

BAB II

KONSEP DAN DEFINISI

Pada bab ini akan dijelaskan konsep dan definisi-definisi yang digunakan dalam metode pada penelitian ini.

2.1 DATA TRANSAKSI

Misalkan $X = \{X_1, X_2, X_3, \dots, X_M\}$ adalah himpunan semua produk yang dijual oleh suatu swalayan. Di mana X_j adalah produk (*item*) ke- j . Untuk suatu transaksi ke- i nilai dari X_j adalah :

$x_{ij} = 1$, jika *item* ke- j dibeli pada transaksi ke- i

$x_{ij} = 0$, jika *item* ke- j tidak dibeli pada transaksi ke- i

dengan $i = 1, 2, \dots, N$ dan $j = 1, 2, \dots, M$

Sehingga untuk suatu database transaksi dapat ditabulasikan dalam bentuk sebagai berikut :

Tabel 2.1. Format Data Transaksi

<i>Item</i> \ <i>Transaksi</i>	X_1	X_2	...	X_M
1	x_{11}	x_{12}	K	x_{1M}
2	x_{21}	x_{22}	K	x_{2M}
N	N	N		N
N	x_{N1}	x_{N2}	K	x_{NM}

Contoh 1 : $X = \{X_1, X_2, X_3\}$

$X_1 = \text{Roti}$, $X_2 = \text{Keju}$ $X_3 = \text{Susu}$

Jika transaksi 1 memuat *item* Roti dan keju tetapi tidak memuat susu, maka nilai $x_{11} = 1$, $x_{12} = 1$, $x_{13} = 0$.

Misal $X = \{X_1, X_2, X_3, \dots, X_M\}$ adalah himpunan semua *item-item* yang diamati dan $A \subseteq X$, $B \subseteq X$ serta $A \cap B = \emptyset$, A dan B saling lepas (*mutually exclusive*), $A = \{X_1, X_2, \dots, X_p\}$, $B = \{X_1, X_2, \dots, X_q\}$, $p \neq q$, $p, q \in \mathbb{Z}$.

Aturan asosiasi berbentuk “jika A maka B”. Intepretasinya adalah “jika seorang membeli *item* yang ada di A maka ia akan membeli *item* yang ada di B”. Dari contoh 1, misalnya $A = \{X_1, X_2\}$, $B = \{X_3\}$, maka aturannya “jika seorang pelanggan membeli roti dan keju, maka ia akan membeli susu”.

2.2 DEFINISI

Berikut ini adalah definisi-definisi yang perlu diketahui dalam metode kaidah asosiasi

Definisi 1

Itemset adalah himpunan dari *item-item* yang terbeli pada suatu transaksi.

Contoh : {susu, roti}, {roti}, {susu roti keju}

Definisi 2

k-itemset adalah *itemset* yang memiliki kardinalitas *k*.

contoh : {susu, roti, keju} = 3-itemset

{roti, keju} = 2-itemset

Misalkan $X = \{X_1, X_2, X_3, \dots, X_M\}$ himpunan dari *item* dan

$A = \{X_1, X_2, X_3, \dots, X_k\}$ adalah sebuah *k-itemset*, $A \subseteq X$, $k \leq M$.

Nilai dari suatu *k-itemset* $A = \{X_1, X_2, X_3, \dots, X_k\}$ untuk suatu transaksi adalah

$\{X_1, X_2, \dots, X_k\} = 1$, jika semua *item* yang ada di himpunan tersebut terjual pada transaksi ke-*i*

= 0, jika paling tidak satu *item* dari himpunan tersebut tidak terjual pada transaksi ke-*i*

atau secara matematis nilai dari k *itemset* $A = \{X_1, X_2, X_3, \dots, X_k\}$ untuk suatu

transaksi ke-i nilai dari suatu k-*itemset* adalah :

$$\{X_1, X_2, X_3, \dots, X_k\} = \prod_{a \in \lambda} X_{ia}; \quad \lambda = \{\text{indeks pada A}\}$$

Definisi 3

Sebuah *itemset* A dikatakan *superset* dari *itemset* B bila setiap anggota dari B adalah anggota A juga. Sebaliknya B adalah *subset* dari A

Contoh : $A = \{X_1, X_2, X_3\}$, maka A *superset* dari *itemset-itemset* berikut :

