

Assignment Project Exam Help

COMP9318 Tutorial 4: Association Rule Mining

<https://eduassistpro.github.io>

Add WeChat edu_assist_pr

Q1 I

Show that if $A \rightarrow B$ does not meet the minconf constraint, $A \rightarrow BC$ does not either.

Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu_assist_pr

Solution to Q1 I

$$\text{conf}(A \rightarrow BC) = \frac{\text{supp}(ABC)}{\text{supp}(A)}$$

Assignment Project Exam Help

Like A

<https://eduassistpro.github.io>

Add WeChat edu_assist_pr

Q2 I

Given the following transactional database

1	C, B, H
2	B, F, S
3	A, F, G
4	C, B, H
5	B, F, G
6	B, E, O

Assignment Project Exam Help

1. We

a
s

<https://eduassistpro.github.io> (C3 . .)

2. Find all the association rules that involves only B, C, H (in either left or right hand side of the rule). The minimum confidence i

Add WeChat edu_assist_pr

Solution to Q2 I

1. Apriori

1.1 $\text{minsup} = 30\% \times 6 = 1.8$. In other words, the support of a frequent itemset must be no less than 2.

1.2 $C_1 = \{A, B, C, E, F, G, H, O, S\}$, scanning the DB and collect the supports

as

A	B	C	E	F	G	H	O	S
1	5	2	1	3	2	2	1	1

<https://eduassistpro.github.io>

Therefore, $L_2 = \{BC, BF, BH, CH, EG\}$

1.4 C_3 is generated from L_2 by a special enum

The result is $\{BCH\}$. Scan the DB and collect the support as

BCH
2

Therefore, $L_3 = \{BCH\}$.

1.5 C_4 will be the empty set, therefore we stop here.

2. We list the frequent itemsets related to B, C, and H below:

Solution to Q2 II

B	C	H	BC	BH	CH	BCH
5	2	2	2	2	2	2

2.1 For BC, we need to consider candidate rules: $B \rightarrow C$, and $C \rightarrow B$. The former has confidence $\frac{\text{supp}(BC)}{\text{supp}(B)} = 40\%$ and does not meet the minconf requirement. The latter rule has confidence: $\frac{\text{supp}(BC)}{\text{supp}(C)} = 100\%$ and it is qualified.

2.2 It is easy to see that any rule in the form of $B \rightarrow \dots$ will not meet the

<https://eduassistpro.github.io>

- $BC \rightarrow H$ (100%)
- ▶ $BH \rightarrow C$ (100%)
- ▶ $CH \rightarrow B$ (100%)
- ▶ $C \rightarrow BH$ (100%)
- ▶ $H \rightarrow BC$ (100%)

Add WeChat edu_assist_pro

Q3 I

Compute the frequent itemset of for the data in Q2 using the FP-growth algorithm.

Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu_assist_pr

Solution to Q3 I

1. Similar to the first step in Apriori, count the support of all items and *normalize* the original transaction db as follows: (by removing non-frequent items and sort items in the decreasing order of their support)

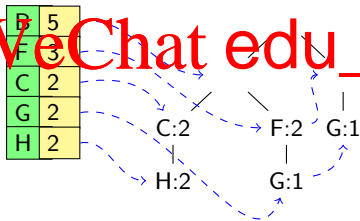
Assignment Project Exam Help

Order	DB				
	B	F	C	G	H
1	B, C, H				
2	B, F				
3	F, G				

<https://eduassistpro.github.io>

2. Construct the FP-tree as:

Add WeChat edu_assist_pro



3. H's conditional pattern base is:

BH : 2

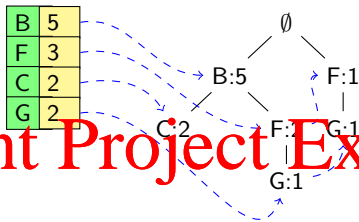
All of the items are frequent, and thus we can output: BH, CH. Construct t



<https://eduassistpro.github.io>

Add WeChat edu_assist_pr

4. We track back and can now safely remove all H nodes from t FP-tree, as shown below.



Assignment Project Exam Help

<https://eduassistpro.github.io>

B F : 1

F : 1

Only F is frequent. We output FS. It is clear that we can stop

5. We track back and can now safely remove all G nodes from t and then process C's conditional pattern base:

B : 2

B is frequent, output BC, and we can stop here.

6. We track back and can now safely remove all C nodes from the FP-tree, and then process F's conditional pattern base:

Solution to Q3 IV

B : 2

B is frequent, output BF, and we can stop here.

7. Since we are left with one item (B) only, we can output stop the whole mining process.

Assignment Project Exam Help

<https://eduassistpro.github.io>

Add WeChat edu_assist_pr