**Data Analytics Laboratory**

**Task 2**

**Association Mining: Finding Association Rules using Apriori principle**

**Introduction**

* Association rules is an unsupervised learning method.
* This is a descriptive, not predictive, method often used to discover interesting relationships hidden in a large dataset.
* The identified relationship is usually called as frequent itemset.
* Association rules are mostly used in mining transaction in databases.
* Each of the uncovered rules is in the form X → Y, meaning that when item X is observed, item Y is also observed.
* Using association rules, patterns can be discovered from the data.
* Association rules are sometimes referred to as market basket analysis.
* An itemset containing k items is called a k-itemset. Use curly braces like
* {item 1, item 2, . . . item k}to denote a k-itemset.
* Apriori as the main focus of the discussion of association rules.
* Apriori is one of the earliest and the most fundamental algorithms for generating association rules.

**Prerequisites**

Define support. If 80% of all transactions contain itemset {bread}, then what is the support value of {bread}?

Define Confidence. {mobile phone, mobile cover, scratch card} has a support count of 0.34. {mobile phone, mobile cover} has a support count of 0.67. What is the Confidence for the above example?

Define Lift. Assuming 1,000 transactions, {milk, eggs} appearing in 300 of them, {milk} appearing in 500, {eggs} appearing in 400, Calculate Lift for the relation (milk →eggs).

**Apply Apriori Algorithm on the below given dataset.**

1. Find which one item repeated mostly in all the transactions.
2. How many pairs of items are there in the above dataset satisfying support value > 0.2.
3. How many pairs of items are there in the above dataset satisfying confidence

value > 0.02.

<https://bit.ly/3qx5rrG>

**Results**

The program is implemented in python and the output is observed.

**Faculty Signature**