true.  
8. Lift indicates the strength of the rule over random chance.  
9. Useful in retail, e-commerce, banking, and healthcare industries.

1. **How Association helps in Boosting the Business Profit**

**ANS:**

1.Identifies customer buying habits and product relationships.  
2. Enables effective cross-selling and up-selling strategies.  
3. Helps design combo offers or product bundles to increase sales.  
4. Improves store layout by placing related items together.  
5. Reduces inventory costs by stocking related high-demand products.  
6. Personalized recommendation systems (e.g., "Customers also bought...").  
7. Targets promotions and discounts more effectively.  
8. Increases customer satisfaction by anticipating needs.  
9. Enhances product placement and merchandising strategies.

1. **Laboratory Exercise:** Implementation of Apriori Algorithm in Java/Python, printout of implementation along with coding and snapshot.
2. **Post-Experiments Exercise**
3. **Questions:**
   * MCQ type test
   * Compare Apriori and FP Tree
4. **Conclusion:**
   * Summary of Experiment
   * Importance of Experiment
   * Application of Experiment

**Reference:** Data Mining: Concept & Techniques, 3rd Edition, Jiawei Han, Micheline  Kamber, Jian Pei, Elsevier.

**CODE:**

import pandas as pd

from mlxtend.preprocessing import TransactionEncoder

from mlxtend.frequent\_patterns import apriori, association\_rules

# Step 1: Define dataset

dataset = [

['eggs', 'milk', 'bread'],

['milk', 'diaper', 'beer', 'bread'],

['milk', 'diaper', 'beer', 'cola'],

['eggs', 'milk', 'diaper', 'beer'],

['eggs', 'milk', 'diaper', 'cola']

]

# Step 2: Transform data into one-hot encoded DataFrame

te = TransactionEncoder()

te\_ary = te.fit(dataset).transform(dataset)

df = pd.DataFrame(te\_ary, columns=te.columns\_)

# Step 3: Get frequent itemsets with minimum support of 0.6

frequent\_itemsets = apriori(df, min\_support=0.6, use\_colnames=True)

# Step 4: Generate association rules with minimum confidence of 0.7

rules = association\_rules(frequent\_itemsets, metric="confidence", min\_threshold=0.7)

# Step 5: Display outputs

print("=== Frequent Itemsets ===")

print(frequent\_itemsets)

print("\n=== Association Rules with All Metrics ===")

print(rules[[

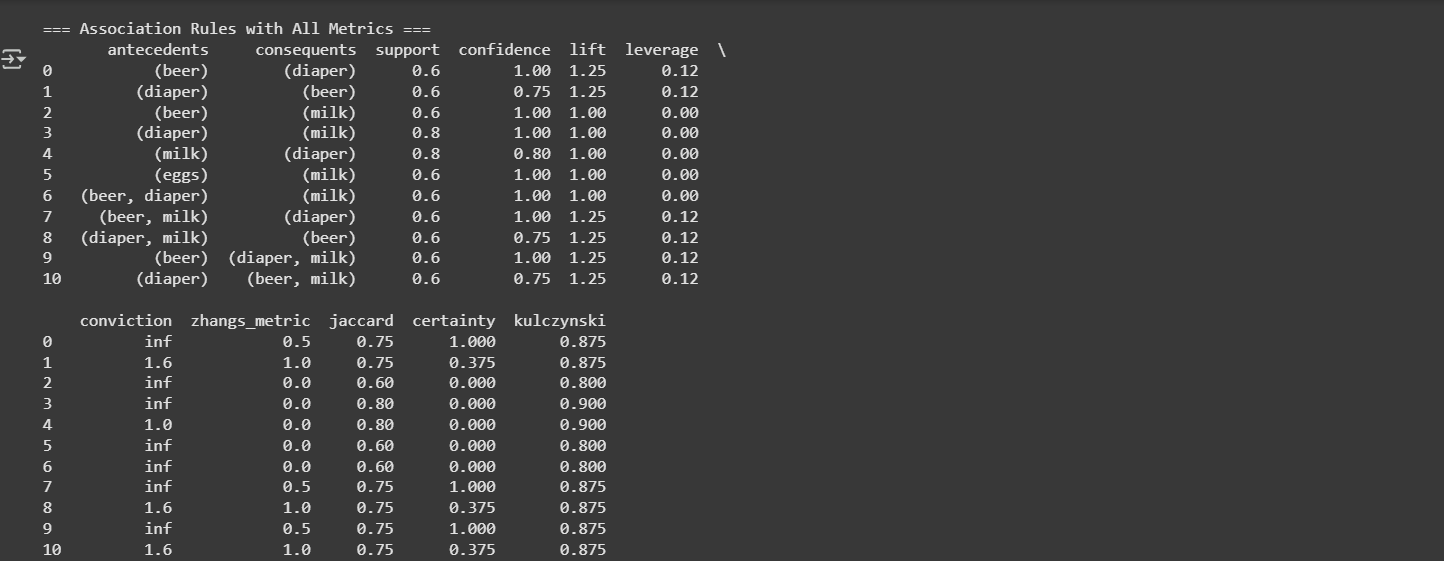
'antecedents', 'consequents', 'support', 'confidence', 'lift',

'leverage', 'conviction', 'zhangs\_metric', 'jaccard',

'certainty', 'kulczynski'

]])

**OUTPUT:**

****