Affinity Analysis.

— Co-occurances. → Medical: Symptom1 (obesity), Symptom2 (Heart dis)

Bread, Butter, Eggs. — 'things that go together'

Laptop, Antivirus. — 'Market Basket Analysis'

Mobile, Cover, Screen guard. — 'Association Rule Mining'.

Association Rule:

If {Bread} → {Butter, Eggs}.
Antecedent Consequent

If {laptop → Antivirus}.

Ip: Set of transaction-

Op: 'Actionable Rules'.
↳ 'Meaningful'

If {Antecedent} then {Consequent}
(A) (C)

Prod A x x x Prd B.

Bud. End

→ Product placement.
→ Offers.

Metrics.

If {A} → {C}.
Antecedent → Consequent.

(a) Support (s) = $\frac{\#(A, C)}{\#(\text{total transactions})}$

→ Frequency

If. {Bread} → {Butter}.

10 → transaction
8 → A, C.

$s = \frac{8}{10} = 80\%$.

(b) Confidence.

(c) = $\frac{\#(A, C)}{\#(A)}$

C. = $\frac{6.\ (\text{Butter also})}{8\ (\text{Bread})}$

= 75%.

(c) Lift: ⇒ $\frac{\text{Confidence}}{\text{Benchmark confidence}}$

↳ $\frac{(\#C)}{\text{total no. of transaction}}$

Apriori Algorithm -

Item Set → set of items.

1-item itemset = {Bread}.

2-item itemset = {Bread, Jam}.

3-item itemset = {Bread, Jam, Eggs}.

Apriori Rule:

"If an itemset is 'frequent', then all its subsets should also be frequent"

"Frequent Itemset" → Support should meet a min-threshold value.

→ Min-support (70%)

Support = $\frac{1}{100}$.

⇒ Frequency

⇒ 1%

If {Bread} Then {Jam}. Min-support = 70%.

10.05 am.

{Bread, Jam} → support -

Support ≥ Min-support -

↳ Frequent Itemset.

Subsets.

{Bread, Butter, Jam} →

↳ Frequent Item set

{Butter} → Frequent

{Jam} → "

{Bread} → "

{Bread, Butter} → "

{Butter, Jam} → "

{Jam, Bread} → "

Min_support = 2 (2/9)

Tid	Item list
1	$I_1 I_2 I_5$
2	$I_2 I_4$
3	$I_2 I_3$
4	$I_1 I_2 I_4$
5	$I_1 I_3$
6	$I_2 I_4$
7	$I_1 I_4$
8	$I_1 I_2 I_3 I_5$ ✓ 1
9	$I_1 I_2 I_3$ — 1

1 item itemset

$\{I_1\}$ — 6 ✓

$\{I_2\}$ — 7

$\{I_3\}$ — 4

$\{I_4\}$ — 4

$\{I_5\}$ — 2

2-item itemsets

$I_1 I_2$ — 4

$I_1 I_3$ — 3

$I_1 I_4$ — 2

$I_1 I_5$ — 2

$I_2 I_3$ — 3

$I_2 I_4$ — 3

$I_2 I_5$ — 2

~~$I_3 I_4$ — 0~~ ✓

~~$I_3 I_5$ — 1~~ ✓

~~$I_4 I_5$ — 0~~

3-item item

$I_1 I_2 I_3 \rightarrow 2$ ✓

$I_1 I_2 I_4 \rightarrow 1$ ✗

$I_1 I_2 I_5 \rightarrow 2$ ✓

$I_1 I_3 I_4 \rightarrow 0$ ✓

$I_1 I_3 I_5 \rightarrow 1$ ✗

$I_1 I_4 I_5 \rightarrow 0$

$I_2 I_3 I_4 \rightarrow 0$

$I_2 I_3 I_5 \rightarrow 1$

$I_2 I_4 I_5 \rightarrow 0$

freq. 3 item itemset

$\{I_1 I_2 I_3\} \rightarrow 2$

$\{I_1 I_2 I_5\} \rightarrow 2$

↳ Frequent itemset

Four-item itemset

$I_1 I_2 I_3 I_5 \rightarrow 1$

Min. confidence = 60%

$\{I_1\ I_2\ I_5\} \rightarrow 2$

$\{I_1\ I_2\ I_3\}$

A C

$I_1 \{I_1 \rightarrow I_2\ I_3\} = \frac{\{I_1\ I_2\ I_3\}}{\#\{I_1\}} = \frac{2}{6} = 33.33\%$

$\{I_2 \rightarrow I_1\ I_3\} = 2/7 = 28.5\%$

$\{I_3 \rightarrow I_1\ I_2\} = 2/4 = 50\%$

$\{I_2\ I_3 \rightarrow I_1\} \Rightarrow 2/3 = 66.66\%$

$\{I_1\ I_3 \rightarrow I_2\} \Rightarrow 2/3 = 66.66\%$

$\{I_1\ I_2 \rightarrow I_3\} \Rightarrow 2/4 \Rightarrow 50\%$

Valid Rules.

$\therefore I_2 I_3 \Rightarrow I_1 \rightarrow$ ✓

$I_1 I_3 \Rightarrow I_2 \rightarrow$

	Confidence	Lift
$I_2 I_3 \rightarrow I_1$	66.66%	1.5
$\Rightarrow I_1 I_3 \rightarrow I_2$	66.66%	2.3

Consequent → Very frequ.
→ High Confidence

⇒ 84 trans.
80 → Milk.
14 → tooth brush.

If {tooth Brush} → {Milk}.
A
Infrequent
Frequent

Confidence = 10 / 14

⇒ 0.71.
⇒ 71 %.

Lift = Confidence / Benchmark confidence.
↳ Support for 'C'.

$$= \frac{(\#A, C) / (\#A)}{(\#C) / \#(\text{total transactions})}.$$

Benchmark Conf. = 80/84
= 0.95.

$$\text{Lift} = \frac{0.71}{0.95}.$$

Lift ⇒ 0.74.

Lift ≥ 1
2, 3,

1. Apriori Algorithm

1. Create 1-item itemsets.
2. Find frequent 1-item itemsets based on Min-support.
3. Create 2-item itemsets from 'Frequent' 1-item itemset.
4. Find frequent 2-item itemsets.
5. Find 3-item itemsets from 'frequent' 2 item itemsets.
6. Find frequent 3-item itemsets.

⋮

2. Find the Confidence for rules created above.
3. Find the lift for those rules.