

Implementasi Algoritma Apriori Untuk Mengetahui Aturan Asosiasi Pada Data Transaksi Penjualan

Daryanto

Abstrak: *Banyaknya persaingan dunia bisnis penjualan palawija menuntut perencanaan strategi bisnis yang tepat, oleh karena itu diperlukan suatu strategi yang dapat meningkatkan penjualan dan pemasaran produk. Salah satunya dengan memanfaatkan data transaksi untuk menemukan pola berupa produk-produk yang sering dibeli bersamaan dalam satu transaksi. Metode yang dapat digunakan untuk menyelesaikan permasalahan ini adalah algoritma Apriori yang termasuk jenis aturan asosiasi pada data mining. Algoritma Apriori adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Diharapkan dengan algoritma Apriori akan diperoleh kombinasi produk-produk yang sering dibeli bersamaan di toko palawija, sehingga dapat ditentukan strategi pemasaran yang tepat.*

Kata Kunci : *Data Mining, Aturan Asosiasi, Algoritma Apriori*

I. PENDAHULUAN

Banyaknya persaingan dalam dunia bisnis khususnya dalam penjualan palawija, diperlukan suatu strategi yang dapat meningkatkan penjualan dan pemasaran barang di toko palawija, salah satunya dengan memanfaatkan data transaksi. Namun dengan adanya kegiatan operasional sehari-hari data semakin lama akan semakin bertambah banyak dan jumlah data yang begitu besar justru bisa menjadi masalah bagi toko palawija tersebut jika tidak bisa dimanfaatkan. Semakin banyak data di toko palawija tersebut maka semakin banyak data yang dapat diolah menjadi informasi. Oleh karena itu diperlukan sebuah aplikasi yang mampu memilah dan memilih data yang besar, sehingga dapat diperoleh informasi yang berguna.

Permasalahan ini dapat diselesaikan dengan data mining yaitu proses mengekstraksi informasi dari data yang ada di dalam database sehingga menghasilkan informasi yang penting. Teknik asosiasi merupakan teknik yang mengadaptasi ilmu data mining menggunakan algoritma Apriori, yang berfungsi untuk membentuk kandidat kombinasi item yang muncul bersamaan, lalu diuji apakah kombinasi tersebut memenuhi parameter support dan confidence minimum yang merupakan nilai ambang yang diberikan oleh user.

Pada penelitian ini, akan dibangun sebuah aplikasi yang mengelompokkan serta mencari pola dari sebuah produk yang sering muncul bersamaan atau cenderung muncul bersama dalam suatu transaksi yang pada umumnya berjumlah besar dengan teknik asosiasi (analisa keranjang pasar) menggunakan algoritma Apriori. Teknik ini digunakan untuk merancang suatu

strategi penjualan dan pemasaran barang di toko palawija melalui proses pencarian asosiasi atau hubungan antar item data dari suatu basis data :

II. TINJAUAN PUSTAKA

2.1 Data Mining

Seiring dengan perkembangan teknologi, semakin berkembang pula kemampuan seseorang dalam mengumpulkan dan mengolah data. Penggunaan sistem komputerisasi dalam berbagai bidang baik itu dalam transaksi-transaksi bisnis, maupun untuk kalangan pemerintah dan sosial, telah menghasilkan data yang berukuran sangat besar. Data-data yang terkumpul ini merupakan suatu tambang emas yang dapat digunakan sebagai informasi dalam dunia bisnis.

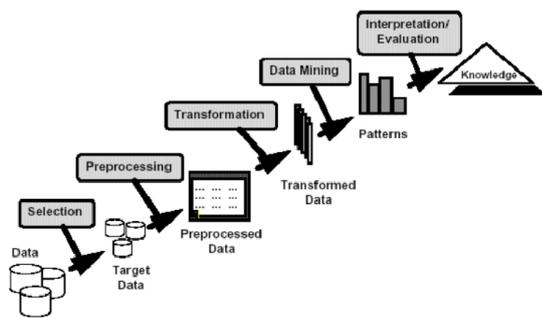
Aplikasi basis data telah banyak diterapkan dalam berbagai bidang, antara lain bidang manajemen, ilmu pengetahuan, administrasi pemerintah dan bidang-bidang lainnya. Akibatnya data yang dihasilkan oleh bidang-bidang tersebut sangatlah besar dan berkembang dengan cepat. Hal ini menyebabkan timbulnya kebutuhan terhadap teknik-teknik yang dapat melakukan pengolahan data sehingga dari data-data yang ada dapat diperoleh informasi penting yang dapat digunakan untuk perkembangan masing-masing bidang tersebut.

