Academic Year: 2025-26 **Semester: V**
Class/Branch: T.E. DS **Subject: DWMLab**

EXPERIMENT NO. 6

1. **Aim:** Implementation of association mining algorithm like Apriori using python.

2. **Objectives:** From this experiment, the student will be able to

- Learn about the association mining algorithm and frequent pattern.
- Learn about Apriori algorithm.

3. **Theory:**

Frequent Pattern Mining in Data Mining:

Frequent pattern mining in data mining is the process of identifying patterns or associations within a dataset that occur frequently. This is typically done by analyzing large datasets to find items or sets of items that appear together frequently.

Techniques for Frequent Pattern Mining:

1. **Apriori algorithm:**
   One of the most popular methods, the Apriori algorithm, uses a step−by−step procedure to find frequent item sets. It starts by creating candidate itemsets of length 1, determining their support, and eliminating any that fall below the predetermined cutoff. The method then joins the frequent itemsets from the previous phase to produce bigger itemsets repeatedly.
   Once no more common item sets can be located, the procedure is repeated. The Apriori approach is commonly used because of its efficiency and simplicity, but because it requires numerous database scans for big datasets, it can be computationally inefficient.

2. FP−growth Algorithm:

FP−growth can be much quicker than Apriori by skipping the construction of candidate itemsets, which lowers the number of runs over the dataset.

The implementation of the Apriori algorithm is explained in the steps below:

Code:

1. Importing required libraries.

```
import pandas as pd
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules
```

2. Provide the input data with transactions and data items.

```
data = {'Milk': [0, 0, 1, 0, 1],
        'Bread': [1, 1, 1, 1, 0],
        'Butter': [1, 1, 1, 0, 0],
        'Eggs': [1, 0, 0, 0, 0],
        'Coke': [0, 0, 0, 1, 1],    }
```

3. Declare a dataframe for the input data.

```
df = pd.DataFrame(data)
```

4. Frequent Itemsets Generation using Apriori Algorithm.

```
frequent_itemsets = apriori(df, min_support=0.3, use_colnames=True)
```

5. Generating Association Rules.

```
rules = association_rules(frequent_itemsets, metric="confidence",
min_threshold=0.8)
```

.

6. Displaying Results.

```
print("Frequent Itemsets:\n", frequent_itemsets)
print("\nAssociation Rules:\n", rules[['antecedents', 'consequents',
'support', 'confidence']])
```

Output:

```
Frequent Itemsets:
    support          itemsets
0       0.4           (Milk)
1       0.8          (Bread)
2       0.6         (Butter)
3       0.4           (Coke)
4       0.6   (Butter, Bread)

Association Rules:
   antecedents consequents  support  confidence
0    (Butter)     (Bread)      0.6         1.0
```

4. Conclusion: