

# Association Rule

Data Mining

# Association Rule

- Analisis asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item.
- Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut affinity analysis atau market basket analysis

# Association Rule

- Analisis asosiasi didefinisikan sebagai suatu proses untuk menemukan semua aturan asosiatif yang memenuhi syarat minimum untuk *support (minimum support)* dan syarat minimum untuk *confidence (minimum confidence)*.
- Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support (nilai penunjang)* yaitu persentase kombinasi item tersebut dalam database dan *confidence (nilai kepastian)* yaitu kuatnya hubungan antar item dalam aturan asosiatif

# Contoh Association Rule

{roti, mentega} → {susu} (support = 40%, confidence = 50%)

- Yang artinya : "50% dari transaksi di database yang memuat item roti dan mentega juga memuat item susu. Sedangkan 40% dari seluruh transaksi yang ada di database memuat ketiga item itu."
- Dapat juga diartikan : "Seorang konsumen yang membeli roti dan mentega punya kemungkinan 50% untuk juga membeli susu. Aturan ini cukup signifikan karena mewakili 40% dari catatan transaksi selama ini."

# Algoritma A Priori

- Algoritma A Priori termasuk jenis aturan asosiasi pada data mining. Selain a priori, yang termasuk pada golongan ini adalah metode *generalized rule induction* dan *algoritma hash based*.

# Metodologi dasar analisis asosiasi :

- 1. Analisa pola frekuensi tinggi**
- 2. Pembentukan aturan assosiatif**

# 1. Analisa pola frekuensi tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database. Nilai support sebuah item diperoleh dengan rumus berikut:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi mengandung A}}{\text{Total Transaksi}}$$

sedangkan nilai support dari 2 item diperoleh dari rumus berikut:

$$\text{Support (A} \cap \text{B)} = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Total Transaksi}}$$

## 2. Pembentukan aturan assosiatif

- Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan assosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan assosiatif  $A \rightarrow B$
- Nilai *confidence* dari aturan  $A \rightarrow B$  diperoleh dari rumus berikut:

$$\text{Confidence} = P(B | A) = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Jumlah Transaksi mengandung A}}$$



# algoritma pada ***Association Rule***

1. Tentukan  $\Phi$
2. Tentukan semua Frekuen Item set
3. Untuk setiap Frekuen Item set lakukan hal sbb:
  - i. Ambil sebuah unsur, namakanlah  $s$
  - ii. Untuk sisanya namakanlah  $ss-s$
  - iii. Masukkan unsur-unsur yang telah diumpamakan ke dalam rule ***If (ss-s) then s***

Untuk langkah ke 3 lakukan untuk semua unsur.

# Contoh Kasus

NO	Transaksi Item yang di beli
1	C, E, D
2	A, F, D
3	D, G, B, F
4	E, D, G, B
5	B, A, C
6	F, A, B, G
7	G, D
8	C, G, E
9	F, A, B
10	B, D

(analisa pola frekwensi tinggi)

Pisahkan masing-masing item yang  
dibeli

NO	Transaksi Item yang di beli
1	C, E, D
2	A, F, D
3	D, G, B, F
4	E, D, G, B
5	B, A, C
6	F, A, B, G
7	G, D
8	C, G, E
9	F, A, B
10	B, D

Item yang di beli
A
B
C
D
E
F
G

# Jumlah pembelian untuk setiap item.

NO	A	B	C	D	E	F	G
1	0	0	1	1	1	0	0
2	1	0	0	1	0	1	0
3	0	1	0	1	0	1	1
4	0	1	0	1	1	0	1
5	1	1	1	0	0	0	0
6	1	1	0	0	0	1	1
7	0	0	0	1	0	0	1
8	0	0	1	0	1	0	1
9	1	1	0	0	0	1	0
10	0	1	0	1	0	0	0
<b>TOT</b>	<b>4</b>	<b>6</b>	<b>3</b>	<b>6</b>	<b>3</b>	<b>4</b>	<b>5</b>

# 1 item-set

$$F1 = \{\{A\}, \{B\}, \{C\}, \{D\}, \{E\}, \{F\}, \{G\}\}$$

Tentukan  $\Phi$ .