Istilah data mining sudah berkembang jauh dalam mengadaptasi setiap bentuk analisa data. Pada dasarnya data mining berhubungan dengan analisa data dan penggunaan teknik-teknik perangkat lunak untuk mencari pola dan keteraturan dalam himpunan data yang sifatnya tersembunyi.

Data Mining diartikan sebagai suatu proses mengekstraksi informasi dari data yang ada di dalam database sehingga menghasilkan informasi yang penting untuk mencari pola dan aturan yang berarti. *Data Mining* sangat diperlukan terutama dalam mengelola data yang sangat besar untuk memudahkan aktifitas *recording* suatu transaksi dan untuk proses data *warehousing* agar dapat memberikan informasi yang akurat bagi penggunaanya [1].

2.2 Tahapan Dalam Proses Data Mining

Dalam proses *data mining* ada beberapa tahapan seperti diagram di bawah ini yang menggambarkan beberapa tahap atau proses yang berlangsung dalam *data mining*. Fase awal di mulai dari data sumber dan berakhir dengan adanya informasi yang dihasilkan dari beberapa tahapan, yaitu: [2].



Gambar 2.2.1 Tahapan Dalam Proses Data Mining

2.3 Association Rule

Analisis asosiasi atau *association rule* adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Analisis asosiasi dikenal juga sebagai salah satu teknik *data mining* lainnya.[3]

Sebagai ilustrasi dari *association rule* pada permasalahan ini yaitu, dapat di ketahuinya berapa besar kemungkinan seorang pelanggan membeli kedelai bersamaan dengan ragi. Dengan pengetahuan tersebut, pemilik toko palawija dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan menggunakan kupon diskon untuk kombinasi barang tertentu atau menambah stok barang lebih banyak agar konsumen tetap beli barang di toko tersebut. Pada ilustrasi tersebut, pola yang ditemukan berpotensi menghasilkan potongan informasi yang menarik dan dibutuhkan oleh perusahaan yang terkait.

2.4 Algoritma Apriori

Algoritma *Apriori* adalah sebuah algoritma pencarian pola dalam teknik penambangan data (*data mining*) untuk menemukan aturan *asosiatif* antara suatu kombinasi item-set yang mempunyai suatu nilai keseringan tertentu sesuai kriteria atau filter yang diinginkan. Algoritma *Apriori* termasuk dalam jenis aturan asosiasi pada data mining. Proses di dalam Algoritma *Apriori* dalam teknik *association rules* adalah mencari aturan-aturan yang memenuhi minimum *support* dan *confidence*. [3]

Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, yaitu *support* dan *confidence*. *Support* (nilai penunjang) adalah suatu persentase kombinasi item tersebut dalam *database*. *Confidence* (nilai kepastian) adalah kuatnya hubungan antar item dalam aturan asosiasi. [3]

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap:

2.4.1 Analisis 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. [3]

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ Transaksi} \times 100\%$$

2.4.2 Pembentukan Aturan Asosiasi

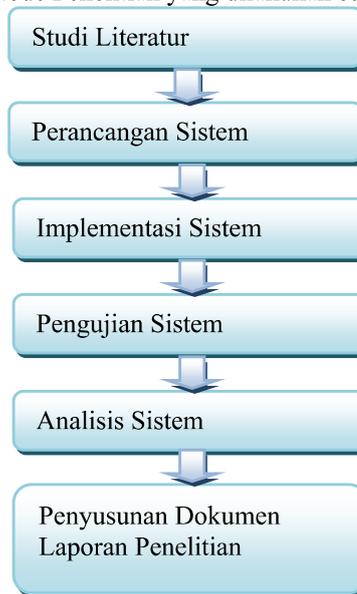
Setelah semua pola frekuensi tinggi di temukan, kemudian dicari aturan asosiasi yang memenuhi syarat minimum *confidence* dengan menghitung *confidence* aturan asosiasi $A \rightarrow B$. [3]

Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus berikut:

$$Confidence = P(P|A) = \frac{\sum Transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi\ mengandung\ A} \times 100\%$$

III. METODE PENELITIAN

Metode Penelitian yang dilakukan sebagai berikut :



Gambar 3.1 Tahapan Penelitian

IV. HASIL DAN PEMBAHASAN

Fungsi Association Rules seringkali disebut dengan "market basket analysis", yang digunakan untuk menemukan relasi atau korelasi diantara himpunan item2. Market Basket Analysis adalah Analisis dari kebiasaan membeli customer dengan mencari asosiasi dan korelasi antara item-item berbeda yang diletakkan customer dalam keranjang belanjanya.

