A-Priori Algorithm

Association Rule Learning
Apriori Intuition

ARL - What is it all about ?

ARL - What is it all about ?

People who bought also bought ...

ARL - Movie Recommendation

User ID	Movies liked
46578	Movie1, Movie2, Movie3, Movie4
98989	Movie1, Movie2
71527	Movie1, Movie2, Movie4
78981	Movie1, Movie2
89192	Movie2, Movie4
61557	Movie1, Movie3

Potential Rules:

Movie1 ⟹ Movie2

Movie2 ⟹ Movie4

Movie1 ⟹ Movie3

ARL - Market Basket Optimisation

Transaction ID	Products purchased
46578	Burgers, French Fries, Vegetables
98989	Burgers, French Fries, Ketchup
71527	Vegetables, Fruits
78981	Pasta, Fruits, Butter, Vegetables
89192	Burgers, Pasta, French Fries
61557	Fruits, Orange Juice, Vegetables
87923	Burgers, French Fries, Ketchup, Mayo

Potential Rules:

Burgers ⟹ French Fries

Vegetables ⟹ Fruits

Burgers, French Fries ⟹ Ketchup

Apriori - Support

Movie Recommendation: $\text{support}(\boldsymbol{M}) = \dfrac{\#\ \text{user watchlists containing } \boldsymbol{M}}{\#\ \text{user watchlists}}$

Market Basket Optimisation: $\text{support}(\boldsymbol{I}) = \dfrac{\#\ \text{transactions containing } \boldsymbol{I}}{\#\ \text{transactions}}$

Apriori - Support

2

5/23/2025
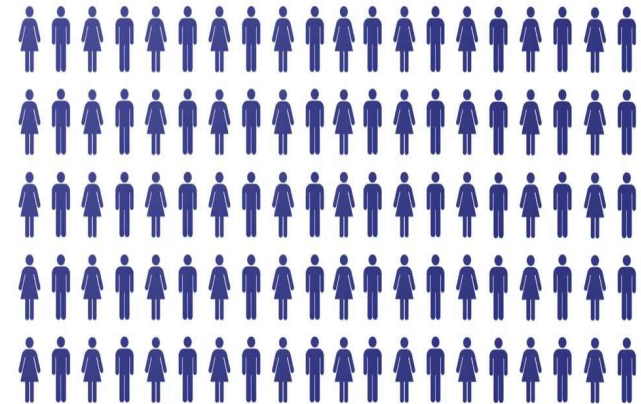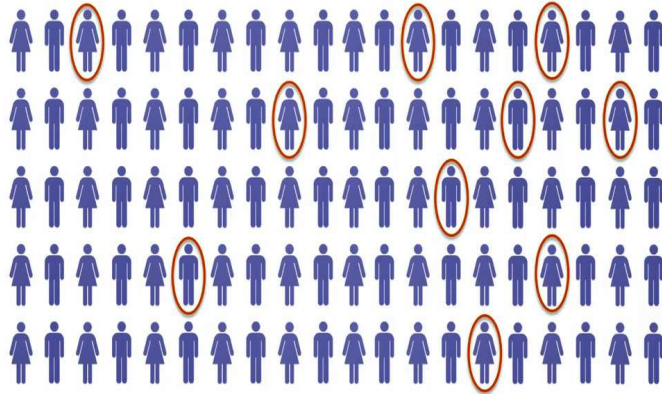
Apriori - Support

Apriori - Support
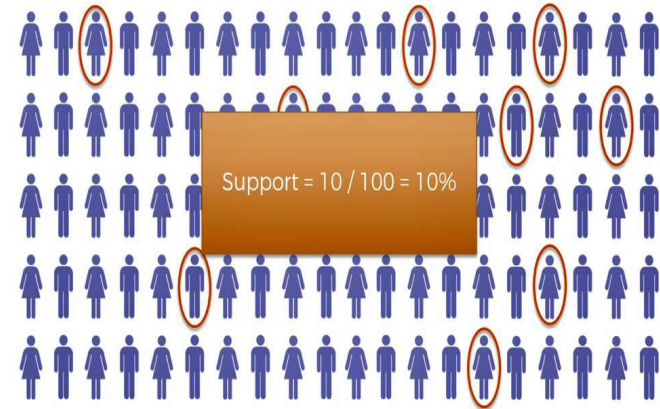
Support = 10 / 100 = 10%

## Apriori - Confidence

Movie Recommendation: $\text{confidence}(M_1 \rightarrow M_2) = \dfrac{\#\ \text{user watchlists containing } M_1 \text{ and } M_2}{\#\ \text{user watchlists containing } M_1}$

