**ASSOCIATION RULE LEARNING**

**(Apriori)**

To understand this algorithm let us take a scenario.

There was a company (or a store) that did some analytics around the products that people are purchasing and analysed thousands of transactions. They found a very interesting thing that very often during certain times of the day mostly after afternoon between 6 and 9 p.m. people who buy diaper also buy beer or vice versa. And it was completely out of the blue that twhy these 2 products were bought together when they are not at all connected to each other. But this was a fact that they came across in the data and one of the possible explanations was that when the husbands get home after work, the husband and the wife take care of the baby and when they run out of diapers the husbands go to buy them. While he is picking up the diapers it is hours after work and as he is already in the convenience store so he also picks up some beer. This is one of the possible explanations. This may or may not be the case. But based on that you can decide where to arrange the products in the store.

So, some stores might decide to put these 2 products closer to entice people to buy a beer when they are buying diaper.

But actually a lot of stores do the opposite. They separate beer and diapers. They may also seperate bread and milk because that way they would really know that these products are bought together and so you actually have to walk through the whole store to pick up. And as you are walking through the store you see more other products and you are more likely to pick up an additional item that you were not actually planning on buying initially.

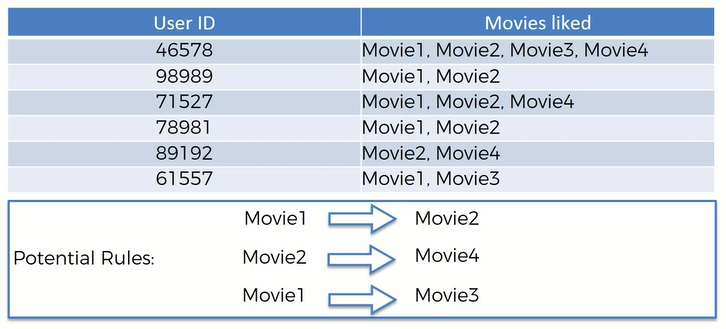
So, there is a lot of interesting marketing tactics that are used based on this data.

But how to get this data?

-> Apriori algorithm is one of the ways to get this data.

So, apriori is based on **“people who bought somethng also bought something else or people who watch something also watch something else or people who did something also did something else”**.

So, t**his association is all about analyzing things that come in pairs or in triplets or combined together for some reason**.



By looking at this data we might come across many rules but we have to find the strongest one for making business decisions.

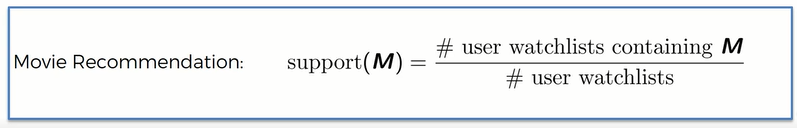
WORKING OF APRIORI ALGORITHM

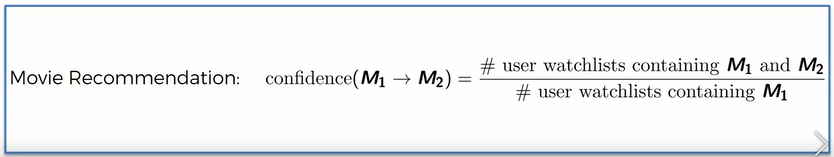
Apriori algorithm has three parts to it.

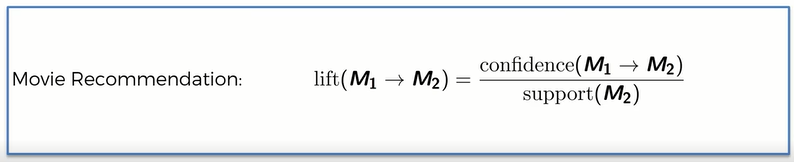
1. Support

2. Confidence

3. Lift







Let there be 10 people who have seen the movie ex - machina out of 100.

Therefore support = 10/100 = 10%

40 people have seen Interstellar, and out of them 7 people have seen ex – machina.

Therefore confidence = 7/40 = 17.5%

Lift is basically confidence divided by support.

(If to a random population we recommend to watch a movie what is the likelihood that they will watch it)

So, let’s recommend people to watch ex – machina only who have seen interstellar.

Therefore lift = (17.5%) / (10%) = 1.75

So, this is the percentage of successful recommendation.

