

DATA MINING ANALISA POLA PEMBELIAN PRODUK DENGAN MENGGUNAKAN METODE ALGORITMA APRIORI

Heroe Santoso¹⁾, I Putu Hariyadi²⁾, Prayitno³⁾

^{1), 2) ,3)}Teknik Informatika STMIK Bumigora Mataram
Jl Ismail Marzuki Mataram

Email : heroe_santoso@yahoo.com¹⁾, putu.hariyadi@gmail.com²⁾, kangpray@gmail.com³⁾

Abstrak

Pengelola swalayan harus mencermati pola-pola pembelian yang dilakukan konsumen. Swalayan mempunyai kekurangan, diantaranya permasalahan peletakkan barang-barang yang tidak sesuai dengan perilaku konsumen dalam membeli barang secara bersamaan dalam satu waktu. Hal ini tentunya mempengaruhi tingkat penjualan. Adanya kegiatan penjualan setiap hari, data transaksi penjualan akan terus bertambah, menyebabkan penyimpanan data semakin besar. Data transaksi penjualan hanya dijadikan arsip tanpa dimanfaatkan dengan baik. Pada dasarnya kumpulan data memiliki informasi-informasi yang sangat bermanfaat.

Berdasarkan permasalahannya diperlukan adanya sistem untuk mengolah data barang berdasarkan kecenderungannya yang muncul bersamaan dalam suatu transaksi menggunakan algoritma apriori. Penerapan algoritma apriori diharapkan akan menemukan pola berupa produk yang sering dibeli bersamaan. Pola tersebut digunakan untuk menempatkan produk yang sering dibeli bersamaan dalam sebuah area yang saling berdekatan. Adapun hasil yang dicapai berupa laporan hasil data mining pola pembelian barang yang sering dibeli secara bersamaan guna pengembangan strategi pemasaran dalam penjualan barang.

Kesimpulannya membangun aplikasi penerapan data mining analisa pola pembelian produk dengan metode algoritma apriori untuk mengetahui perilaku konsumen dalam membeli produk barang secara bersamaan, sebagai alternative keputusan dalam menentukan penempatan barang di area yang saling berdekatan sesuai perilaku konsumen dalam membeli barang secara bersamaan.

Kata kunci: Data Mining, algoritma apriori, perilaku konsumen

1. Pendahuluan

Jumlah pasar swalayan yang terus berkembang membuat para pengelola swalayan juga dituntut untuk menerapkan strategi pemasaran yang lebih baik. Untuk itu maka para pengelola harus mencermati pola-pola pembelian yang dilakukan oleh konsumen. Penulis menemukan banyak kekurangan yang terjadi, diantaranya permasalahan

peletakkan barang-barang yang tidak sesuai dengan perilaku kebiasaan konsumen dalam membeli barang secara bersamaan dalam satu waktu. Hal ini tentu akan mempengaruhi tingkat penjualan barang.

Data transaksi penjualan akan terus bertambah setiap harinya dan menyebabkan penyimpanan data yang sangat besar. Kebanyakan data transaksi penjualan hanya dijadikan arsip saja tanpa dimanfaatkan dengan baik. Padahal kumpulan data tersebut memiliki informasi yang sangat bermanfaat.

Berdasarkan permasalahan di atas, maka diperlukan adanya aplikasi untuk mengelompokkan data barang berdasarkan kecenderungannya yang muncul bersamaan dalam suatu transaksi menggunakan algoritma apriori. Penerapan algoritma apriori dalam penelitian ini diharapkan akan menemukan pola berupa produk yang sering dibeli bersamaan. Pola tersebut bisa digunakan untuk menempatkan produk yang sering dibeli bersamaan dalam sebuah area yang saling berdekatan, merancang kupon diskon pada produk tertentu untuk menarik daya beli konsumen.

Rumusan Masalah

Adapun rumusan masalahnya adalah bagaimana membuat aplikasi data mining analisa pola pembelian produk dengan metode algoritma apriori untuk memberikan gambaran keterkaitan antar barang dengan menganalisis data transaksi penjualan berdasarkan pola berupa produk barang yang sering dibeli secara bersamaan.

Tujuan Penulisan

Tujuan penulisannya adalah membangun aplikasi penerapan data mining analisa pola pembelian produk dengan metode algoritma apriori untuk menemukan pola pembelian barang yang sering dibeli secara bersamaan guna pengembangan strategi pemasaran dalam penjualan barang.

