

MARKET BASKET ANALYSIS (MBA)

Given transaction table:-

Transaction ID	Item Bought
1	{milk, kaya, soya}
2	{bread, butter, milk}
3	{milk, soya, cookies}
4	{bread, butter, cookies}
5	{kaya, cookies, soya}
6	{milk, soya, bread, butter}
7	{bread, butter, soya}
8	{kaya, soya}
9	{milk, soya, bread, butter}
10	{kaya, cookies}

Item = milk, kaya, soya, bread, butter, cookies = 6

- Steps**
1. List item
 2. do frequency board.

do frequency board:

ID	MILK	KAYA	SOYA	BREAD	BUTTER	COOKIES
1	1	1	1	0	0	0
2	1	0	0	1	1	0
3	1	0	1	0	0	1
4	0	0	0	1	1	1
5	0	1	1	0	0	1
6	1	0	1	1	1	0
7	0	0	1	1	1	0
8	0	1	1	0	0	0
9	1	0	1	1	1	0
10	0	1	0	0	0	1

Layman: How many times X appears with in total transactions N

If support high \rightarrow eliminate that the rule did not simply occur by chance

Support \rightarrow how often the rule is applicable in a given dataset.

conf \rightarrow how frequently items in X appear in the transactions that contain X

Find the support for 1 size, 2 size, & 3 size itemsets.

1 size =

- milk = 5
- soya = 7
- butter = 5
- kaya = 4
- bread = 5
- cookies = 4

2 size =

{milk, kaya} = 1

{milk, soya} = 4

{kaya, soya} = 3

3 size =

{milk, kaya, soya} = 1

{bread, butter, milk} = 3

* total bilangan yang keluar sekali \rightarrow support count

kalah 1 size = gabung 1 item

kalah 2 size = gabung 2 item

kalah 3 size = gabung 3 item

3. Find the largest support for 2 size. (gabungan 2 item yg keluar paling banyak)

$$\{bread, butter\} = 5\#$$

4. Find the count for min conf ≥ 0

$$\{bread, butter\} \rightarrow \{soya\}$$

$$\frac{bread, butter, soya}{bread, butter} = \frac{3}{5} = 0.6\#$$

bilangan bread, butter, soya keluar sekali

bilangan bread, butter keluar sekali

5. What is max no of association rules that can be expected?

$$R = 3^d - 2^{d+1} + 1 \quad d: \text{no of item}$$

$$R = 3^d - 2^{d+1} + 1$$

$$= 3^6 - 2^{6+1} + 1$$

$$= 3^6 - 2^7 + 1$$

$$= 602\#$$

6. What is max * of size 3 itemset from the table?

$$\frac{n!}{k! (n-k)!}$$

n: total item

k: size yang diminta.

$$\frac{n!}{k! (n-k)!} = \frac{6!}{3! (6-3)!}$$

$$= \frac{6!}{3! (3!)}$$

$$= \frac{1 \times 2 \times 3 \times 4 \times 5 \times 6}{(1 \times 2 \times 3) (1 \times 2 \times 3)}$$

$$= \frac{4 \times 5 \times 6}{1 \times 2 \times 3} = 20\#$$

7. What is max size of frequent itemsets that can be extracted?

* size gabungan item yang paling besar.

$$\{milk, soya, bread, butter\} = \text{gabungan 4 itemset}\#$$

d the confident of associate rule of:-

$\{ \text{bread, butter} \} \rightarrow \text{milk}$

$$\frac{\text{bread, butter, milk}}{\text{bread, butter}} = \frac{3}{5} = 0.6 \quad \#$$

9. Find the pair of items a & b, such that the rules $\{a\} \rightarrow \{b\}$ & $\{b\} \rightarrow \{a\}$ have the same confident

$$\{ \text{bread, butter} \} \rightarrow \text{milk} = \frac{3}{5} = 0.6 \quad \#$$

$$\{ \text{milk} \} \rightarrow \{ \text{bread, butter} \} = \frac{3}{5} = 0.6 \quad \#$$

10. Support + count

cth: $\{ \text{milk, kaya} \} = 1$ (sebab ada 1 transaction shj)

so, $\frac{1}{10}$ (total transaction) *

11. confident

cth: $\{ \text{milk, diapers} \} \rightarrow \{ \text{beer} \}$

$$\text{Support} = \frac{\text{milk, diapers, beers}}{\text{milk, diapers}} = \frac{3}{5} = 0.6 \quad \#$$

so, confident $\frac{\{ \text{milk, diapers, beers} \}}{\{ \text{milk, diapers} \}}$ support count 2 kali

support count 3 kali

$$= \frac{2}{3}$$
$$= 0.67 \quad \#$$

p/s: with no ⑩ dan ⑪ contoh lain.

* Thank you Naslaa Sadig (Msc. candidate 2013/2014) for compiling the answer from the white board. You've got superb eyes!