Misl  $\Phi=3$

NO	A	B	C	D	E	F	G
1	0	0	1	1	1	0	0
2	1	0	0	1	0	1	0
3	0	1	0	1	0	1	1
4	0	1	0	1	1	0	1
5	1	1	1	0	0	0	0
6	1	1	0	0	0	1	1
7	0	0	0	1	0	0	1
8	0	0	1	0	1	0	1
9	1	1	0	0	0	1	0
10	0	1	0	1	0	0	0
<b>TOT</b>	<b>4</b>	<b>6</b>	<b>3</b>	<b>6</b>	<b>3</b>	<b>4</b>	<b>5</b>

## 2 item-set

Untuk  $k=2$  diperlukan tabel untuk tiap-tiap pasang item.

$\{A,B\}$ ,  $\{A,C\}$ ,  $\{A,D\}$ ,  $\{A,E\}$ ,  $\{A,F\}$ ,  $\{A,G\}$ ,

$\{B,C\}$ ,  $\{B,D\}$ ,  $\{B,E\}$ ,  $\{B,F\}$ ,  $\{B,G\}$ ,

$\{C,D\}$ ,  $\{C,E\}$ ,  $\{C,F\}$ ,  $\{C,G\}$ ,

$\{D,E\}$ ,  $\{D,F\}$ ,  $\{D,G\}$ ,

$\{E,F\}$ ,  $\{E,G\}$ ,  $\{F,G\}$

# k = 2 item set

T	A	B	Hasil
1	0	0	F
2	1	0	F
3	0	1	F
4	0	1	F
5	1	1	T
6	1	1	T
7	0	0	F
8	0	0	F
9	1	1	T
10	0	1	F
TOT T			3

T	A	C	Hasil
1	0	1	F
2	1	0	F
3	0	0	F
4	0	0	F
5	1	1	T
6	1	0	F
7	0	0	F
8	0	1	F
9	1	0	F
10	0	0	F
TOT T			1

T	A	D	Hasil
1	0	1	F
2	1	1	T
3	0	1	F
4	0	1	F
5	1	0	F
6	1	0	F
7	0	1	F
8	0	0	F
9	1	0	F
10	0	1	F
TOT T			1

T	A	E	Hasil
1	0	1	F
2	1	0	F
3	0	0	F
4	0	1	F
5	1	0	F
6	1	0	F
7	0	0	F
8	0	1	F
9	1	0	F
10	0	0	F
TOT T			0

T	A	F	Hasil
1	0	0	F
2	1	1	T
3	0	1	F
4	0	0	F
5	1	0	F
6	1	1	T
7	0	0	F
8	0	0	F
9	1	1	T
10	0	0	F
TOT T			3

T	A	G	Hasil
1	0	0	F
2	1	0	F
3	0	1	F
4	0	1	F
5	1	0	F
6	1	1	T
7	0	1	F
8	0	1	F
9	1	0	F
10	0	0	F
TOT T			1

T	B	C	Hasil
1	0	1	F
2	0	0	F
3	1	0	F
4	1	0	F
5	1	1	T
6	1	0	F
7	0	0	F
8	0	1	F
9	1	0	F
10	1	0	F
TOT T			1

T	B	D	Hasil
1	0	1	F
2	0	1	F
3	1	1	T
4	1	1	T
5	1	0	F
6	1	0	F
7	0	1	F
8	0	0	F
9	1	0	F
10	1	1	T
TOT T			3

T	B	E	Hasil
1	0	1	F
2	0	0	F
3	1	0	F
4	1	1	T
5	1	0	F
6	1	0	F
7	0	0	F
8	0	1	F
9	1	0	F
10	1	0	F
TOT T			1

T	B	F	Hasil
1	0	0	F
2	0	1	F
3	1	1	T
4	1	0	F
5	1	0	F
6	1	1	T
7	0	0	F
8	0	0	F
9	1	1	T
10	1	0	F
TOT T			3

T	B	G	Hasil
1	0	0	F
2	0	0	F
3	1	1	T
4	1	1	T
5	1	0	F
6	1	1	T
7	0	1	F
