

## **PENERAPAN DATA MINING UNTUK ANALISIS POLA BELANJA KONSUMEN MENGGUNAKAN ALGORITMA APRIORI PADA MALL CPM JAKARTA**

**Persis Haryo Winasis**

Mahasiswa Program Pascasarjana Magister Ilmu Komputer

Universitas Budi Luhur

Email: [persis.winasis@gmail.com](mailto:persis.winasis@gmail.com)

### **ABSTRACT**

*Companies engaged in retail such as malls that have a lot of transaction data and sales transactions that are very much. Every purchase transaction made by consumers will be recorded and purchased in one database. Processing data in this study was carried out using a priori algorithm and using the help of the Weka application. The results of data mining in this study are expected to be able to produce new information about spending patterns in a certain period that can be used by the mall manager and store manager to support each related product promotion or organizing an event to increase the number of consumers in a certain period.*

*Keywords: apriori algorithm, data mining, shopping pattern analysis.*

### **PENDAHULUAN**

Persaingan perusahaan yang bergerak dibidang retail seperti *mall* di Indonesia khususnya di kota besar seperti di Jakarta menuntut pihak pengelola *mall* untuk bisa menemukan strategi agar mampu terus menerus menarik minat pengunjung untuk datang dan berbelanja di toko yang ada. Keberlangsungan banyaknya transaksi penjualan dari masing masing toko juga tidak terlepas dari peran aktif pihak pengelola *mall* untuk bisa memberikan tawaran promosi maupun acara acara yang diadakan di dalam *mall*. Konsumen atau pengunjung cenderung akan lebih memilih berbelanja di suatu *mall* yang bisa memberikan penawaran promosi yang sesuai dengan kebutuhan konsumen saat ini. Analisis terhadap data transaksi mampu memberikan informasi mengenai pola belanja konsumen di dalam suatu *mall* dalam periode tertentu. Analisis mengenai toko apa saja yang sering dikunjungi dalam satu waktu yang sama oleh seorang konsumen dengan menggunakan aturan asosiasi dan algoritma *apriori*

berfungsi untuk menemukan kombinasi yang mungkin terjadi. Nantinya kombinasi tersebut akan diuji apakah memenuhi nilai ambang yang sudah ditentukan melalui parameter *minimum support* dan *minimum confidence*. Tujuan dari analisis tersebut adalah mendapatkan informasi mengenai pola belanja konsumen dalam suatu *mall*, sehingga informasi tersebut bisa digunakan pihak pengelola *mall* maupun penjual sebagai pertimbangan dalam menentukan promosi barang maupun acara yang perlu diselenggarakan untuk meningkatkan penjualan.

## **TINJAUAN PUSTAKA**

### **1. *Data Mining***

*Data Mining* merupakan proses untuk menggali nilai tambah berupa informasi yang selama ini belum didapatkan secara manual dari suatu *database*. *Data mining* menggali pola-pola dari data untuk dimanipulasi menjadi sebuah informasi baru yang diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik (Apostolakis, 2010) (gsbipb, 2013).

Data mining memiliki beberapa fungsi yang paling umum, fungsi tersebut adalah (Larose, 2005) (Muchlisin Riadi, n.d.):

#### *a. Description*

Tujuan deskripsi adalah mengidentifikasi pola yang muncul secara berulang dari suatu data serta kemudian mengubah pola menjadi aturan dan kriteria yang mudah dimengerti.

#### *b. Clasification*

Merupakan proses menemukan model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

c. *Prediction*

Prediksi hamper sama dengan klasifikasi, namun data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang.

d. *Estimation*

Pada fungsi estimasi, model dibuat menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya pada pemeriksaan berikutnya, estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

e. *Clustering*

Tujuan *clustering* adalah untuk menghasilkan kelompok objek yang mirip satu sama lain dalam beberapa kelompok yang ada. Semakin besar kemiripan objek dalam suatu *cluster* dan semakin besar perbedaan tiap *cluster* maka kualitas analisis *cluster* akan semakin baik.

f. *Association*

Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam suatu waktu. Selain itu tugas dari asosiasi adalah berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