Market Basket Optimisation: $\text{confidence}(I_1 \rightarrow I_2) = \dfrac{\#\ \text{transactions containing } I_1 \text{ and } I_2}{\#\ \text{transactions containing } I_1}$
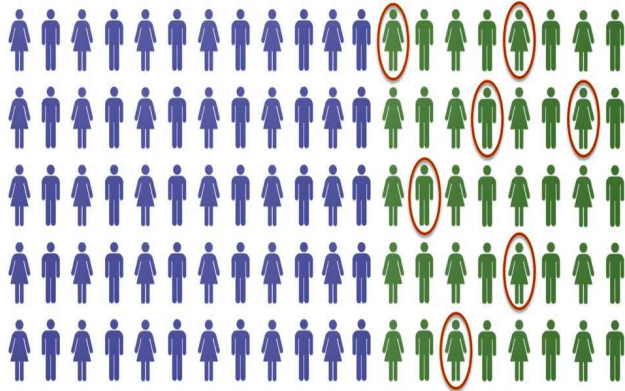
Apriori - Confidence

3

Apriori - Confidence

Apriori - Confidence

Confidence = 7 / 40 = 17.5%
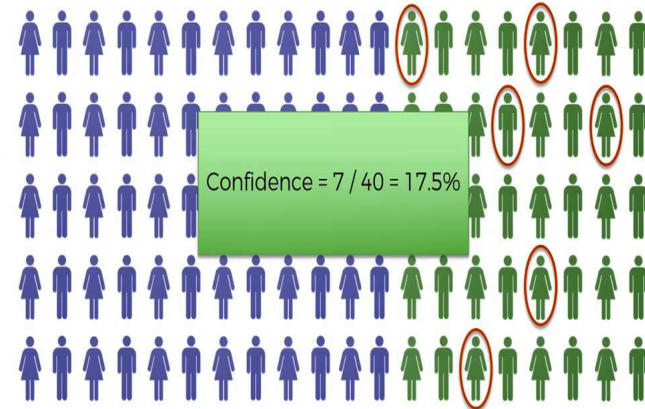
Apriori - Lift

Movie Recommendation: $\text{lift}(M_1 \to M_2) = \dfrac{\text{confidence}(M_1 \to M_2)}{\text{support}(M_2)}$

Market Basket Optimisation: $\text{lift}(I_1 \to I_2) = \dfrac{\text{confidence}(I_1 \to I_2)}{\text{support}(I_2)}$

Apriori - Lift

Lift = 17.5% / 10% = 1.75

4
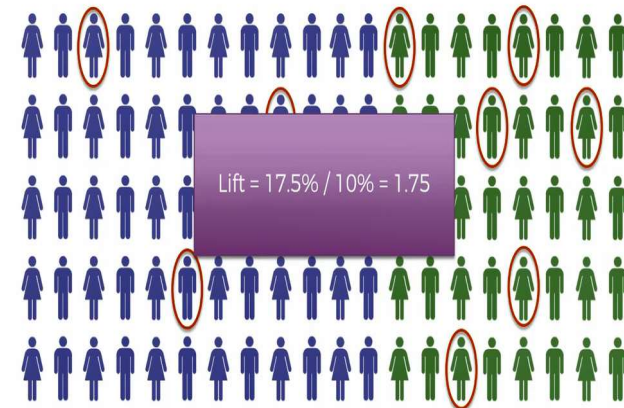
Apriori - Algorithm

Step 1: Set a minimum support and confidence

Step 2: Take all the subsets in transactions having higher support than minimum support

Step 3: Take all the rules of these subsets having higher confidence than minimum confidence

Step 4: Sort the rules by decreasing lift

- **T**he Apriori Algorithm is an algorithm for mining frequent itemsets for Boolean association rules.

- **A**priori uses a "bottom up" approach, where frequent subsets are extended one item at a time which is known as candidate generation, and groups of candidates are tested against the data.

Example

Design the association rules for the given table using Apriori Algorithm.

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Support = 2
Confidence = 50%

Candidate C1

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Itemset	Support
{A}	3
{B}	2
{C}	2
{D}	1
{E}	1
{F}	1

Support = 2
Confidence = 50%

Slide 21

Candidate C1

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Itemset	Support
{A}	3
{B}	2
{C}	2
{D}	1
{E}	1
{F}	1

Support = 2
Confidence = 50%

Slide 22

Candidate C1

L1 [Frequent Itemset]

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Itemset	Support
{A}	3
{B}	2
{C}	2

Support = 2
Confidence = 50%

Slide 23

Candidate C2

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Itemset	Support
{A}	3
{B}	2
{C}	2

Itemset	Support
{A,B}	1
{B,C}	1
{A,C}	2

Support = 2
Confidence = 50%

Slide 24

Candidate C2

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Itemset	Support
{A}	3
{B}	2
{C}	2

Itemset	Support
{A,B}	1
{B,C}	1
{A,C}	2

Support = 2
Confidence = 50%

Slide 25

Candidate C2

Transaction ID	Itemsets
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Itemset	Support
{A}	3
{B}	2
{C}	2

L2 [Frequent Itemset]

Itemset	Support
{A,C}	2

Support = 2
Confidence = 50%

Slide 26

Candidate C2

Itemset	Support
{A,C}	2

Associative Rule	Support	Confidence	Confidence %
A -> C	2	2/3 = 0.66	66%
C -> A	2	2/2 = 1	100%