8	0	1	F
9	1	0	F
10	1	0	F
TOT T			3

# k = 2

T	C	D	Hasil
1	1	1	T
2	0	1	F
3	0	1	F
4	0	1	F
5	1	0	F
6	0	0	F
7	0	1	F
8	1	0	F
9	0	0	F
10	0	1	F
TOT T			1

T	C	E	Hasil
1	1	1	T
2	0	0	F
3	0	0	F
4	0	1	F
5	1	0	F
6	0	0	F
7	0	0	F
8	1	1	T
9	0	0	F
10	0	0	F
TOT T			2

T	C	F	Hasil
1	1	0	F
2	0	1	F
3	0	1	F
4	0	0	F
5	1	0	F
6	0	1	F
7	0	0	F
8	1	0	F
9	0	1	F
10	0	0	F
TOT T			0

T	C	G	Hasil
1	1	0	F
2	0	0	F
3	0	1	F
4	0	1	F
5	1	0	F
6	0	1	F
7	0	1	F
8	1	1	T
9	0	0	F
10	0	0	F
TOT T			1

T	D	E	Hasil
1	1	1	T
2	1	0	F
3	1	0	F
4	1	1	T
5	0	0	F
6	0	0	F
7	1	0	F
8	0	1	F
9	0	0	F
10	1	0	F
TOT T			2

T	D	F	Hasil
1	1	0	F
2	1	1	T
3	1	1	T
4	1	0	F
5	0	0	F
6	0	1	F
7	1	0	F
8	0	0	F
9	0	1	F
10	1	0	F
TOT T			2

T	D	G	Hasil
1	1	0	F
2	1	0	F
3	1	1	T
4	1	1	T
5	0	0	F
6	0	1	F
7	1	1	T
8	0	1	F
9	0	0	F
10	1	0	F
TOT T			3

T	E	F	Hasil
1	1	0	F
2	0	1	F
3	0	1	F
4	1	0	F
5	0	0	F
6	0	1	F
7	0	0	F
8	1	0	F
9	0	1	F
10	0	0	F
TOT T			0

T	E	G	Hasil
1	1	0	F
2	0	0	F
3	0	1	F
4	1	1	T
5	0	0	F
6	0	1	F
7	0	1	F
8	1	1	T
9	0	0	F
10	0	0	F
TOT T			2

T	F	G	Hasil
1	0	0	F
2	1	0	F
3	1	1	T
4	0	1	F
5	0	0	F
6	1	1	T
7	0	1	F
8	0	1	F
9	1	0	F
10	0	0	F
TOT T			2



# Frekuensi Item set F2

*Dari tabel diatas diatas frequent k2-item set*

$$F2 = \{\{A,B\}, \{A,F\}, \{B,D\}, \{B,F\}, \{B,G\}, \{D,G\}\}$$

# K=3 item set

$$F2 = \{\{A,B\}, \{A,F\}, \{B,D\}, \{B,F\}, \{B,G\}, \{D,G\}\}$$

Itemset-itemset yang dapat digabungkan adalah itemset-itemset yang memiliki kesamaan dalam k-1 item pertama. Misalnya {A,B} dan {A,F} memiliki itemset k-1 pertama yg sama, yaitu A, maka dapat digabungkan menjadi 3-itemset baru yaitu {A, B, F}.

$$\{A, B, F\}, \{B, D, F\}, \{B, D, G\}, \{B, F, G\}$$

# F3

T	A	B	F	Hasil
1	0	0	0	F
2	1	0	1	F
3	0	1	1	F
4	0	1	0	F
5	1	1	0	F
6	1	1	1	T
7	0	0	0	F
8	0	0	0	F
9	1	1	1	T
10	0	1	0	F
TOT T				2

