# RIORI ALGORITHM.

Min. Support Count = 2

| TID | Items |
|---|---|
| $T_1$ | 1,3,4 |
| $T_2$ | 2,3,5 |
| $T_3$ | 1,2,3,5 |
| $T_4$ | 2,5 |
| $T_5$ | 1,3,5 |

→

C1 (1-itemset)

| Item | Support |
|---|---|
| {1} | 3 |
| {2} | 3 |
| {3} | 4 |
| {4} | 1 |
| {5} | 4 |

→

L1

| Item | Support |
|---|---|
| {1} | 3 |
| {2} | 3 |
| {3} | 4 |
| {5} | 4 |

←

C2 (2-itemset)

| | |
|---|---|
| {1,2} | 1 |
| {1,3} | 3 |
| {1,5} | 2 |
| {2,3} | 2 |
| {2,5} | 3 |
| {3,5} | 3 |

→

L2

| | |
|---|---|
| {1,3} | 3 |
| {1,5} | 2 |
| {2,3} | 2 |
| {2,5} | 3 |
| {3,5} | 3 |

↙

(Pruning) C3 (3-itemset)

| | |
|---|---|
| {1,2,3} {1,2} {1,3} {2,3} | No |
| {1,2,5} {1,2} {1,5} {2,5} | No |
| {1,3,5} {1,3} {3,5} {1,3} | Yes |
| {2,3,5} {2,3} {2,5} {3,5} | Yes |

→

L3

| | |
|---|---|
| {1,3,5} | 2 |
| {2,3,5} | 2 |

↙

C4 (4-itemset)

| | |
|---|---|
| {1,2,3,5} | 1 |

→

L4 Level (Rejected)

| |
|---|
| {1,2,3,5} |

# Rule framing:

| Itemset | Support |
|---|---|
| {1,3,5} | 2 |
| {2,3,5} | 2 |

Subset - S

for I = {1,3,5}, subsets are {1,3} {1,5} {3,5} {1,} {3} {5}

for I = {2,3,5}, subsets are {2,3} {2,5} {3,5} {2} {3} {5}

for every subset S of I, output the rule:

S → (I - S) (S recommends I-S) if

Support(I) / Support(S) >= min_confidence Value.

Applying rules: min-confidence Value = 60%

Rule 1 : {1,3} → ({1,3,5} - {1,3})

Confidence = support(1,3,5) / support(1,3)
= 2/3 = 66.66% > 60% ✓

Rule 2 : {1,5} → ({1,3,5} - {1,5})

Confidence = Support(1,3,5) / Support(1,5)
= 2/2 = 100% > 60% ✓

Rule 3 : {3,5} → ({1,3,5} - {3,5})

Confidence = Support(1,3,5) / support(3,5)
= 2/3 = 66.66% ≥ 60% ✓

Rule 4 : {1} → ({1,3,5} - {1})

Confidence = Support(1,3,5) / support(1)
= 2/3 = 66.66% > 60% ✓

Rule 5 : {3} → ({1,3,5} - {3})

Confidence = Support(1,3,5) / support(3)
= 2/4 = 50% < 60% ✗

Find all the frequent item sets using Apriori algorithm where the Min. support count is 2 in the following table.

| TID | List of items. |
|---|---|
| $T_{100}$ | $I_1, I_2, I_5$ |
| $T_{200}$ | $I_2, I_4$ |
| $T_{300}$ | $I_2, I_3$ |
| $T_{400}$ | $I_1, I_2, I_4$ |
| $T_{500}$ | $I_1, I_3$ |
| $T_{600}$ | $I_2 I_3$ |
| $T_{700}$ | $I_1, I_3$ |
| $T_{800}$ | $I_1, I_2, I_3, I_5$ |
| $T_{900}$ | $I_1, I_2, I_3$ |

Example for Apriori Algorithm:

Min Support Count = 2.