$$\{X_1, X_2\}, \{X_1, X_3\}, \{X_2, X_3\}, \{X_1\}, \{X_2\}, \{X_3\}$$

Definisi 4

Itemset frequency adalah jumlah transaksi yang berisi *itemset* tertentu

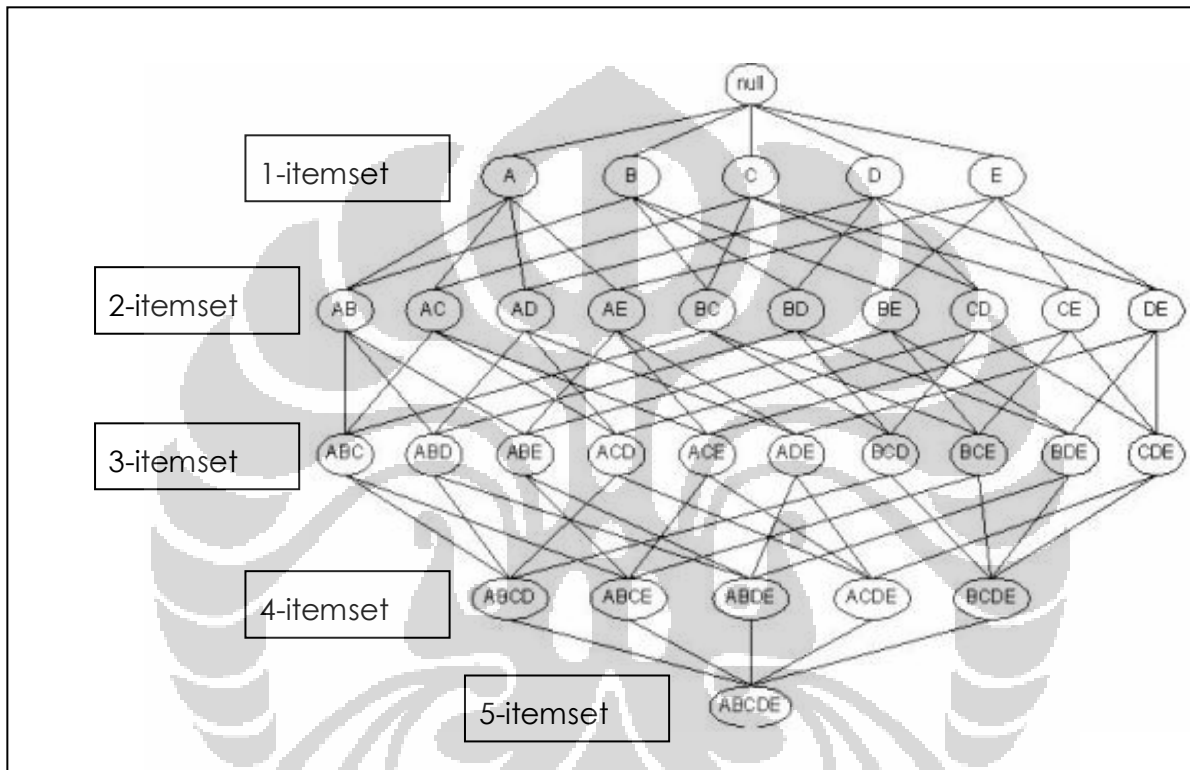
Misal A adalah *itemset* dan $\lambda = \{\text{integer pada A}\}$, maka:

$$\text{Itemset frequency (A)} = \sum_{i=1}^N \prod_{a \in \lambda} X_{ia} \quad (1)$$

Itemset frequency dapat juga diartikan berapa kali *item-item* yang ada pada *itemset* tersebut terbeli secara bersamaan dalam suatu observasi.

Itemset dibentuk dari *join* antar *item* yang diamati. Untuk lebih jelasnya dapat dijelaskan pada Gambar 2.1 :

Misalkan ada lima buah *item* A, B, C, D, E, maka *itemset* yang terbentuk adalah



Gambar 2.1. Kemungkinan *Itemset* yang Terbentuk dari Lima *Item*

Jadi, untuk *item* sebanyak M, *itemset* yang terbentuk adalah sebanyak

$$C_1^M + C_2^M + C_3^M + \dots + C_M^M = \sum_{i=1}^M C_i^M .$$

Definifisi 5

Support dari suatu *itemset* A adalah proporsi dari kejadian semua *item* di himpunan A terbeli secara bersamaan. Dirumuskan dengan

$$S(A) = \frac{\text{itemset frequency}(A)}{\text{total transaksi}} \quad (2)$$

$$= \frac{1}{N} \sum_{i=1}^N \prod_{a \in \lambda} x_{ia}$$

Dalam pemrosesannya, penelitian ini bertujuan untuk memilih *itemset* yang sering dibeli oleh pelanggan (*itemset* yang *frequent*) dari sebanyak $\sum_{i=1}^M C_i^M$ *itemset* yang terbentuk. Oleh karena hal tersebut, diperlukan suatu indikator untuk membedakan antara *itemset* yang *frequent* dan tidak. Indikator yang dimaksud adalah *minimum support*.

Definifisi 6

Minimum support adalah *support* minimum yang dicapai suatu *itemset* yang *frequent*. Nilai dari *minimum support* diberikan oleh pihak manajemen minimarket. Nilai dari *minimum support* berbeda untuk tiap minimarket atau supermarket

Definifisi 7

Suatu *itemset* A dikatakan *frequent* jika dan hanya jika memiliki *support* lebih besar dari *minimum support*.

$$S(A) \geq \text{minimum support}$$
$$\frac{1}{N} \sum_{i=1}^N \prod_{a \in A} x_{ia} \geq \text{minimum support}$$