Pengujian pada aplikasi keranjang pasar ini menggunakan 300 buah transaksi yang diambil dari Toko Trijaya Jember. Data yang diambil berupa data jenis barang yang terdiri dari 12 buah jenis barang dan 45 buah nama produk beserta stoknya. Pada pengujian ini beberapa nilai awal maupun asumsi yang digunakan dalam analisis data sebagai berikut:

- min-support : 30%
- min-confidence : 60%

Dari hasil pengolahan terhadap 300 data transaksi, dengan menggunakan algoritma *apriori* untuk masalah penggalian asosiasi didapat hasil sebagai berikut:

Iterasi 1:

Menentukan kandidat untuk 1-item_set dengan menentukan *support*. Data *support* disajikan dalam tabel 4.1. Misal pada baris 2 dibaca sebagai jumlah pembelian kedelai USA sebanyak 64 buah atau 21,3 % dari transaksi.

Support 1	$37/300 \times 100\% = 12,3\%$
Support 2	$64/300 \times 100\% = 21,3\%$
Support 3	$48/300 \times 100\% = 16\%$
.	
.	
.	
Support 44	$8/300 \times 100\% = 2,7\%$
Support 45	$18/300 \times 100\% = 6\%$

Tabel 4.1 Support Untuk Kandidat 1-Itemset

Nama Barang	Jumlah	Support %
Kedelai Galunggung	37	12,3
Kedelai USA BW	64	21,3
Kacang India	48	16
Gula DR	106	35,3
Tepung Cakra	39	13
Kopi BB	30	10
Tepung Payung	23	7,7
Kanji 99	24	8
Wijen India	29	9,6
Tepung Segitiga	11	3,7
Kacang Ijo Burek	15	5
Ketan Siam	14	4,7
Kanji SPM	13	4,3
Ketan Vietnam	26	8,7
Tepung Beras Rose	9	3
Tepung Kunci Biru	2	0,7
Tepung Lencana	15	5
Tepung CK Emas	13	4,3
Mentega Amanda	14	4,7
Kacang Ijo Kecil LN	15	5
Otok Putih Besar	9	3
Gula SBR	25	8,3
Beras Jagung Polos	12	4
Kacang Ijo Kupas	19	6,3
Kacang local	8	2,7
Ketan Item Local	10	3,3
Kedelai USA Bola	23	7,7
Beras Niki legi	8	2,7
Beras Kelinci	7	2,3
Gula thailand	11	3,7
Beras Padi Mas	9	3
Ragi Tempe	45	15
Ketan Item Bali	10	3,3
Beras IR Macan	11	3,7
Beras Dua Putra	12	4
Beras jagung super	11	3,7

Tepung Ketan Rose	8	2,7
Kacang Ijo Kecil	8	2,7
Jagung Madura	10	3,3
Jagung Super	9	3
Ketan Ayon	13	4,3
Tepung Segitiga Biru	17	5,7
Kopi GB Local	11	3,7
Kopi GB Nangka	8	2,7
Tepung Lencana	18	6

Keterangan:

 Item yang tidak memenuhi *min_support*.

 Item yang memenuhi *min_support*.

Menentukan frequent 1-itemset, dengan membuang itemset yang tidak memenuhi *min_support* 10%. Tampak dari tabel 4.1, item yang tidak diberi warna adalah item yang tidak memenuhi *min_support* dan berjumlah 39 buah jenis barang, sehingga dapat dipangkas. Frequent 1-itemset yang didapat berjumlah 6 buah jenis barang.

Tabel 4.2 *Support_count* untuk kandidat 1-itemset

Nama Barang	Jml	Sup
Kedelai Galunggung	37	12,3
Kedelai USA BW	64	21,3
Kacang India	48	16
Gula DR	106	35,3
Tepung Cakra Kembar	39	13
Kopi BB	30	10

Iterasi 2:

Menentukan kandidat 2-itemset dengan menentukan *support* dan *confidence*. Data untuk kandidat frequent 2-itemset disajikan dalam tabel 4.3, berisi data yang telah dipangkas pada pencarian frequent 2-itemset.