## 2. Algoritma *Apriori*

Algoritma *apriori* digunakan untuk mendapatkan aturan asosiasi dan mencari pola hubungan antar satu atau lebih item dalam suatu data (INFORMATIKALOGI, 2017). Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi (*frequent pattern mining*). Suatu asosiasi dapat diketahui penting atau tidak dapat diukur menggunakan parameter *support* dan *confidence*. *Support* atau nilai penunjang adalah ukuran yang menunjukkan tingkat dominasi suatu item dari keseluruhan transaksi,

sedangkan *confidence* atau nilai kepastian adalah kuatnya hubungan antar-item dalam aturan asosiasi (Kusrini, 2009).

### 1. Analisis Pola Frekuensi

Analisis ini dilakukan untuk mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam *database*. Nilai *support* dengan menggunakan satu buah item diperoleh dengan menggunakan rumus berikut:

$$\text{support}(A) = \frac{\text{jumlahtransaksimengandungA}}{\text{totaltransaksi}}$$

Sedangkan untuk nilai support dengan dua buah *item* dapat diperoleh menggunakan rumus berikut:

$$\begin{aligned} \text{suppotr}(A, B) &= P(A \cap B) \\ \text{support}(A, B) &= \frac{\sum \text{transaksimengandungAdanB}}{\sum \text{transaksi}} \end{aligned}$$

### 2. Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, langkah berikutnya adalah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A – B menggunakan rumus berikut:

$$\begin{aligned} \text{confidence} &= P(B|A) \\ &= \frac{\sum \text{transaksi mengandung A dan B}}{\sum \text{transaksi mengandung A}} \end{aligned}$$

Nilai yang dihasilkan kemudian diurutkan berdasarkan *support* dan *confidence* guna menentukan aturan asosiasi yang akan dipilih. Aturan diambil sebanyak “n” aturan yang memiliki hasil atau nilai terbesar.

## **METODE PENELITIAN**

### **1. Pengumpulan Data**

*Mall CPM* memiliki sekitar 500 toko didalamnya. Jumlah transaksi setiap hari mencapai ratusan bahkan ribuan transaksi penjualan. Data yang digunakan dalam penelitian ini adalah data transaksi penjualan periode 1 Juli 2019 sampai dengan 7 Juli 2019. Data tersebut diperoleh dari database aplikasi *membership* yang digunakan untuk melakukan pendataan transaksi penjualan pada *mall CPM*. Data tersebut kemudian diolah untuk mendapatkan jumlah kemunculan toko dalam periode tertentu.

### **2. Pengolahan Data Menggunakan Algoritma *Apriori***

Pengolahan data dilakukan dengan akumulasi transaksi penjualan pada *mall CPM* dalam periode yang sudah ditentukan dan kemudian dilakukan pencarian nilai menggunakan algoritma *apriori*. Tahapan yang dilakukan dalam algoritma *apriori* terdiri dari beberapa tahap, yaitu:

