Association Rule Learning
Apriori Intuition

ARL - What is it all about ?

ARL - What is it all about ?

This is about a story that tells people who usually buy the diaper between 6pm until 9pm, they also buy bear. This is something that come out of noting and from data.

ARL - What is it all about ?

People who bought also bought ...

This kind of data is perfect for stores in order to put the some good next to each other.

ARL - Movie Recommendation

User ID	Movies liked
46578	Movie1, Movie2, Movie3, Movie4
98989	Movie1, Movie2
71527	Movie1, Movie2, Movie4
78981	Movie1, Movie2
89192	Movie2, Movie4
61557	Movie1, Movie3

Potential Rules:

Movie1 ⇒ Movie2

Movie2 ⇒ Movie4

Movie1 ⇒ Movie3

People who watch Movie 1 are likely to like Movie 2 and so on.

ARL - Market Basket Optimisation

Transaction ID	Products purchased
46578	Burgers, French Fries, Vegetables
98989	Burgers, French Fries, Ketchup
71527	Vegetables, Fruits
78981	Pasta, Fruits, Butter, Vegetables
89192	Burgers, Pasta, French Fries
61557	Fruits, Orange Juice, Vegetables
87923	Burgers, French Fries, Ketchup, Mayo

Potential Rules:

Burgers ⇒ French Fries

Vegetables ⇒ Fruits

Burgers, French Fries ⇒ Ketchup

Apriori algorithm has three parts: Support, Confident and Lift.

Apriori - Support

Movie Recommendation: $\text{support}(\boldsymbol{M}) = \frac{\#\text{ user watchlists containing } \boldsymbol{M}}{\#\text{ user watchlists}}$

Market Basket Optimisation: $\text{support}(\boldsymbol{I}) = \frac{\#\text{ transactions containing } \boldsymbol{I}}{\#\text{ transactions}}$

Number of users who watches movie M, over total number of users.

Apriori - Support

Total of 100 people

Apriori - Support

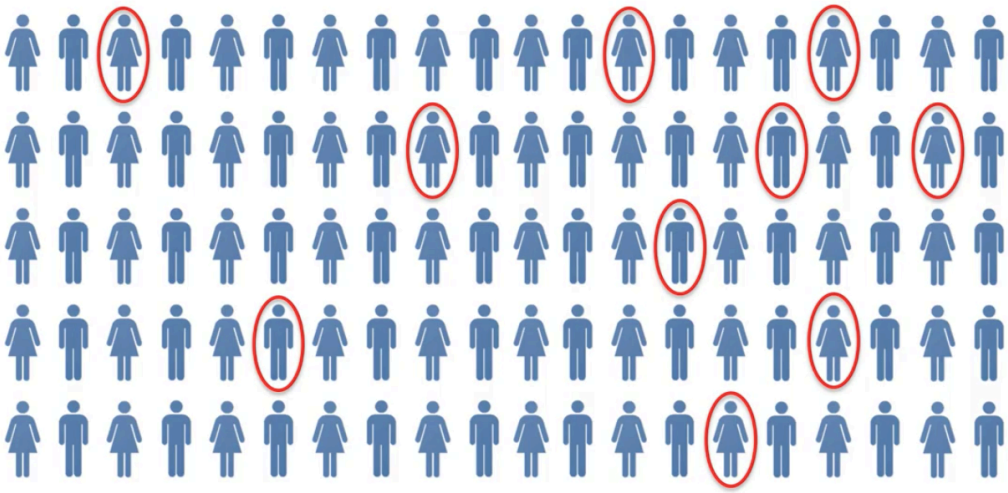
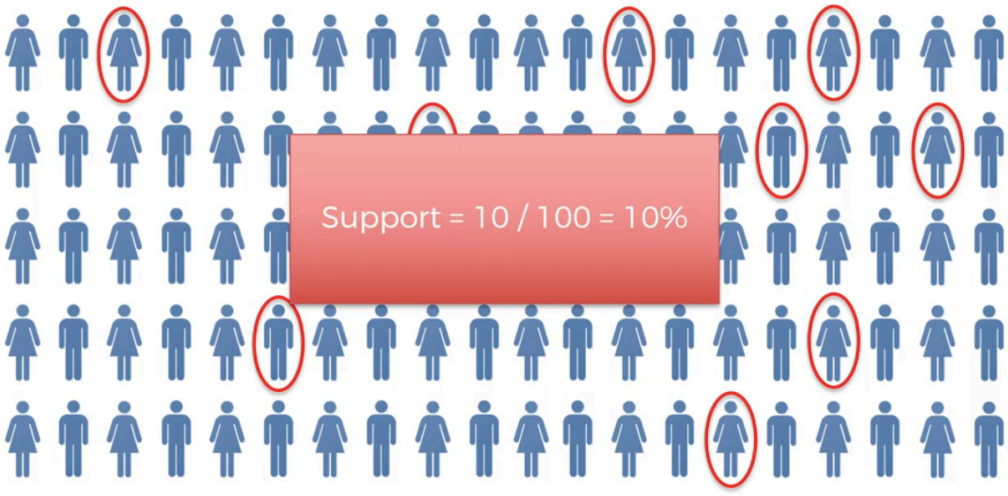


