

LATIHAN 1 APRIORI KONSEP APLIKASI DATA MINING

1. STUDI KASUS

T1	{roti, selai, mentega}
T2	{roti, mentega}
T3	{roti, susu, mentega}
T4	{coklat, roti}
T5	{coklat, susu}

- Suatu supermarket mempunyai sejumlah transaksi seperti dalam tabel
- Buatlah Association rule dari data tersebut dengan cara menghitung support dan confidence untuk semua itemset
- Jika minimum support = 0,3 dan confidence = 0,8, maka bagaimana kaidah asosiasinya?

2. PENYELESAIAN

Transaksi	A	B	C	D	E
T1	1	1	1	0	0
T2	1	0	1	0	0
T3	1	0	1	1	0
T4	1	0	0	0	1
T5	0	0	0	1	1

Keterangan :

A = Roti

D = Susu

B = Selai

E = Coklat

C = Mentega

2.1. Association Rule

a. Itemset A

- Jumlah Transaksi $|D| = 5$
- Kemunculan item A pada transaksi ($|T_A|$) sebanyak 4 kali, yaitu pada T1, T2, T3, T4
- $\text{Support}(A) = \frac{|T_A|}{|D|} = \frac{4}{5} = 0,8$

b. Itemset B

- Jumlah Transaksi $|D| = 5$
- Kemunculan item B pada transaksi ($|T_B|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(B) = \frac{|T_B|}{|D|} = \frac{1}{5} = 0,2$

c. Itemset C

- Jumlah Transaksi $|D| = 5$
- Kemunculan item C pada transaksi ($|T_C|$) sebanyak 3 kali, yaitu pada T1, T2, T3
- $\text{Support}(C) = \frac{|T_C|}{|D|} = \frac{3}{5} = 0,6$

d. Itemset D

- Jumlah Transaksi $|D| = 5$
- Kemunculan item D pada transaksi ($|Td|$) sebanyak 2 kali, yaitu pada T3, T5
- $\text{Support}(D) = \frac{|Td|}{|D|} = \frac{2}{5} = 0,4$

e. Itemset E

- Jumlah Transaksi $|D| = 5$
- Kemunculan item E pada transaksi ($|Te|$) sebanyak 2 kali, yaitu pada T4, T5
- $\text{Support}(E) = \frac{|Te|}{|D|} = \frac{2}{5} = 0,4$

f. Itemset AB

- Jumlah Transaksi $|D| = 5$
- Kemunculan item AB pada transaksi ($|Tab|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(AB) = \frac{|Tab|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } A \rightarrow B = \frac{|Tab|}{|Ta|} = \frac{1}{4} = 0,25$
- $\text{Lift } (A \rightarrow B) = \frac{\text{Confidence } A \rightarrow B}{\text{Support}(AB)} = \frac{0,25}{0,2} = 1,25$

g. Itemset AC

- Jumlah Transaksi $|D| = 5$
- Kemunculan item AC pada transaksi ($|Tac|$) sebanyak 3 kali, yaitu pada T1, T2, T3
- $\text{Support}(AC) = \frac{|Tac|}{|D|} = \frac{3}{5} = 0,6$
- $\text{Confidence } A \rightarrow B = \frac{|Tabc|}{|Ta|} = \frac{1}{4} = 0,25$
- $\text{Lift } (A \rightarrow B) = \frac{\text{Confidence } A \rightarrow B}{\text{Support}(AB)} = \frac{0,25}{0,2} = 1,25$

h. Itemset AD

- Jumlah Transaksi $|D| = 5$
- Kemunculan item AB pada transaksi ($|Tab|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(AD) = \frac{|Tad|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } A \rightarrow D = \frac{|Tab|}{|Ta|} = \frac{1}{4} = 0,25$
- $\text{Lift } (A \rightarrow D) = \frac{\text{Confidence } A \rightarrow D}{\text{Support}(AD)} = \frac{0,25}{0,2} = 1,25$

i. Itemset AE

- Jumlah Transaksi $|D| = 5$
- Kemunculan item AE pada transaksi ($|Tae|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(AE) = \frac{|Tae|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } A \rightarrow E = \frac{|Tae|}{|Ta|} = \frac{1}{4} = 0,25$

