

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/340105166>

Association Rule Mining--Apriori Algorithm Solved Problems

Method · March 2020

DOI: 10.13140/RG.2.2.35174.78407

CITATIONS

0

READS

2,638

1 author:



Mahendra Patil

Atharva College of Engineering, Malad

36 PUBLICATIONS 13 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Heart Beat Sensing Without Physical Contact Using Signal and Image Processing [View project](#)



Covert Communication in TCP/IP [View project](#)

Association Rule Mining--Apriori Algorithm Solved Problems

Q.1) For the following given Transaction Data-set, Generate Rules using Apriori Algorithm.

Consider the values as Support=50% and Confidence=75%

Transaction ID	Items Purchased
1	Bread, Cheese, Egg, Juice
2	Bread, Cheese, Juice
3	Bread, Milk, Yogurt
4	Bread, Juice, Milk
5	Cheese, Juice, Milk

Answer:

Given Support=50% and Confidence=75%

Step 1) Find Frequent Item Set and their support

Item	Frequency	Support (in %)
Bread	4	$4/5=80\%$
Cheese	3	$3/5=60\%$
Egg	1	$1/5=20\%$
Juice	4	$4/5=80\%$
Milk	3	$3/5=60\%$
Yogurt	1	$1/5=20\%$

Support (item) = Frequency of item/Number of transactions

Step 2) Remove all the items whose support is below given minimum support.

Item	Frequency	Support (in %)
------	-----------	----------------

Bread	4	$4/5=80\%$
Cheese	3	$3/5=60\%$
Juice	4	$4/5=80\%$
Milk	3	$3/5=60\%$

Step 3) Now form the two items candidate set and write their frequencies.

Items Pair	Frequency	Support (in %)
Bread, Cheese	2	$2/5=40\%$
Bread, Juice	3	$3/5=60\%$
Bread, Milk	2	$2/5=40\%$
Cheese, Juice	3	$3/5=60\%$
Cheese, Milk	1	$1/5=20\%$
Juice, Milk	2	$2/5=40\%$

Step 4) Remove all the items whose support is below given minimum support.

Items Pair	Frequency	Support (in %)
Bread, Juice	3	$3/5=60\%$
Cheese, Juice	3	$3/5=60\%$

Step 5) Generate rules

For Rules we consider item pairs:

- a) (Bread, Juice)
Bread->Juice and Juice->Bread
- b) (Cheese, Juice)
Cheese->Juice and Juice->Cheese

Confidence (A->B) = support (AUB)/support (A)

Therefore,

1. Confidence (Bread->Juice) = support (Bread U Juice)/support (Bread)
= $3/5 * 5/4 = 3/4 = 75\%$

2. Confidence (Juice->Bread) = support (Juice U Bread)/support (Juice)

= $3/5 * 5/4 = 3/4 = 75\%$

3. Confidence (Cheese->Juice) = support (Cheese U Juice)/support (Cheese)

= $3/5 * 5/3 = 1 = 100\%$

4. Confidence (Juice->Cheese) = support (Juice U Cheese)/support (Juice)

= $3/5 * 5/4 = 3/4 = 75\%$

All the above rules are good because the confidence of each rule is greater than or equal to the minimum confidence given in the problem.

Q.2) For the following given transaction data set, generate rules using Apriori Algorithm. Consider the values as Support=22% and Confidence= 70%

Transaction ID	Items Purchased
1	I1,I2,I5
2	I2,I4
3	I2,I3
4	I1,I2,I4
5	I1,I3
6	I2,I3
7	I1,I3
8	I1,I2,I3,I5
9	I1,I2,I3

Answer:

Given minimum support=22% and confidence=70%

Step 1) Find Frequent Item Set and their support

Item	Frequency	Support (in %)
I1	6	$6/9=66\%$
I2	7	$7/9=80\%$
I3	6	$6/9=66\%$
I4	2	$2/9=22.2\%$
I5	2	$2/9=22.2\%$

Support (item) = Frequency of item/Number of transactions

Step 2) Remove all the items whose support is below given minimum support and form the two items candidate set and write their frequencies.

Item	Frequency	Support (in %)
I1,I2	4	$4/9=44.4\%$
I1,I3	4	$4/9=44.4\%$
I1,I4	1	$1/9=11.1\%$
I1,I5	2	$2/9=22.2\%$
I2,I3	4	$4/9=44.4\%$
I2,I4	2	$2/9=22.2\%$
I2,I5	2	$2/9=22.2\%$
I3,I4	0	$0/9=0\%$
I3,I5	1	$1/9=11.1\%$
I4,I5	0	$0/9=0\%$

Step 3) Remove all the items whose support is below given minimum support and form the three items candidate set and write their frequencies

Items Pair	Frequency	Support (in %)
I1,I2,I3	2	$2/9=22.2\%$
I1,I2,I5	2	$2/9=22.2\%$

Step 4) Generate rules and find the confidence of each rule. If the confidence is below 70% then discard that rule. Keep only the rules whose confidence is equal to or above 70%.

Confidence (A->B) = support (AUB)/support (A)

For the item set (I1,I2,I3) following rules are possible:

Rule	Confidence
(I1,I2)-> I3	$2/9*9/4=2/4=50\%$
(I1,I3)-> I2	$2/9*9/4=2/4=50\%$
(I2,I3)-> I1	$2/9*9/4=2/4=50\%$
I1-> (I2,I3)	$2/9*9/6=2/6=33.33\%$
I2-> (I1,I3)	$2/9*9/7=2/7=28.57\%$
I3->(I1,I2)	$2/9*9/6=2/6=33.33\%$

For the item set (I1,I2,I5) following rules are possible:

Rule	Confidence
(I1,I2)-> I5	$2/9*9/4=2/4=50\%$
(I1,I5)-> I2	$2/9*9/2=2/2=100\%$

(I2,I5)-> I1	$2/9 * 9/2 = 2/2 = 100\%$
I1-> (I2,I5)	$2/9 * 9/6 = 2/6 = 33.33\%$
I2-> (I1,I5)	$2/9 * 9/7 = 2/7 = 28.57\%$
I5->(I1,I2)	$2/9 * 9/2 = 2/2 = 100\%$

From above rules only following rules are valid (because it's confidence $\geq 75\%$):

(I1,I5)-> I2

(I2,I5)-> I1

I5->(I1,I2)

Q.3) Assume min. Support ort = 40%

TID	Items
01	A,B,D,E
02	B,C,D
03	A,B,E
04	B,D
05	A,B,C,E

Marks: 10 M

Year: Dec 2015

Ans:

Support count = 40%

$x/5 * 100 = 50$

$x = 3$

Step 1: Generating 1-itemset frequent pattern

Scan D for count of each candidate

C1

Item set	Support count
{A}	3
{B}	5
{D}	3
{C}	2
{E}	3

Compare candidate support count with minimum support count

L1

Item set	Support count
{A}	3
{B}	5
{D}	3
{E}	3

Step 2: Generate C2- item set Frequent Pattern

Generate C2 candidate from L1

C2

Item set	Support count
{A,B}	3
{A,D}	1
{A,E}	3
{B,D}	3
{B,E}	3

Compare candidate support count with minimum support count

L2

Item set
{A,B}
{A,E}
{B,D}
{B,E}

Step 3: Generating 3- item set Frequent Pattern

C3

Item set	Support count
{A,B,E}	2
{A,B,D}	1
{A,E,B,D}	1

Compare candidate support count with minimum support count.

As the support count generated is less than minimum support count. So, there is no item set with minimum support count.