Metodologi Penelitian

Metode penelitian yang digunakan dalam penerapan data mining adalah metode Waterfall. Pengembangan metode Waterfall sendiri melalui beberapa tahapan yaitu Analisis, Desain, Coding, Testing [1].

Tinjauan Pustaka

Pengertian Data Mining

- Data mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar [2].
- Data mining*, sering juga disebut sebagai *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [3].
- Data mining* didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar.

Karakteristik data mining :

- Data mining* berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
- Data mining* biasa menggunakan data yang sangat besar.
- Data mining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi [2].

Pengenalan Pola, Data Mining, dan Machine Learning

Keluaran dari *data mining* ini bisa dipakai untuk memperbaiki pengambilan keputusan di masa depan. Sehingga istilah *pattern recognition* jarang digunakan karena termasuk bagian dari *data mining* [4]. *Machine Learning* adalah suatu area dalam *artificial intelligence* atau kecerdasan buatan yang berhubungan dengan pengembangan teknik-teknik yang bisa diprogramkan dan belajar dari data masa lalu.

Tahap-Tahap Data mining

- Pembersihan data (*data cleaning*)
Merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.
- Integrasi data (*data integration*)
Merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.
- Seleksi Data (*Data Selection*)
Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.
- Data Transformation*
Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*
- Proses *mining*.

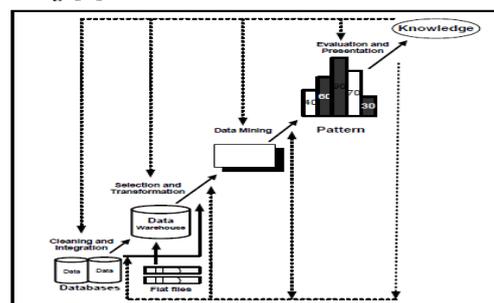
Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

- Evaluasi pola (*pattern evaluation*).

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan, bermanfaat.

- Presentasi pengetahuan.

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil *data mining* [5].



Gambar 1. Tahap-tahap Data Mining

Algoritma Apriori.

Algoritma apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule* [6]. Untuk mencari *association rule* dari suatu kumpulan data, tahap pertama yang harus dilakukan adalah mencari *frequent itemset* terlebih dahulu. *Frequent itemset* adalah sekumpulan item yang sering muncul secara bersamaan. Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu *support* dan *confidence*. *Support* adalah nilai penunjang atau persentase kombinasi sebuah *item* dalam database, sedangkan *confidence* adalah nilai kepastian yaitu kuatnya hubungan antar *item* dalam aturan asosiasi. Proses utama yang dilakukan dalam algoritma apriori untuk mendapat *frequent itemset* [6] yaitu :

1. *Join* (penggabungan).

Proses ini dilakukan dengan cara pengkombinasian item dengan yang item lainnya hingga tidak bisa terbentuk kombinasi lagi.

2. *Prune* (pemangkasan).

Proses pemangkasan yaitu hasil dari item yang telah dikombinasikan kemudian dipangkas dengan menggunakan *minimum support* yang telah ditentukan.

Konsep Aturan Asosiasi Rule.

Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiasi antara kombinasi item. Contoh dari aturan asosiasi dari analisa pembelian di suatu pasar swalayan adalah dapat diketahui berapa besar kemungkinan seseorang membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur

penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

Metodologi dasar analisis asosiasi

a. Analisis Pola Frekuensi Tinggi

Tahapan 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 2 berikut.

$$\text{Support (A,B)} = \frac{\sum \text{Transaksi Mengandung A dan B}}{\sum \text{Transaksi}}$$

b. Pembentukan Aturan Asosiasi

Pembentukan Aturan Asosiasi Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan *asosiasif* “ jika A maka B “. Nilai *confidence* dari aturan “ jika A maka B “ diperoleh dari rumus berikut :

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

2. Pembahasan

a. *Database* dari transaksi belanja pasar swalayan seperti ditunjukkan dalam tabel 1.

Tabel 1. Tabel Transaksi

Transaksi	Item yang di beli
1	Susu, Teh, Gula
2	Teh, Gula, Roti
3	Teh, Gula
4	Susu, Roti
5	Susu, Gula, Roti
6	Teh, Gula
7	Gula , Kopi, Susu
8	Gula , Kopi, Susu
9	Susu, Roti, Kopi
10	Gula, Teh, Kopi

b. Data pada Tabel 1 dalam *database* transaksi direpresentasikan seperti Tabel 2.