Support = 2
Confidence = 50%

Example 2

- **A** database has 5 transactions
- **M**inimum support = 60%
- **M**inimum confidence = 80%

T_ID	Itemsets
T_1000	M,O,N,K,E,Y
T_1001	D,O,N,K,E,Y
T_1002	M,A,K,E
T_1003	M,U,C,K,Y
T_1004	C,O,O,K,E

Slide 28

- **Candidate C1**

Itemset	Support
{M}	3
{O}	3
{N}	2
{K}	5
{E}	4
{Y}	3
{D}	1
{A}	1
{U}	1
{C}	2

Slide 29

- L1

Itemset	Support
{M}	3
{O}	3
{N}	2
{K}	5
{E}	4
{Y}	3
{D}	1
{A}	1
{U}	1
{C}	2

Slide 30

- L1

Itemset	Support
{M}	3
{O}	3
{K}	5
{E}	4
{Y}	3

Slide 31

- Candidate C2

Itemset	Support
{M,O}	1
{M,K}	3
{M,E}	2
{M,Y}	2
{O,K}	3
{O,E}	3
{O,Y}	2
{K,E}	4
{K,Y}	3
{E,Y}	2

Slide 32

- L2

Itemset	Support
{M,O}	1
{M,K}	3
{M,E}	2
{M,Y}	2
{O,K}	3
{O,E}	3
{O,Y}	2
{K,E}	4
{K,Y}	3
{E,Y}	2

8

- **L2**

Itemset	Support
{M,K}	3
{O,K}	3
{O,E}	3
{K,E}	4
{K,Y}	3

- **Candidate C3**

Itemset	Support
{M,O,K}	1
{M,K,O,E}	1
{M,K,E}	2
{M,K,Y}	2
{O,K,E}	3
{O,K,Y}	2
{O,K,E,Y}	2
{K,E,Y}	2

- **L3**

Itemset	Support
{M,O,K}	1
{M,K,O,E}	1
{M,K,E}	2
{M,K,Y}	2
{O,K,E}	3
{O,K,Y}	2
{O,K,E,Y}	2
{K,E,Y}	2

- **L3**

Itemset	Support
{O,K,E}	3

- **Association rules**

Association Rules	Support	Confidence	Confidence %
O → K^E	3	3/3 = 1	100
K → O^E	3	3/5 = 0.6	60
E → K^O	3	3 / 4 = 0.75	75

Slide 37

- **A**ssociation rules

The minimum confidence is 80%

Association Rules	✚ Support	Confidence	Confidence %
O → K^E	3	3/3 = 1	100
K → O^E	3	3/5 = 0.6	60
E → K^O	3	3 / 4 = 0.75	75

O → K^E = ACCEPTED

37

Slide 38

Design association rules for a database has 9 transactions

Minimum support = 2

Minimum confidence = 70% ✓

ID	Itemsets
100	1,2,5
200	2,4
300	2,3
400	1,2,4
500	1,3
600	1,3
700	✚ 1,3,2,5
800	1,3
900	1,2,3

38

Slide 39

- **C**andidate C1

Itemset	Support
{1}	✚ 7
{2}	6
{3}	6
{4}	2
{5}	2

39

Slide 40

- **C**andidate C2

Itemset	Support
{1,2}	3
{1,3}	5
{1,4}	1
{1,5}	2
{2,3}	3
{2,4}	2
{2,5}	2
{3,4}	-
{3,5}	1
{4,5}	-

40

L2

Itemset	Support
{1,2}	3
{1,3}	5
{1,4}	1
{1,5}	2
{2,3}	3
{2,4}	2
{2,5}	2
{3,4}	-
{3,5}	1
{4,5}	-

41

L2

Itemset	Support
{1,2}	3
{1,3}	5
{1,5}	2
{2,3}	3
{2,4}	2
{2,5}	2

42

Candidate C3

Itemset	Support
{1,2,3}	2
{1,2,5}	2
{1,2,4}	1
{1,3,5}	1
{1,2,3,4}	-
{1,2,4,5}	1
{1,2,3,5}	-
{2,3,4}	-
{2,3,5}	1
{2,4,5}	-

43

L3

Itemset	Support
{1,2,3}	2
{1,2,5}	2
{1,2,4}	1
{1,3,5}	1
{1,2,3,4}	-
{1,2,4,5}	1
{1,2,3,5}	-
{2,3,4}	-
{2,3,5}	1
{2,4,5}	-

44

- **L3**

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

- Candidate C4

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

- To design association rule

- Association rules

Association Rules	Support	Confidence	Confidence %
1 → 2^3	2	2/7= 0.28	28
2 → 1^3	2	2/6= 0.33	33
3 → 1^2	2	2/6= 0.33	33
1 → 2^5	2	2/7= 0.28	28
2 → 1^5	2	2/6= 0.33	33
5 → 1^2	2	2/2=1.00	100

The minimum confidence is 70%