Image 10 people out of 100 watched Ex-Machina

Apriori - Support



Apriori - Confidence

Movie Recommendation: $\text{confidence}(M_1 \rightarrow M_2) = \frac{\#\text{ user watchlists containing } M_1 \text{ and } M_2}{\#\text{ user watchlists containing } M_1}$

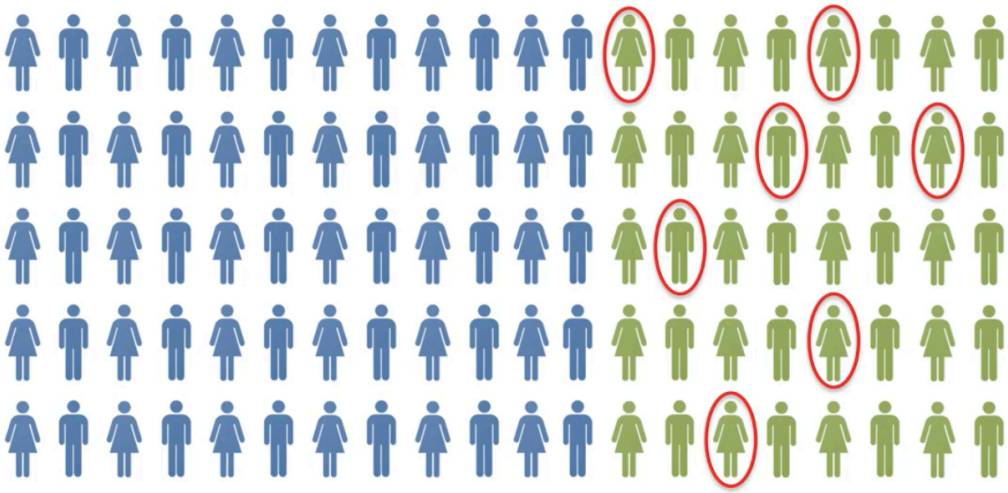
Market Basket Optimisation: $\text{confidence}(I_1 \rightarrow I_2) = \frac{\#\text{ transactions containing } I_1 \text{ and } I_2}{\#\text{ transactions containing } I_1}$

Number of people who watched movie M1 and M2, over total number of people who watched M1.