Tabel 2. Tabel Representasi Transaksi

Transaksi	Item yang dibeli
1	Susu
1	Teh
1	Gula
2	Teh
2	Gula
2	Roti
3	Teh

3	Gula
4	Susu
4	Roti
5	Susu
5	Gula
5	Roti
6	Teh
6	Gula
7	Gula
7	Kopi
7	Susu
8	Gula
8	Kopi
8	Susu
9	Susu
9	Roti
9	Kopi
10	Gula
10	Teh
10	Kopi

c. Data transaksi item yang dibeli dibuat dalam bentuk tabular seperti pada Tabel 3.

Tabel 3. Tabel Tabular

Transaksi	Teh	Gula	Kopi	Susu	Roti
1	1	1	0	1	0
2	1	1	0	0	1
3	1	1	0	0	0
4	0	0	0	1	1
5	0	1	0	1	1
6	1	1	0	0	0
7	0	1	1	1	0
8	0	1	1	1	0
9	0	0	1	1	1
10	1	1	1	0	0

d. Calon 2 *itemset* pada setiap data transaksi dan frekuensi masing-masing dihitung sesuai dengan data tabular pada table 3. Tabel 4 berikut ini menunjukkan calon 2 *itemset* dari data transaksi pada tabel 3 di atas.

Tabel 4. Tabel calon 2 itemset

Kombinasi	Jumlah
Teh, gula	5
Teh, kopi	1
Teh, susu	1
Teh, roti	1
Gula, kopi	3
Gula, susu	4
Gula, roti	2
Kopi, susu	3
Kopi, roti	1
Susu, roti	3

e. Dari data tersebut diatas tahap selanjutnya seleksi frekuensi \geq batas minimal yang ditentukan. Jika minimal *supportnya* ditetapkan nilai = 2, maka : $F_2 = \{ \{teh, gula\}, \{gula, kopi\},$

{gula, susu}, {gula, roti}, {kopi, susu}, {susu,roti} } seperti pada Tabel 5.

Tabel 5. Tabel calon 2 itemset

Kombinasi	Jumlah
Teh, gula	5
Gula, kopi	3
Gula, susu	4
Gula, roti	2
Kopi, susu	3
Susu, roti	3

f. Kombinasi 2 *itemset* dalam F_2 pada tabel 5 dapat di gabungkan menjadi calon 3 *itemset*. *Itemset-itemset* dari F_2 yang dapat digabungkan adalah *itemset-itemset* yang memiliki kesamaan dalam k-2 item pertama. Calon 3 *itemset* yang dapat dibentuk dari F_2 tampak pada Tabel 6.

Tabel 6. Tabel calon 3 itemset

Kombinasi	Jumlah
Teh, gula, susu	1
Teh, gula, roti	1
Gula, susu, roti	1
Gula, kopi, susu	2
Kopi, susu, roti	1
Teh, gula, kopi	1

Dengan demikian $F_3 = \{gula, kopi, susu\}$, karena hanya kombinasi inilah yang memiliki frekuensi kemunculan ≥ 2 .

g. Pembentukan Aturan Asosiasi

Pembentukan Aturan Asosiasi setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan *asosiasif* “ jika A maka B “. Nilai *confidence* dari aturan “ jika A maka B “

Berdasarkan Tabel 3, dapat dilihat besarnya nilai *support* dan *confidence* dari calon aturan asosiasi seperti tampak pada tabel 7.

Tabel 7. Tabel Aturan Asosiasi 3 itemset

Aturan	Confidence	
Jika membeli gula dan susu, maka akan membeli kopi	2/4	50%
Jika membeli gula dan kopi, maka akan membeli susu	2/3	67%
Jika membeli kopi dan susu, maka akan membeli gula	2/3	67%

Misalkan ditetapkan nilai *minimum confidence* 60%, maka aturan yang bisa terbentuk adalah dengan dua aturan berikut.

- Jika membeli gula dan kopi, maka akan membeli susu

- Jika membeli kopi dan susu, maka akan membeli gula

Calon aturan asosiasi dari F_2 dapat dilihat pada tabel 8 berikut.

