

Big data-2

Project-1

Jawahar - 170004

Question 1: Give an example where you can apply the Apriori algorithm.

Ans-

When you talk of data mining, the discussion would not be complete without the mentioning of the term, 'Apriori Algorithm.' Apriori algorithm, a classic algorithm, is useful in mining frequent itemsets and relevant association rules. Usually, you operate this algorithm on a database containing a large number of transactions.

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.

Question2: What happens when we decrease the support level? Why?

Ans-

Support level gives the frequency (no. of times the item occurred) of the item in the dataset. 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)

From here we can infer that Support level directly proportional to frequency, so when we decrease the support level, we decrease the amount of popularity we need.

Question 3. What happens when we increase the confidence level? Why?

Ans-

Confidence can be interpreted as the likelihood of purchasing both the products A and B. Confidence is calculated as the number of transactions that include both A and B divided by the number of transactions includes only product A.

Confidence (A=>B) = (Number of transactions includes both A and B)/ (Number of transactions includes only product A)

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 $\{X \rightarrow Y\}$. Calculated by the proportion of number of transactions in which both (X & Y) occurs to support of the item X.

When we increase the confidence level, we tend to increase the probability of that item to be transacted with the compared item.

Question 10. What recommendations would you give to the owner of the bakery?

Ans-

Recommendations from my Analysis:

- People who buy a coffee and tea have most chance in buying a cake, so make a combo offer on three of them.
- The most famous combo in the findings are hot chocolate and cake, so make a combo offer for this also.
- High demand Items are Pairs of cake- coffee and pastry- coffee.
- Remember to not recommend customers to buy brownie with bread.
- If a person buys a toast, there is chance of 70% that he/she will buy coffee.