T	B	D	F	Hasil
1	0	1	0	F
2	0	1	1	F
3	1	1	1	T
4	1	1	0	F
5	1	0	0	F
6	1	0	1	F
7	0	1	0	F
8	0	0	0	F
9	1	0	1	F
10	1	1	0	F
TOT T			1	

T	B	D	G	Hasil
1	0	1	0	F
2	0	1	0	F
3	1	1	1	T
4	1	1	1	T
5	1	0	0	F
6	1	0	1	F
7	0	1	1	F
8	0	0	1	F
9	1	0	0	F
10	1	1	0	F
TOT T			2	

T	B	F	G	Hasil
1	0	0	0	F
2	0	1	0	F
3	1	1	1	T
4	1	0	1	F
5	1	0	0	F
6	1	1	1	T
7	0	0	1	F
8	0	0	1	F
9	1	1	0	F
10	1	0	0	F
TOT T			2	

Dari tabel-tabel di atas, didapat  $F3 = \{ \}$ , karena tidak ada  $\Sigma \geq \Phi$  sehingga  $F4, F5, F6$  dan  $F7$  juga merupakan himpunan kosong.

# Pembentukan aturan assosiatif

- Rule yang dipakai adalah *if x then y, dimana x adalah antecedent dan y adalah consequent.* Berdasarkan rule tersebut, maka dibutuhkan 2 buah item yang mana salah satunya sebagai antecedent dan sisanya sebagai consequent.
- Dari langkah sblmnya didapat 1 buah Fk yaitu F2. F1 tidak disertakan karena hanya terdiri dari 1 item saja. Untuk antecedent boleh lebih dari 1 unsur, sedangkan untuk consequent terdiri dari 1 unsur.

# Pembentukan aturan assosiatif

- Tentukan  $(ss-s)$  sebagai antecedent dan  $s$  sebagai consequent dari  $F_k$  yang telah di dapat berdasarkan rule pada langkah sebelumnya.
- Pada  $F_2$  didapat himpunan  $F_2 = \{\{A,B\}, \{A,F\}, \{B,D\}, \{B,F\}, \{B,G\}, \{D,G\}\}$

# Pembentukan aturan assosiatif

- Untuk  $\{A,B\}$ :
  - Jika  $(ss-s) = A$ , Jika  $s = B$ , Maka If buy A then buy B
  - Jika  $(ss-s) = B$ , Jika  $s = A$ , Maka If buy B then buy A
- Untuk  $\{A,F\}$ :
  - Jika  $(ss-s) = A$ , Jika  $s = F$ , Maka If buy A then buy F
  - Jika  $(ss-s) = F$ , Jika  $s = A$ , Maka If buy F then buy A
- Untuk  $\{B,D\}$ :
  - Jika  $(ss-s) = B$ , Jika  $s = D$ , Maka If buy B then buy D
  - Jika  $(ss-s) = D$ , Jika  $s = B$ , Maka If buy D then buy B
- Untuk  $\{B,F\}$ :
  - Jika  $(ss-s) = B$ , Jika  $s = F$ , Maka If buy B then buy F
  - Jika  $(ss-s) = F$ , Jika  $s = B$ , Maka If buy F then buy B
- Untuk  $\{B,G\}$ :
  - Jika  $(ss-s) = B$ , Jika  $s = G$ , Maka If buy B then buy G
  - Jika  $(ss-s) = G$ , Jika  $s = B$ , Maka If buy G then buy B
- Untuk  $\{D,G\}$ :
  - Jika  $(ss-s) = D$ , Jika  $s = G$ , Maka If buy D then buy G
  - Jika  $(ss-s) = G$ , Jika  $s = D$ , Maka If buy G then buy D