Apriori - Confidence

Green people watched Interstellar

Apriori - Confidence



Red circle people watched Ex-Machina

Apriori - Confidence

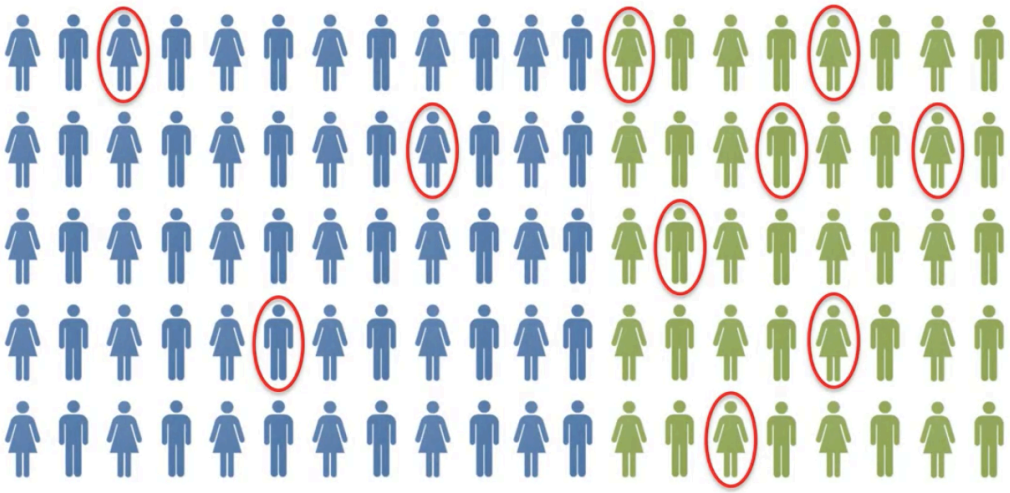
Apriori - Lift

Movie Recommendation: $\text{lift}(M_1 \rightarrow M_2) = \frac{\text{confidence}(M_1 \rightarrow M_2)}{\text{support}(M_2)}$

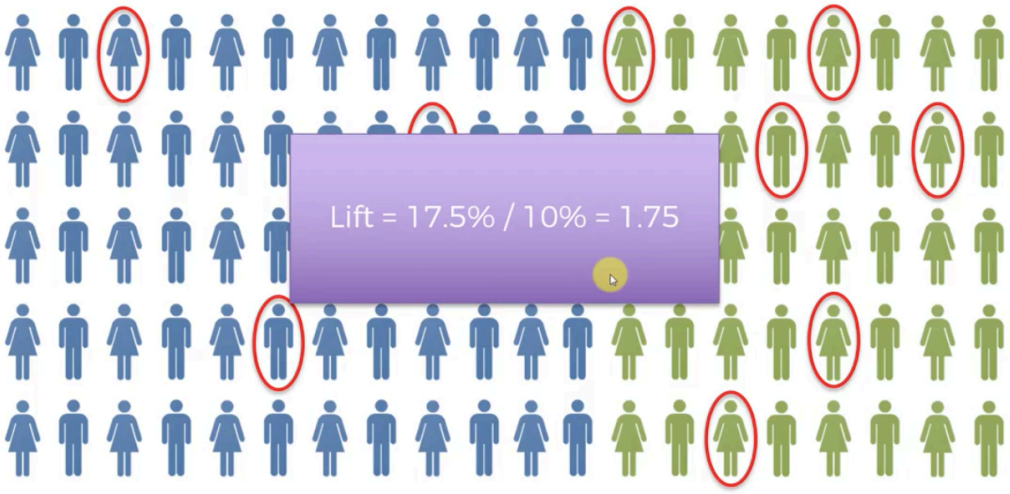
Market Basket Optimisation: $\text{lift}(I_1 \rightarrow I_2) = \frac{\text{confidence}(I_1 \rightarrow I_2)}{\text{support}(I_2)}$

Apriori - Lift

Apriori - Lift



Apriori - Lift



Apriori - Algorithm

Step 1: Set a minimum support and confidence

Step 2: Take all the subsets in transactions having higher support than minimum support

Step 3: Take all the rules of these subsets having higher confidence than minimum confidence

Step 4: Sort the rules by decreasing lift