Tabel 8. Tabel Aturan Asosiasi 2 itemset

Aturan	Confidence	
Jika membeli teh, maka akan membeli gula	5/5	100%
Jika membeli gula, maka akan membeli teh	5/8	62.5%
Jika membeli gula, maka akan membeli kopi	3/8	37.5%
Jika membeli kopi, maka akan membeli gula	3/4	75%
Jika membeli gula, maka akan membeli susu	4/8	50%
Jika membeli susu, maka akan membeli gula	4/6	67%
Jika membeli gula, maka akan membeli roti	2/8	25%
Jika membeli roti, maka akan membeli gula	2/4	50%
Jika membeli kopi, maka akan membeli susu	3/4	75%
Jika membeli susu, maka akan membeli kopi	3/6	50%
Jika membeli susu, maka akan membeli roti	3/6	50%
Jika membeli roti, maka akan membeli susu	3/4	75%

Aturan Asosiasi Final terurut berdasarkan *support x confidence* terbesar dapat dilihat pada Tabel 9

Tabel 9. Tabel Aturan Asosiasi Final

Aturan	Support	Confidence	Support x Confidence
Jika membeli teh, maka akan membeli gula	50%	100%	50.0%
Jika membeli gula, maka akan membeli teh	50%	62.50%	31.3%
Jika membeli susu, maka akan membeli gula	40%	67%	26.8%
Jika membeli kopi, maka akan membeli gula	30%	75%	22.5%
Jika membeli kopi, maka akan membeli susu	30%	75%	22.5%
Jika membeli roti, maka akan membeli susu	30%	75%	22.5%
Jika membeli gula dan kopi, maka akan membeli susu	20%	67%	13.4%

Jika membeli kopi dan susu, maka akan membeli gula	20%	67%	13.4%
--	-----	-----	-------

Tabel *final association rule* menjelaskan tentang *support* dan *confidence* dari masing-masing kombinasi 2 *itemsets* dan 3 *itemsets*. Hasil perhitungan support pada tabel *final association rule* didapatkan dari jumlah transaksi mengandung A dan B dibagi total transaksi. Sedangkan *confidence* didapatkan dari jumlah transaksi mengandung A dan B dibagi jumlah transaksi mengandung A. Hasil perkalian *support* dan *confidence* itulah yang menjadi hasil akhir dari algoritma apriori.

3. Kesimpulan

Data Mining dapat di implementasikan dengan menggunakan *database* penjualan produk barang untuk dapat menemukan kecenderungan pola kombinasi *itemsets* sehingga dapat dijadikan sebagai informasi untuk mengetahui perilaku konsumen dalam membeli produk barang secara bersamaan, sebagai *alternative* alat bantu keputusan dalam menentukan penempatan barang di area yang saling berdekatan sesuai perilaku konsumen dalam membeli barang secara bersamaan, membantu untuk mengetahui produk barang yang jarang dibeli konsumen dan sebagai alat *alternative* dalam meningkatkan strategi pemasaran dengan cara membuat diskon barang tertentu yang jarang di beli untuk menarik minat beli konsumen.

Saran

Penelitian selanjutnya diharapkan dapat menyajikan sistem proses *output itemset* barang lebih dari 3 *itemset*

Daftar Pustaka

- [1] Pressman, Roger S, 2002, *Rekayasa perangkat Lunak Pendekatan Praktisi (buku satu)*, Andi, Yogyakarta.
- [2] Davies, and Paul Beynon, 2004, "Database System Third Edition", Palgrave Macmillan, New York.
- [3] Pramudiono, I., 2007, *Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data*. diakses tanggal 15 mei 2015, dari <http://www.ilmukomputer.org/wp-content/uploads/2006/08/iko-datamining>.
- [4] Santoso, B., 2007, *Data Mining : Teknik pemanfaatan data untuk keperluan bisnis*, Graha Ilmu, Yogyakarta.
- [5] Han, J and Kamber, M, 2006 " *Data Mining Concepts and Techniques, Second Edition* ". Morgan Kaufman, San Francisco. <http://id.wikipedia.org/wiki/Elektronika>, 23 Maret 2015. <https://charistasfibriani.wordpress.com>
- [6] Erwin, 2009. *Analisis Market Basket dengan Algoritma Apriori dan FP-Growth*. Jurnal Generik Vol.4 No2. juli 2009.

Biodata Penulis

Heroe Santoso, M.Kom, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika dan Komputer STIKI Malang, lulus tahun 1995. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Ilmu Komputer Universitas Gajah Mada Yogyakarta, lulus tahun

2004. Saat ini menjadi Dosen di STMIK Bumigora Mataram.

I Putu Hariyadi, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika Sekolah Tinggi Teknik Atlas Nusantara (STTAR) Malang, lulus tahun 2004. Memperoleh gelar Magister Komputer (M.Kom) Program Pasca Sarjana Magister Teknologi Informasi Sekolah Tinggi Teknik Surabaya (STTS), lulus tahun 2015. Saat ini menjadi Dosen di STMIK Bumigora Mataram.

Prayitno, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika STMIK Bumigora Mataram, lulus tahun 2015.