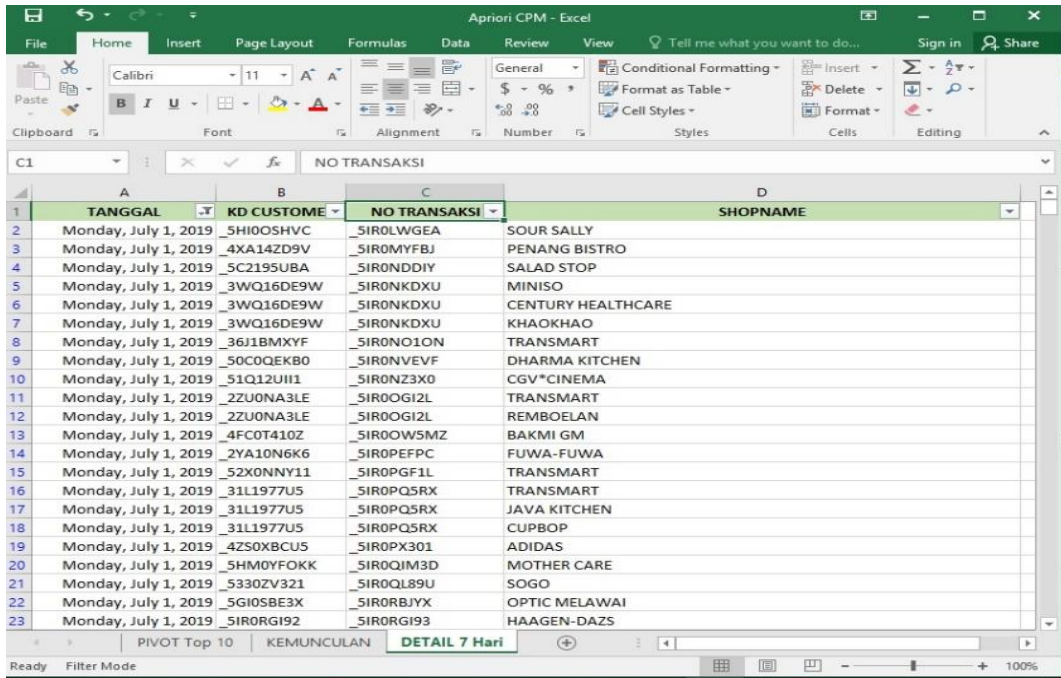
- a. Menentukan nilai dari *minimum support* dari setiap item yang ada.
- b. Menentukan nilai dari *minimum support* untuk 2 item set.
- c. Menentukan nilai dari *minimum confidence*.
- d. Proses pembentukan aturan asosiasi.

Dalam proses pembentukan aturan asosiasi, peneliti menggunakan bantuan aplikasi *Weka* versi 38.

## **HASIL DAN PEMBAHASAN**

### **1. Pengumpulan Data**

Pada penelitian ini proses pengolahan data dilakukan dengan data yang sudah diekspor kedalam bentuk tabel *Microsoft Excel*. Pengolahan data menggunakan data transaksi selama 1 minggu. Dari data tersebut didapatkan total sebanyak 3.304 data transaksi penjualan yang kemudian diolah kembali dengan cara mengambil 10 toko dengan penjualan teratas selama 1 minggu. Dari total 10 toko dengan peringkat penjualan teratas tersebut didapatkan jumlah data transaksi sebanyak 407 data.



Gambar 1. Data transaksi penjualan dalam kurun waktu 1 minggu.

## 2. Analisis Pencarian Pola Frekuensi Tertinggi

Pada tahapan ini akan dilakukan proses pembentukan 1 itemset atau disebut C1. Pada proses ini data akan diolah menggunakan rumus sebagai berikut:

$$support(A) = \frac{jumlahtransaksimengandungA}{totaltransaksi}$$

Dari rumus diatas, maka akan didapat *data support* dari masing masing item sebagai berikut:

Tabel 1. Data support itemset

No	Nama Toko	Count	Support (%)
1	TRANSMART	319	78.38%
2	ACEHOMECENTER	250	61.43%
3	GRAMEDIA	198	48.65%
4	URBANKITCHEN	174	42.75%
5	SOGO	140	34.40%
6	ZARA	124	27.27%
7	H&M	49	12.04%
8	CGV*CINEMA	48	11.79%
9	SUSHITEI	36	8.85%
10	STARBUCKS	32	7.86%

Langkah berikutnya, berdasarkan tabel 1 akan dicari nilai yang memenuhi nilai *minimum support* yang telah ditentukan yaitu  $\geq 30\%$ , maka akan didapat data yang tercantum didalam tabel berikut:

**Tabel 2. Data support itemset dengan nilai minimum support  $\geq 30\%$**

No	Nama Toko	Count	Support (%)
1	TRANSMART	319	78.38%
2	ACEHOMECENTER	250	61.43%
3	GRAMEDIA	198	48.65%
4	URBANKITCHEN	174	42.75%
5	SOGO	140	34.40%

