Date __/__/____

Assignment - 3

* **Title :-** Apply a-priory algorithm to find frequently occuring items from given data and generate strong association rules using support and confidence thresholds.
  Example :- Market Basket Analysis.

* **Objective :-** Implement a-priory algorithm on dataset using python. Finding frequently occuring items from given data.

* **Outcomes :-** Will be able to implement a-priory algorithm in python.

* **H/W & S/W Requirements :-** Python 3.8, Jupyter Notebook, 64-bit OS, Machine with 64-bit processor.

* **Theory :-**

A-priory Algorithm :- A-priory is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to large and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by A-priory can be used to determined association rules which highlight general trends in the database. This have applications in domains such as market basket analysis.

3) Finding itemsets with high support :-
Using the a-priory principle, the no. of itemsets that have to be examined can be proved, and the list of popular itemsets can be

obtained in these steps :-

(i) Start with itemsets containing just a single item.

(ii) Determine the support for itemsets. Keep the itemsets that meet your minimum support threshold.

(iii) Using the itemset generate all possible itemset configurations.

(iv) Repeat steps 2 & 3 until there are no more new items.

⇒ Finding item rules with high confidence or lift :

We have seen how the a-priory algorithm can be used to identify itemsets with high support. The same principle can also be used to identify item associations with high confidence or lift. Finding rules with high confidence or lift is less computationally costing once high support itemsets have been identified, because confidence and lift values are calculated using support values.

Example:- If the rule – {Bacon, Soup → Soda} has low confidence, all other rules with the same consistent items and with apple in the right hand side would have low confidence too.

* Conclusions- Thus, we have applied a-priory algorithm to find frequently occuring items from given data and generated strong association rules using support & confidence thresholds.