Data Mining Homework - Association Analysis

Instructions

Please finish the homework assigned below and submit your homework electronically to Blackboard Learn before the deadline. Any submissions after the deadline will be given a zero grade.

Questions

1. In the following, we have ten market-basket transactions:

Transaction ID	Items Bought
1	{A,B,D,E}
2	{B,C,D}
3	{A,B,D,E}
4	{A,C,D,E}
5	{B,C,D,E}
6	{B,D,E}
7	{C,D}
8	{A,B,C}
9	{A,D,E}
10	{B,D}

   (a) What is the maximum size of frequent itemsets that can be extracted (assuming $minsup = 0$)?

   (b) What is the maximum number of association rules that can be extracted from this data set (including rules that have zero support)?

   (c) Compute the confidence for the association rule $\{A, D\} \rightarrow \{E\}$

   (d) Find all the frequent itemsets assuming $minsup\ count = 2$.

   (e) Find an itemset (of size 2 or larger) that has the largest support.

   (f) Find a pair of items, say x and y, such that the rules $\{x\} \rightarrow \{y\}$ and $\{y\} \rightarrow \{x\}$ have the same confidence.

2. (a) `June 5 class meeting` is scheduled for your course project presentation.

   (b) You will receive questions/comments from the audience of your presentation. Please address them in your final project report.

   (c) The report will be due by `June 12 on Blackboard`.

   (d) Attached there are two papers:

   Whitesides' Group: Writing A Paper: This is very helpful for you to plan and organize your research, presentation, and report.

   Final Project Report Sample: This is a very good example. You should cover similar elements logically in your own report.

   Please confirm that you are clear about the above schedules, and spend sufficient time and efforts in preparing your project presentation and report.