Proses selanjutnya adalah melakukan pembentukan nilai 2 *itemset* atau C2 dengan cara melakukan kombinasi dari *itemset* yang sudah terbentuk sebelumnya. Pengolahan data dengan 2 *itemset* dilakukan dengan menggunakan rumus sebagai berikut:

$$support(A, B) = \frac{P(A \cap B)}{\sum \text{transaksimengandungAdanB}} = \frac{\sum \text{transaksi}}{\sum \text{transaksi}}$$

Dari pengolahan data tersebut diatas maka akan menghasilkan nilai 2 *itemset* seperti tertera didalam tabel berikut ini:

**Tabel 3. Data support 2 itemset**

Kode Join	Nama Toko	Count	Support (%)
A	TRANSMART, ACEHOMECENTRE	221	54.30%
B	TRANSMART, GRAMEDIA	183	44.96%
C	TRNASMART, URBANKITCHEN	148	36.36%
D	TRANSMART, SOGO	102	25.06%
E	ACEHOMECENTER, GRAMEDIA	170	41.77%
F	ACEHOMECENTER, URBANKITCHEN	114	28.01%
G	ACEHOMECENTER, SOGO	51	12.53%
H	GRAMEDIA, URBANKITCHEN	96	23.59%
I	GRAMEDIA, SOGO	46	11.30%
J	URBANKITCHEN, SOGO	49	12.04%

### 3. Pembentukan Aturan Asosiasi

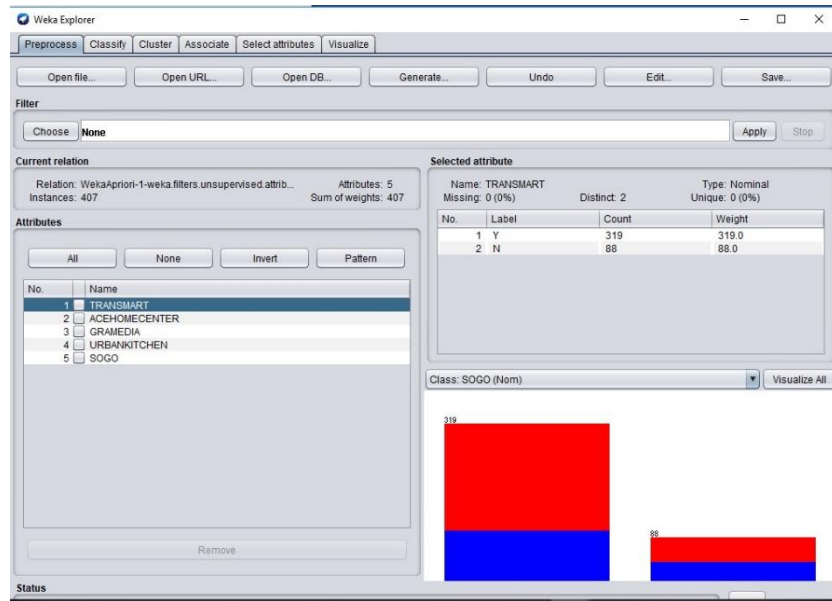
Kemudian untuk mencari nilai *confidence* dan aturan asosiasi yang dibutuhkan kita dapat menggunakan bantuan aplikasi *Weka* versi 3.8. Sebelumnya perlu disiapkan terlebih dahulu data tabular yang nantinya akan di unggah kedalam aplikasi *Weka*. Data tabular yang dimaksud dapat dilihat pada gambar 2 dibawah ini:

