Lab 7 – SECTION A, BATCH 1 Date: 19th OCT. 2022

Exer 1: Association Rule Mining

1. Use the "groceries.csv" dataset and answer the following:
2. How many transactions and items are there in the data set?
3. Prepare the data for finding association rules. Each transaction will contain a list of item in the transaction.
   [['citrus fruit', 'semi-finished bread', 'margarine', 'ready soups'],
   ['tropical fruit', 'yogurt', 'coffee'],…….
   ['whole milk']]

4. Use Python library *mlxtend* and convert the transactions into a format that can be used in the Apriori method for finding frequent itemsets.
   pip install mlxtend
   from mlxtend.preprocessing import TransactionEncoder
   from mlxtend.frequent_patterns import apriori, association_rules
5. Find top selling items with minimum support of 2%.
6. Find all frequent itemsets with minimum support of 5%.
7. Find all frequent itemsets of length 2 with minimum support of 2%.
8. Find the top 10 association rules with minimum support of 2%, sorted by confidence in descending order.
9. Find association rules with minimum support of 2% and lift of more than 1.0.