# Pembentukan aturan assosiatif

- Dari langkah sblmnya, kita mendapatkan 12 rule yang dapat digunakan yaitu

If antecedent then consequent	Support	Confidence
If buy A then buy B		
If buy B then buy A		
If buy A then buy F		
If buy F then buy A		
If buy B then buy D		
If buy D then buy B		
If buy B then buy F		
If buy F then buy B		
If buy B then buy G		
If buy G then buy B		
If buy D then buy G		
If buy G then buy D		

# Hitung support dan confidence.

$$SUPPORT = \frac{\Sigma \text{item yang dibeli sekaligus}}{\Sigma \text{jumlah seluruh transaksi}} \times 100\%$$

$$CONFIDENCE = \frac{\Sigma \text{item yang dibeli sekaligus}}{\Sigma \text{jumlah transaksi pada bagian antecedent}} \times 100\%$$



# Pembentukan aturan assosiatif

Untuk  $\Sigma$  item yang dibeli sekaligus pada **If buy A then buy B**, ada 3 transaksi. Jumlah seluruh transaksi adalah 10 transaksi, sehingga support-nya adalah:

$$\text{Support} = (3/10) * 100\% = 30\%$$

sedangkan jumlah transaksi yang membeli A adalah 4 transaksi, sehingga confidence-nya adalah:

$$\text{Confidence} = (3/4) * 100\% = 75\%$$

If antecedent then consequent	Support		Confidence	
<b>If buy A then buy B</b>	<b><math>(3/10) \times 100\% =</math></b>	<b>30.00%</b>	<b><math>(3/4) \times 100\% =</math></b>	<b>75.00%</b>
If buy B then buy A	$(3/10) \times 100\% =$	30.00%	$(3/6) \times 100\% =$	50.00%
If buy A then buy F	$(3/10) \times 100\% =$	30.00%	$(3/4) \times 100\% =$	75.00%
If buy F then buy A	$(3/10) \times 100\% =$	30.00%	$(3/4) \times 100\% =$	75.00%
If buy B then buy D	$(3/10) \times 100\% =$	30.00%	$(3/6) \times 100\% =$	50.00%
If buy D then buy B	$(3/10) \times 100\% =$	30.00%	$(3/6) \times 100\% =$	50.00%
If buy B then buy F	$(3/10) \times 100\% =$	30.00%	$(3/6) \times 100\% =$	50.00%
If buy F then buy B	$(3/10) \times 100\% =$	30.00%	$(3/4) \times 100\% =$	75.00%
If buy B then buy G	$(3/10) \times 100\% =$	30.00%	$(3/6) \times 100\% =$	50.00%
If buy G then buy B	$(3/10) \times 100\% =$	30.00%	$(3/5) \times 100\% =$	60.00%
If buy D then buy G	$(3/10) \times 100\% =$	30.00%	$(3/6) \times 100\% =$	50.00%
If buy G then buy D	$(3/10) \times 100\% =$	30.00%	$(3/5) \times 100\% =$	60.00%

# Pembentukan aturan assosiatif

- Setelah di dapat support dan confidence untuk masing-masing kandidat, lakukan perkalian antara support dan confidence, dimana confidence-nya diambil 70% ke atas, sehingga di dapat tabel sbb:

If antecedent then consequent	Support	Confidence	Support X Confidence
If buy A then buy B	30.00%	75.00%	0.225
If buy A then buy F	30.00%	75.00%	0.225
If buy F then buy A	30.00%	75.00%	0.225
If buy F then buy B	30.00%	75.00%	0.225

# Pembentukan aturan assosiatif

- Hasil paling besar dari perkalian perkalian tersebut merupakan rule yang dipakai pada saat menjual
- Maka rule yg dihasilkan:
  - *Jika membeli A maka akan membeli B dengan support 33,33% dan confidence 75%*
  - *Jika membeli A maka akan membeli F dengan support 33,33% dan confidence 75%*
  - *Jika membeli F maka akan membeli A dengan support 33,33% dan confidence 75%*
  - *Jika membeli F maka akan membeli B dengan support 33,33% dan confidence 75%*