Scan D for Count of each candidate →

$C_1$ (1-item)

| Item | Support |
|---|---|
| {I1} | 6 |
| {I2} | 7 |
| {I3} | 6 |
| {I4} | 2 |
| {I5} | 2 |

→ $L_1$

| Item | Support |
|---|---|
| {I1} | 6 |
| {I2} | 7 |
| {I3} | 6 |
| {I4} | 2 |
| {I5} | 2 |

→ $C_2$

| Item | Support |
|---|---|
| I1 I2 | 4 |
| I1 I3 | 4 |
| I1 I4 | 1 |
| I1 I5 | 2 |
| I2 I3 | 4 |
| I2 I4 | 2 |
| I2 I5 | 2 |
| I3 I4 | 0 |
| I3 I5 | 1 |
| I4 I5 | 0 |

→ $L_2$

| | |
|---|---|
| I1 I2 | 4 |
| I1 I3 | 4 |
| I1 I5 | 2 |
| I2 I3 | 4 |
| I2 I4 | 2 |
| I2 I5 | 2 |

→ $C_3$

| | |
|---|---|
| I1 I2 I3 | 2 |
| I1 I2 I5 | 2 |

→ $L_3$

| |
|---|
| I1 I2 I3 |
| I1 I2 I5 |

→ $C_4$

| Item | Support Count |
|---|---|
| I1 I2 I3 I5 | 1 → ✗ (Less than Min support count) |

Min Confidence Value : 50% ✓

Confidence :-

Rule :1 $I_1 I_2 \rightarrow I_3$ [$I_1 I_2 I_3 - I_1 I_2$] = Support $(I_1 I_2 I_3)$ / Support $(I_1 I_2)$ = 2/4 = 50% ✓

:2 $I_1 I_3 \rightarrow I_2$ [$I_1 I_2 I_3 - I_1 I_3$] = Support $(I_1 I_2 I_3)$ / Support $(I_1 I_3)$ = 2/4 = 50% ✓

:3 $I_2 I_3 \rightarrow I_1$ [$I_1 I_2 I_3 - I_2 I_3$] = Support $(I_1 I_2 I_3)$ / Support $(I_2 I_3)$ = 2/4 = 50% ✓

:4 $I_1 \rightarrow I_1 I_2 I_3 - I_1$ = Support $(I_1 I_2 I_3)$ / Support $(I_1)$ = 2/6 = 0.33 = 33%

:5 $I_2 \rightarrow I_1 I_2 I_3$ = 2/7 = 28%

:6 $I_3 \rightarrow I_1 I_2 I_3$ = 2/6 = 33%

$$\text{Confidence}(A \Rightarrow B) = P(B|A) = \frac{\text{Support count}(A \cup B)}{\text{Support count}(A)}$$

Resulting Association rules for Itemset $\{I_1, I_2 I_5\}$ are:

$\{I_1, I_2\} \Rightarrow I_5$ Confidence = 2/4 = 50% ✓

$\{I_1, I_5\} \Rightarrow I_2$ Confidence = $\frac{\text{Support count}(I_1 I_2 I_5)}{\text{Support count}(I_1 I_5)} = \frac{2}{2} = 100\%$ ✓

$\{I_2 I_5\} \Rightarrow I_1$ Confidence = $\frac{SC(I_1 I_2 I_5)}{SC(I_2 I_5)} = \frac{2}{2} = 100\%$ ✓

$\{I_1\} \Rightarrow \{I_2 I_5\}$ = 2/6 = 33%

$\{I_2\} \Rightarrow \{I_1, I_5\}$ = 2/7 = 29%

Framing the Rules:

Rule: 1 $(I_1, I_5)$

If customer buys $(I_1 I_5)$ then customer buys $I_3$.

Rule: 2

If Customer buys $(I_1, I_3)$ then Customer buys $I_2$

Rule: 3

If Customer buys $(I_2 I_3)$ then Customer buys $I_1$

Rule: 4

If Customer buys $I_1, I_2$ then $I_5$ is also bought

Rule: 5

If Customer buys $I_1, I_5$ then $I_2$ is also bought

Rule: 6

If Customer buys $I_2 I_5$ then $I_1$ is also bought

Rule: 7

If Customer buys $I_5$ then Customer also buys $I_1$ & $I_2$.

---