Association Rule

Association rule mining is one of the most popular data mining methods. However, mining association rules often results in a very large number of found rules, leaving the analyst with the task to go through all the rules and discover interesting ones.

Many business enterprises accumulate large amounts of data in their day-to-day operations.Take an example of a Super Market where customers can buy variety of items. Usually, there is a pattern in what the customers buy. For instance, mothers with babies buy baby products such as milk and diapers. Damsels may buy makeup items whereas bachelors may buy beers and chips etc. In short, transactions involve a pattern. More profit can be generated if the relationship between the items purchased in different transactions can be identified.

For instance, if item X and Y are bought together more frequently then several steps can be taken to increase the profit. For example:

1. X and Y can be placed together so that when a customer buys one of the product he doesn't have to go far away to buy the other product.
2. People who buy one of the products can be targeted through an advertisement campaign to buy the other.
3. Collective discounts can be offered on these products if the customer buys both of them.
4. Both X and Y can be packaged together.

The process of identifying an associations between products is called association rule mining.

Theory of Apriori Algorithm

> There are four major components of Apriori algorithm : **Support, Confidence, Lift, lenght**

We will explain these three concepts with the help of an example.

Suppose we have a record of 1 thousand customer transactions, and we want to find the Support, Confidence, and Lift for two items. burgers and ketchup. Out of one thousand transactions, 100 contain ketchup while 150 contain a burger. Out of 150 transactions where a burger is purchased, 50 transactions contain ketchup as well. Using this data, we want to find the support, confidence, and lift.

Support

Support refers to the default popularity of an item and can be calculated by finding number of transactions containing a particular item divided by total number of transactions. Suppose we want to find support for item Y. This can be calculated as:

Support(Y) = (Transactions containing (Y))/(Total Transactions)

For instance if out of 1000 transactions, 100 transactions contain Ketchup then the support for item Ketchup can be calculated as:

Support(Ketchup) = (Transactions containing Ketchup)/(Total Transactions)

Support(Ketchup) = 100/1000 = 10%

Confidence

Confidence refers to the likelihood that an item Y is also bought if item X is bought. It can be calculated by finding the number of transactions where X and Y are bought together, divided by total number of transactions where X is bought. Mathematically, it can be represented as:

Confidence(X→Y) = (Transactions containing both (X and Y))/(Transactions containing X)

Coming back to our problem, we had 50 transactions where Burger and Ketchup were bought together. While in 150 transactions, burgers are bought. Then we can find likelihood of buying ketchup when a burger is bought can be represented as confidence of Burger -> Ketchup and can be mathematically written as:

Confidence(Burger→Ketchup) = (Transactions containing both (Burger and Ketchup))/(Transactions containing X)

Confidence(Burger→Ketchup) = 50/150 = 33,3%

You may notice that this is similar to what you'd see in the Naive Bayes Algorithm, however, the two algorithms are meant for different types of problems.

Lift

Lift(X -> Y) refers to the increase in the ratio of sale of Y when X is sold. Lift(X –> Y) can be calculated by dividing Confidence(X -> Y) divided by Support(X). Mathematically it can be represented as:

Lift(X→Y) = (Confidence (X→Y))/(Support (Y))

Coming back to our Burger and Ketchup problem, the Lift(Burger -> Ketchup) can be calculated as:

Lift(Burger→Ketchup) = (Confidence (Burger→Ketchup))/(Support (Ketchup))

Lift(Burger→Ketchup) = 33.3/10 =3.33

Lift basically tells us that the likelihood of buying a Burger and Ketchup together is 3.33 times more than the likelihood of just buying the ketchup. X Lift of 1 means there is no association between products X and Y. Lift of greater than 1 means products X and Y are more likely to be bought together. Finally, Lift of less than 1 refers to the case where two products are unlikely to be bought together.