Team37-Report

Raghavendra Reddy (2019101074)

Harshita Upadhyay (2019101126)

Task-1

In this particular task, we are asked to implement Apriori Algorithm OR FP-growth Algorithm and to find out the association rules X->Y from the either one of the algorithms.

Approach – We chose the Apriori Algorithm for this task and implemented it by considering Minimum support and minimum confidence as 0.1.

Task-2

For this task, we were asked to create two lists:
- The first list consists of top-100 association rules sorted based on the support.
- The second list consists of top-100 rules based on confidence.

Then we need to select the rules which are common to both lists and sort the common rules based on confidence.

Approach – We started with one data-frame in which we selected the rules from Apriori and sort them by using confidence and stored it in the dataframe-1. Then we took another data-frame and sorted values by using support. Then, we picked the common rules from both the data-frames 1 and 2 and sorted them.

Task-3

In this task, we need to select association rules of the form X→Y, where X is the movie in the training set for each user in the test set.

Then we need to compute the average precision and average recall by varying the number of rules from 1 to 10 and plot the graphs.

Approach – We got the rules from Apriori Algorithm from task-1. By using them, we are clipping the rows with respect to user set and then we will get the desired test set. The intersection of test set and association rules from Apriori Algorithm will be hit-set. Then using the hit-set, we will compute the average precision and recall and plot the graphs.