

Q1 Solution : Recommendation of the Users

User 0: [], {"Captain America":1.000, "Non-Stop":1.000, "The Wolf of Wall Street":1.000}

0.3 0.5 Captain America -> Frozen

0.3 0.5 Non-Stop -> 300 Rise of an Empire

As the confidence value for all the rules pertaining to this user has value 0.5 which states that it is equally likely that both Non-stop and Captain America can be seen by the user and we recommend him Frozen and 300 Rise of the empire respectively

User 1: [], {"Non-Stop":1.000, "300 Rise of an Empire":1.000, "THOR":1.000}

0.3 0.5 Non-Stop -> The Wolf of Wall Street

0.3 0.5 Non-Stop -> Captain America

If the user watches Non-Stop then it is equally likely that he can see The Wolf of Wall Street and Captain America.

User 2: [], {"Captain America":1.000, "Frozen":1.000}

0.3 0.5 Captain America -> The Wolf of Wall Street

0.3 0.5 Captain America -> Non-Stop

If the User 2 watches Captain America, we can recommend him The wolf of Wall Street and Non-Stop equally likely.

User 3: [], {"Captain America":1.000, "Non-Stop":1.000, "300 Rise of an Empire":1.000}

0.3 0.5 Captain America -> Frozen

0.3 0.5 Non-Stop -> The Wolf of Wall Street

0.3 0.5 Captain America -> The Wolf of Wall Street

If user 3 watches captain america, then we can recommend him Frozen or The wolf of wall street with equally likely probability but if user 3 watch Non-stop then we surely recommend The wolf of Wall Street.

User 4: [], {"Captain America":1.000, "The Wolf of Wall Street":1.000, "Frozen":1.000}

0.3 0.7 The Wolf of Wall Street -> Non-Stop

0.3 0.5 Captain America -> Non-Stop

we recommend user 4 to watch Non-stop with high surety if the user 4 has watched The wolf of the Wall Street.

User 5: [], {"Non-Stop":1.000, "The Wolf of Wall Street":1.000}

0.3 0.5 Non-Stop -> 300 Rise of an Empire

0.3 0.7 The Wolf of Wall Street -> Captain America

0.3 0.5 Non-Stop -> Captain America

If user has watched The Wolf of the wall street then with high probability I recommend Captain america else if user has watched Non-Stop then there is equally likely chance of watching 300 Rise of an Empire and Captain America.

Q2 : Solution

For completing the Q2, I have used mushroom data set from UCI. The objective of using this dataset was to analyze the importance of support and confidence value in the python code. For a very low support and confidence, we are getting a lot of tuples close to 521567 whereas with high support and confidence gives us less number of tuples. So how to analysis the association rules. This question helped me in formulating the notion of policy research, a project being done as a part of Data mining by myself and Mr. Abhishek. When analyzing the impact of policies on crime, we need to consider very low support and confidence as the policy is a positive or neutral feature of the sentence and crime is a negative feature of the sentence. Hence a very low confidence will yield better results regarding policy research. Further Application which can be derived merely by replacing the features from the mushroom data set of UCI are :

1. Market Basket Analysis
2. Product Marketing
3. Demand Forecasting
4. Movie Recommendation
5. Music Recommendation (buy youtube)
6. Geotagging of comments