	A	B	C	D	E	F
1	NO TRANSAKSI	TRANSMART	ACEHOMECENTER	GRAMEDIA	URBANKITCHEN	SOGO
2	_5IR0RZ0DN	Y	N	N	N	Y
3	_5IROXARSY	Y	Y	N	N	Y
4	_5IS15URPF	N	N	N	Y	Y
5	_5IS19HB4B	Y	N	N	N	N
6	_5IS1A4UON	Y	N	N	N	Y
7	_5IT0XPNGW	Y	Y	N	N	Y
8	_5IU0W4UAZ	Y	N	Y	Y	N
9	_5IU0ZNE0Q	Y	N	N	N	Y
10	_5IV0ZSIQ0	N	Y	N	N	Y
11	_5IV12GUUN	Y	N	N	N	Y
12	_5IV17DXD5	Y	Y	N	N	N
13	_5IV18V74H	Y	N	N	N	Y
14	_5IW177F9L	N	N	Y	N	Y
15	_5IW1818QH	N	N	Y	N	Y
16	_5IW197OB0	Y	N	N	N	Y
17	_5IW19J28C	Y	Y	Y	Y	N
18	_5IX0UWKQK	N	N	N	N	Y
19	_5IX0ZVRO6	Y	N	Y	N	N
20	_5IX127PHO	N	N	N	N	Y
21	_5IR0TCET6	Y	N	Y	N	N
22	_5IROTE8K9	Y	N	N	N	Y

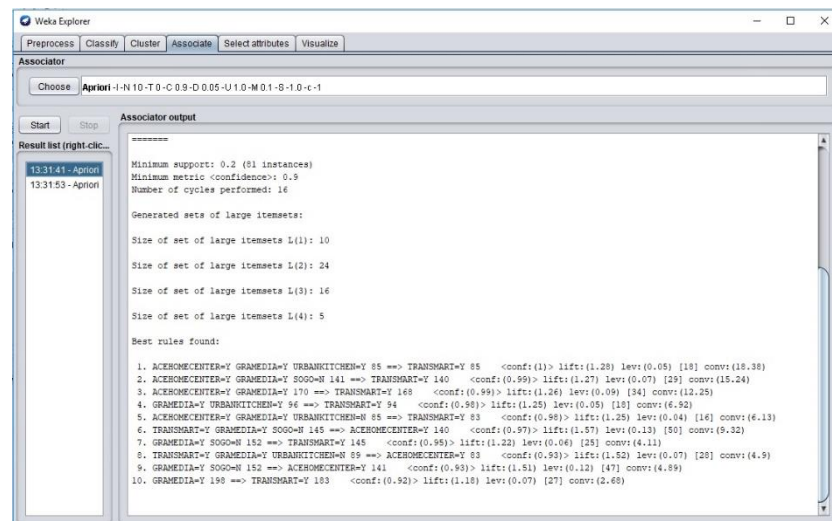
Gambar 2. Data tabular

Dalam proses pembentukan aturan asosiasi menggunakan aplikasi *Weka*, parameter nilai *minimum support* maupun nilai *confidence* dapat diatur sesuai dengan kebutuhan pengolahan data. Implementasi pengolahan data menggunakan aplikasi *Weka* dapat dilihat pada gambar 3 dan 4 dibawah ini:





Gambar 3. Implementasi menggunakan aplikasi Weka



Gambar 4. Hasil pengolahan data menggunakan aplikasi Weka

Dari hasil pengolahan data menggunakan aplikasi *Weka*, maka didapat hasil sebagai berikut:

```

=== Associator model (full training set) ===
Apriori
=====
Minimum support: 0.2 (81 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 16

Generated sets of large itemsets:

Size of set of large itemsets L(1): 10
Size of set of large itemsets L(2): 24
Size of set of large itemsets L(3): 16
Size of set of large itemsets L(4): 5

Best rules found:
1. ACEHOMECENTER=Y GRAMEDIA=Y URBANKITCHEN=Y 85 ==> TRANSMART=Y 85 <conf:(1)> lift:(1.28) lev:(0.05) [18] conv:(18.38)
2. ACEHOMECENTER=Y GRAMEDIA=Y SOGO=N 141 ==> TRANSMART=Y 140 <conf:(0.99)> lift:(1.27) lev:(0.07) [29] conv:(15.24)
3. ACEHOMECENTER=Y GRAMEDIA=Y 170 ==> TRANSMART=Y 168 <conf:(0.99)> lift:(1.26) lev:(0.09) [34] conv:(12.25)
4. GRAMEDIA=Y URBANKITCHEN=Y 96 ==> TRANSMART=Y 94 <conf:(0.98)> lift:(1.25) lev:(0.05) [18] conv:(6.92)
5. ACEHOMECENTER=Y GRAMEDIA=Y URBANKITCHEN=N 85 ==> TRANSMART=Y 83 <conf:(0.98)> lift:(1.25) lev:(0.04) [16] conv:(6.13)
6. GRAMEDIA=Y SOGO=N TRANSMART=Y 145 ==> ACEHOMECENTER=Y 140 <conf:(0.97)> lift:(1.57) lev:(0.13) [50] conv:(9.32)
7. GRAMEDIA=Y SOGO=N 152 ==> TRANSMART=Y 145 <conf:(0.95)> lift:(1.22) lev:(0.06) [25] conv:(4.11)
8. GRAMEDIA=Y TRANSMART=Y URBANKITCHEN=N 89 ==> ACEHOMECENTER=Y 83 <conf:(0.93)> lift:(1.52) lev:(0.07) [28] conv:(4.9)
9. GRAMEDIA=Y SOGO=N 152 ==> ACEHOMECENTER=Y 141 <conf:(0.93)> lift:(1.51) lev:(0.12) [47] conv:(4.89)
10. GRAMEDIA=Y 198 ==> TRANSMART=Y 183 <conf:(0.92)> lift:(1.18) lev:(0.07) [27] conv:(2.68)
    
```