Dengan menggunakan rumus :

$$\text{Support}(A) = \frac{\text{Jumlah transaksi mengandung } A}{\text{Total Transaksi}} \times 100\%$$

$$\text{Confidence} = P(P|A) = \frac{\sum \text{Transaksi mengandung } A \text{ dan } B}{\sum \text{Transaksi mengandung } A} \times 100\%$$

Support 1	$90/300 \times 100\% = 30,03\%$
Confidence 1	$90/37 \times 100\% = 85\%$
Support 2	$90/300 \times 100\% = 30,03\%$
Confidence 2	$83/64 \times 100\% = 79,6\%$

Support 12 $90/300 \times 100\% = 30,80\%$
 Confidence 12 $90/17 \times 100\% = 74,34\%$

Hasil Asosiasi dengan *Minsup* 30% dan *Mincof* 60%

Tabel 4.3 Hasil Asosiasi dari Data Transaksi Penjualan

Nama Barang	Support	Confidence
Kedelai Galunggung → Gula DR	30.03%	85.00%
Gula DR → Kedelai Galunggung	30.03%	79.06%
Kedelai USA BW → Gula DR	30.03%	85.00%
Gula DR → Kedelai USA BW	30.03%	63.77%
Kacang India → Gula DR	30.03%	85.00%
Gula DR → Kacang India	30.03%	72.83%
Gula DR → Tepung Cakra Kembar	30.13%	85.28%
Gula DR → Kacang Ijo Burek	31.33%	77.36%
Gula DR → Tepung Lencana Merah	31.33%	77.36%
Gula DR → Kacang Ijo Kecil LN	31.33%	77.36%
Gula DR → Kacang Ijo Kupas	30.27%	71.32%
Gula DR → Kopi BB	30.80%	74.34%

Display untuk layout dari hasil asosiasi di atas untuk penempatan barang dengan ketentuan *minimum support* barang A dengan barang B harus sama diperoleh sebagai berikut:

Tabel 4.4 Hasil Display

-----display-----
Kedelai Galunggung -- Gula DR
Kedelai USA BW -- Gula DR
Kacang India -- Gula DR
Gula DR -- Tepung Cakra Kembar
Gula DR -- Kacang Ijo Burek
Gula DR -- Tepung Lencana Merah
Gula DR -- Kacang Ijo Kecil LN
Gula DR -- Kacang Ijo Kupas
Gula DR -- Kopi BB

Dari hasil asosiasi tersebut didapatkan strategi pemasaran salah satunya untuk *discount* dengan ketentuan barang yang muncul bersamaan terdapat salah satu *support* harus lebih besar seperti dibawah ini:

Tabel 4.5 Hasil Discount

-----Discount-----
Gula DR → Kedelai Galunggung
Gula DR → Kedelai USA BW
Gula DR → Kacang India
Gula DR → Tepung Cakra Kembar
Gula DR → Kacang Ijo Burek
Gula DR → Tepung Lencana Merah
Gula DR → Kacang Ijo Kecil LN
Gula DR → Kacang Ijo Kupas
Gula DR → Kopi BB

V. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pengujian yang dilakukan pada bagian sebelumnya, maka kesimpulan yang dapat diambil adalah:

1. Semakin banyak data transaksi, maka semakin sulit memenuhi *support* di atas 30%. Karena jika data barang yang di beli bersamaan sedikit, maka nilai *support*-nya semakin kecil.
2. Semakin banyak barang yang dibeli dan yang mendekati salah satu transaksi yang dibeli bersamaan maka semakin tinggi nilai *confidence*-nya.

Semakin sedikit barang yang dibeli bersamaan maka semakin rendah pula nilai *confidence*-nya

VI. DAFTAR PUSTAKA

- [1] Bertalya. (2009). *Konsep Data Mining*. Universitas Gunadarma. Depok
- [2] Dinda, Devi. (2009). *Penggunaan Metode Apriori Untuk Analisa Keranjang Pasar Pada Data Transaksi Penjualan Minimarket Menggunakan Java & Mysql*. Teknik Informatika. Universitas Gunadarma, Depok.
- [3] Kusri dan taufiq, Emha. (2009). *Algoritma Data Mining*. CV. Andi Offset, Jogjakarta