**Vivek Deswal** **BDA Project** **Section B**
**180002**

**Q1: Give an example where you can apply the Apriori algorithm.**
**Answer:** Apriori algorithm is a classical algorithm in data mining. It is used for mining frequent item sets and relevant association rules. It is devised to operate on a database containing a lot of transactions, for instance, items brought by customers in a store. It helps the customers buy their items with ease, and enhances the sales performance of the departmental store. This algorithm has utility in the field of healthcare as it can help in detecting adverse drug reactions (ADR) by producing association rules to indicate the combination of medications and patient characteristics that could lead to ADRs.

**Q2: What happens when we decrease the support level? Why?**
**Answer:** When we decrease the support level, we decrease the amount of popularity we need. Support represents the popularity of that product of all the product transactions. Support of the product is calculated as the ratio of the number of transactions includes that product and the total number of transactions. Support of the product = (Number of transactions includes that product)/ (Total number of transactions).

**Q3: What happens when we increase the confidence level? Why?**
**Answer:** When we increase the confidence level, we tend to increase the probability of that item to be transacted with the compared item. This explains how likely Y is purchased when X is purchased. This defines association between two items. For example, when a person buys milk is more likely to buy bread as well or vice versa. This is measured by the proportion of transactions with item X, in which item Y also appears. Expressed as {X -> Y}. Calculated by the proportion of number of transactions in which both (X & Y) occurs to support of the item X.

**Q4: How many rules are generated with a support level of 5% and a confidence level of 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%? Also visualized the generated rules.**
**Answer:**

```
Answer-4:
Support Level of 5%:

Confidence          count
(-0.001, 10.0]       0
(10.0, 20.0]         2
(20.0, 30.0]         1
(30.0, 40.0]         0
(40.0, 50.0]         0
(50.0, 60.0]         1
(60.0, 70.0]         0
(70.0, 80.0]         0
(80.0, 90.0]         0
Name: confidence, dtype: int64
```

**Q5: How many rules are generated with a support level of 1% and a confidence level of 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%? Also visualized the generated rules.**

**Answer:**

```
Answer 5:
Support Level of 1%:

Confidence              count
(-0.001, 10.0]             30
(10.0, 20.0]               14
(20.0, 30.0]               11
(30.0, 40.0]                3
(40.0, 50.0]                5
(50.0, 60.0]               10
(60.0, 70.0]                0
(70.0, 80.0]                1
(80.0, 90.0]                0
Name: confidence, dtype: int64
```

**Q6: How many rules are generated with a support level of 0.5% and a confidence level of 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%? Also visualized the generated rules.**

**Answer:**

```
Answer 6:
Support Level of 0.5%:

Confidence              count
(-0.001, 10.0]             92
(10.0, 20.0]               46
(20.0, 30.0]               21
(30.0, 40.0]               11
(40.0, 50.0]                7
(50.0, 60.0]               13
(60.0, 70.0]                2
(70.0, 80.0]                1
(80.0, 90.0]                1
Name: confidence, dtype: int64
```

**Q7: Analyze the results of questions 4, 5, 6 and choose the optimal threshold value for support and confidence. What values do you choose for support and confidence? Why?**

```
Answer 7:
Rules:

   antecedents  consequents  antecedent support  ...      lift  leverage  conviction
0      (Bread)  (Alfajores)            0.327205  ...  0.870657 -0.001538    0.995145
1  (Alfajores)      (Bread)            0.036344  ...  0.870657 -0.001538    0.940818
2  (Alfajores)     (Coffee)            0.036344  ...  1.130235  0.002264    1.135648
3     (Coffee)  (Alfajores)            0.478394  ...  1.130235  0.002264    1.004936
4      (Bread)    (Brownie)            0.327205  ...  0.822508 -0.002326    0.992651

[5 rows x 9 columns]


pairs with best confidence:

         antecedents consequents   support  confidence      lift
41           (Toast)    (Coffee)  2.366614   70.440252  1.472431
38  (Spanish Brunch)    (Coffee)  1.088220   59.883721  1.251766
26       (Medialuna)    (Coffee)  3.518225   56.923077  1.189878
30          (Pastry)    (Coffee)  4.754358   55.214724  1.154168
2        (Alfajores)    (Coffee)  1.965135   54.069767  1.130235


pairs with best support:

   antecedents consequents   support  confidence      lift
15    (Coffee)      (Cake)  5.472795   11.439929  1.101515
14      (Cake)    (Coffee)  5.472795   52.695829  1.101515
30    (Pastry)    (Coffee)  4.754358   55.214724  1.154168
31    (Coffee)    (Pastry)  4.754358    9.938163  1.154168
```

```
As We can see from the above tables, The best possible pairs would be Cake & Coffee, Paestry & Cake

pairs with worst confidence:

   antecedents       consequents   support  confidence      lift
0      (Bread)       (Alfajores)  1.035394    3.164353  0.870657
51    (Coffee)   (Bread, Pastry)  1.119915    2.340989  0.802807
39    (Coffee)  (Spanish Brunch)  1.088220    2.274735  1.251766
47    (Coffee)     (Bread, Cake)  1.003698    2.098057  0.898557
57    (Coffee)       (Cake, Tea)  1.003698    2.098057  0.882582


pairs with best lift:

        antecedents      consequents   support  confidence      lift
55           (Cake)    (Tea, Coffee)  1.003698    9.664293  1.937977
54    (Tea, Coffee)           (Cake)  1.003698   20.127119  1.937977
16           (Cake)  (Hot chocolate)  1.141046   10.986775  1.883874
17  (Hot chocolate)           (Cake)  1.141046   19.565217  1.883874
19            (Tea)           (Cake)  2.377179   16.666667  1.604781


pairs with worst lift:

        antecedents      consequents   support  confidence      lift
5         (Brownie)          (Bread)  1.077655   26.912929  0.822508
7         (Cookies)          (Bread)  1.447438   26.601942  0.813004
```

**Q9: What recommendations would you give to the owner of the bakery?**
**Answer:**

- There is 70% chance that he/she will buy coffee.
- Never recommend your customers to buy brownie with bread.