**Gambar 5. Hasil pengolahan data menggunakan aplikasi Weka**

Dari data tersebut diatas kita mendapatkan 10 aturan asosiasi dengan nilai *confidence* yang berbeda. Nilai *confidence* terbesar didapat pada asosiasi 3 *itemset* yang menyebutkan bahwa jika terdapat pembelian pada toko Acehomecenter, Gramedia, Urbankitchen maka konsumen akan berbelanja juga di toko Transmart. Sedangkan nilai *confidence* terbesar untuk 2 *itemset* menyebutkan bahwa jika konsumen berbelanja di toko Gramedia maka konsumen tersebut juga berbelanja di toko Transmart.

## KESIMPULAN DAN SARAN

Dari penelitian diatas dapat disimpulkan, penggunaan algoritma *apriori* dapat memetakan kecenderungan pola belanja konsumen pada periode tertentu. Data yang dihasilkan dapat digunakan oleh pihak pengelola *mall* maupun pengelola serta pemilik toko untuk menganmbil keputusan seperti:

1. Menentukan lokasi atau area toko yang efektif pada masing masing lantai di dalam *mall* agar lalu lintas keramaian pengunjung *mall* dapat merata, sehingga bisa membantu peningkatan penjualan bagi toko yang penjualannya belum optimal.
2. Menentukan promo belanja yang sesuai dengan pola belanja konsumen pada periode yang diinginkan.

3. Menentukan *event* yang dapat diselenggarakan di area mall untuk menunjang peningkatan transaksi penjualan.

Saran untuk penelitian berikutnya terkait pola belanja konsumen *mall* adalah penelitian dapat menggunakan data minimal 1 tahun untuk dapat mendapatkan pola belanja konsumen yang lebih akurat dan dapat di korelasikan dengan pola belanja per bulan bahkan per tahun.

## **REFERENSI**

Apostolakis, J. (2010). An introduction to data mining. *Structure and Bonding*.  
[https://doi.org/10.1007/430\\_2009\\_1](https://doi.org/10.1007/430_2009_1)

gsbipb. (2013). Pengertian dan Konsep Data Mining. Retrieved from  
<https://gsbipb.com/?p=821>

INFORMATIKALOGI. (2017). Algoritma Apriori (Association Rule). Retrieved  
from <https://informatikalogi.com/algoritma-apriori-association-rule/>

Kusrini, E. T. L. (2009). Algoritma Data Mining. In *Data Mining*.

Larose, D. T. (2005). Discovering Knowledge in Data: An Introduction to Data Mining. In *Discovering Knowledge in Data: An Introduction to Data Mining*. <https://doi.org/10.1002/0471687545>

Muchlisin Riadi. (n.d.). *Pengertian, Fungsi, Proses dan Tahapan Data Mining*. Retrieved from <https://www.kajianpustaka.com/2017/09/data-mining.html>