- $\text{Lift}(A \rightarrow E) = \frac{\text{Confidence } A \rightarrow E}{\text{Support}(AE)} = \frac{0,25}{0,2} = 1,25$

j. Itemset BC

- Jumlah Transaksi $|D| = 5$
- Kemunculan item BC pada transaksi ($|T_{bc}|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(BC) = \frac{|T_{bc}|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } B \rightarrow C = \frac{|T_{bc}|}{|T_b|} = \frac{1}{1} = 1$
- $\text{Lift}(B \rightarrow C) = \frac{\text{Confidence } B \rightarrow C}{\text{Support}(BC)} = \frac{1}{0,2} = 5$

k. Itemset CD

- Jumlah Transaksi $|D| = 5$
- Kemunculan item CD pada transaksi ($|T_{cd}|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(CD) = \frac{|T_{cd}|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } C \rightarrow D = \frac{|T_{cd}|}{|T_c|} = \frac{1}{3} = 0,3$
- $\text{Lift}(C \rightarrow D) = \frac{\text{Confidence } C \rightarrow D}{\text{Support}(CD)} = \frac{0,3}{0,2} = 1,5$

l. Itemset DE

- Jumlah Transaksi $|D| = 5$
- Kemunculan item DE pada transaksi ($|T_{de}|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(DE) = \frac{|T_{de}|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } D \rightarrow E = \frac{|T_{de}|}{|T_d|} = \frac{1}{2} = 0,5$
- $\text{Lift}(D \rightarrow E) = \frac{\text{Confidence } D \rightarrow E}{\text{Support}(DE)} = \frac{0,5}{0,2} = 2,5$

m. Itemset ABC

- Jumlah Transaksi $|D| = 5$
- Kemunculan item ABC pada transaksi ($|T_{abc}|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(ABC) = \frac{|T_{abc}|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } A, B \rightarrow C = \frac{|T_{abc}|}{|T_{ab}|} = \frac{1}{4} = 0,25$
- $\text{Lift}(A, B \rightarrow C) = \frac{\text{Confidence } A, B \rightarrow C}{\text{Support}(ABC)} = \frac{0,25}{0,2} = 1,25$

n. Itemset ACD

- Jumlah Transaksi $|D| = 5$
- Kemunculan item ACD pada transaksi ($|T_{acd}|$) sebanyak 1 kali, yaitu pada T1
- $\text{Support}(ACD) = \frac{|T_{acd}|}{|D|} = \frac{1}{5} = 0,2$
- $\text{Confidence } A, C \rightarrow D = \frac{|T_{acd}|}{|T_{ac}|} = \frac{1}{4} = 0,25$
- $\text{Lift}(A, C \rightarrow D) = \frac{\text{Confidence } A, C \rightarrow D}{\text{Support}(ACD)} = \frac{0,25}{0,2} = 1,25$

2.2.Kaidah Asosiasi Minsup 0,2

a. 1 itemset

Itemset	Support
A	0,8
B	0,2
C	0,6
D	0,4
E	0,4

Dengan batasan minsup 0,2, maka itemset yang dapat diikutkan pada tahap selanjutnya adalah

Itemset	Support
A	0,8
B	0,2
C	0,6
D	0,4
E	0,4

b. 2 itemset

Itemset	Support
AB	0,2
AC	0,6
AD	0,2
AE	0,2
BC	0,2
CD	0,2
DE	0,2

Dengan batasan minsup 0,2, maka itemset yang dapat diikutkan pada tahap selanjutnya adalah

Itemset	Support
AB	0,2
AC	0,6
AD	0,2
AE	0,2
BC	0,2
CD	0,2
DE	0,2

c. 3 itemset

Itemset	Support
ABC	0,2
ACD	0,2

Dengan batasan minsup 0,2, maka itemset yang dapat diikutkan pada tahap selanjutnya adalah

Itemset	Support
ABC	0,2
ACD	0,2

2.3.Kaidah Asosiasi Confidence 0,8

Itemset	Confidence
AB	0,25
AC	0,25
AD	0,25
AE	0,25
BC	1
CD	0,25
DE	0,25
ABC	0,25
ACD	0,25

Dari hasil tabel asosiasi, maka dapat digenerasi suatu aturan asosiasi sebagai berikut:

$B \rightarrow C$

$C \rightarrow B$