



APRIORI ALGORITHM USE FOR A CONSUMER BEHAVIOR IN THE PURCHASE OF GOODS

Azwar Anas¹

¹STIE - Graha Karya Muara Bulian-Jambi Indonesia

<http://dx.doi.org/10.22216/jsi.2015.02.01.494-1111>

<i>Article History</i>	<i>Abstract</i>
Received : August 2015	<i>Data owned by an institution is one of the assets of the institution. Their daily trade transactions will further multiply the number of transaction data. The amount of transaction data so large it could be a problem if traders can not use it wisely. In this study, researchers will utilize a priori algorithm to group data based on the purchase of goods tendency to appear together in a purchase of goods. The method I use in this case is a priori algorithm. Meanwhile, to conduct tests on the a priori algorithm analysis, the authors use data mining software Orange. The end result of this research is the description of consumer behavior towards the purchase of goods, so that could be a reference to the placement of goods in accordance with their preferences purchased by consumers. In this study the authors simply using a priori algorithm, the authors further expected to be able to combine a couple of similar algorithms, in order to produce more accurate information.</i>
Accepted : September 2015	
Published : December 2015	

<i>Keywords</i>
<i>Apriori Algorithm; Conusmen; Data Mining; Transaction.</i>

ALGORITHMMA APRIORI UNTUK MENDAPATKAN PERILAKU KONSUMEN DALAM PEMBELIAN BARANG

Abstrak

Data yang dimiliki oleh suatu institusi merupakan salah satu aset dari institusi tersebut. Adanya transaksi perdagangan sehari-hari akan semakin memperbanyak jumlah data transaksi. Jumlah data transaksi yang begitu besar justru bisa menjadi masalah bila pedagang tidak bisa dengan bijak memanfaatkannya. Dalam penelitian ini, peneliti akan memanfaatkan algoritma apriori untuk mengelompokkan data pembelian barang berdasarkan kecenderungannya yang muncul bersamaan dalam suatu transaksi pembelian barang. Metode yang penulis gunakan dalam hal ini adalah algoritma apriori. Sementara untuk melakukan uji coba terhadap analisis algoritma apriori, penulis menggunakan software data mining Orange. Hasil akhir dari penelitian ini adalah gambaran perilaku konsumen terhadap pembelian barang, sehingga bisa menjadi rujukan penempatan barang sesuai dengan kecenderungannya dibeli oleh konsumen. Pada penelitian ini penulis hanya menggunakan metode algoritma apriori, bagi penulis selanjutnya diharapkan mampu mengkombinasikan beberapa algoritma sejenis, agar menghasilkan informasi yang lebih akurat.

Corresponding author:
email: aans_07@yahoo.co.id

ISSN : 2459-9549
e-ISSN : 2502-096X

PENDAHULUAN

Perdagangan adalah adanya transaksi antara penjual dan pembeli. Proses jual-beli akan terjadi apabila adanya transaksi pembelian barang oleh konsumen kepada pedagang, baik pedagang di pasar tradisional, modern, mini market ataupun took-toko kecil lainnya. Dalam pembelian barang, konsumen biasanya membeli berbagai jenis barang sekaligus. Hal ini berdasarkan upaya pemenuhan kebutuhan pokok sehari-hari.

Saat ini, hampir semua mini market telah menggunakan program komputer untuk merekam data transaksi pembelian barang. Setiap transaksi pembelian barang, akan tersimpan dalam sebuah bank data pada komputer. Dengan adanya transaksi setiap hari, semakin lama data akan semakin banyak dan menumpuk pada bank data. Adanya transaksi perdagangan sehari-hari akan semakin memperbanyak jumlah data transaksi. Jumlah data transaksi yang begitu besar justru bisa menjadi masalah bila pedagang tidak bisa dengan bijak memanfaatkannya. Data yang menumpuk hanya akan menjadi sampah apabila tidak digunakan dengan baik. Padahal data yang menumpuk merupakan salah satu aset untuk diambil manfaatnya.

Untuk dapat mengambil manfaat dari data masa lalu yang menumpuk tersebut, perlu dilakukan penambangan data atau istilah lainnya adalah data mining. Dalam penelitian ini, peneliti akan memanfaatkan algoritma apriori untuk mengelompokkan data pembelian barang berdasarkan kecenderungannya yang muncul bersamaan dalam suatu transaksi pembelian barang.

Dalam penelitian ini, peneliti akan menggunakan algoritma apriori untuk menganalisa data transaksi perdagangan dan menggunakan *software Orange* untuk melakukan uji coba terhadap hasil analisa tersebut.

Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Selain apriori, yang termasuk dalam golongan ini adalah metode *Generalized Rule Induction* dan *Algoritma Hash Based*. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut *affinity analysis* atau *market basket analysis*.

Penerapan *algoritma apriori* telah banyak digunakan sebelumnya untuk mendapatkan informasi berharga dari sejumlah frekuensi data. Berikut beberapa penerapan algoritma apriori yang telah dilakukan :

1. Penerapan Data Mining Algoritma Asosiasi untuk Meningkatkan Penjualan (Emha Taufiq Luthfi, 2009).
2. Analisis *Market Basket* dengan Algoritma Apriori dan FP-Growth (Erwin (Jurusan Teknik Informatika, Fakultas Ilmu Komputer, 2009).
3. Implementasi *Association Rule* terhadap Penyusunan *Layout* Makanan dan Penentuan Paket Makanan Hemat di RM Roso Echo dengan Algoritma Apriori (Widiati & Evita Dewi, 2014).
4. Implementasi Data Mining pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori (Siregar, 2014).
5. Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat (Yanto & Khoiriah, 2015).

Tugas dari data mining adalah untuk menghasilkan semua kaidah asosiasi pada suatu tabel transaksional, yang mempunyai nilai *support* lebih dari *minimum support*. Kaidah tersebut juga harus mempunyai *support* yang lebih besar dari *confidences*.

Definisi Data Mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Studi, Informasi, Ilmu, Univer-

sitas, & Kuning, 2016). Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Pane, 2013).

Teknik – teknik *Data Mining*

Data mining dibagi menjadi beberapa teknik berdasarkan tugas yang dapat dilakukan, yaitu (Tampubolon et al., 2013):

1. Deskripsi
Terkadang analis/peneliti ingin mendeskripsikan pola dan trend yang tersimpan dalam data.
2. Estimasi
Estimasi mirip dengan klasifikasi, kecuali variabel tujuan yang lebih ke arah numeric daripada kategori. Misalnya, akan dilakukan estimasi tekanan darah *systolic* dari pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah.
3. Prediksi
Prediksi memiliki kemiripan dengan estimasi dan klasifikasi. Hanya saja, prediksi hasilnya menunjukkan sesuatu yang belum terjadi (mungkin terjadi di masa depan).
4. Klasifikasi
Dalam klasifikasi variabel, tujuan bersifat kategorik. Misalnya, kita akan mengklasifikasikan pendapatan dalam 3 kelas, yaitu mendapatkan tinggi, pendapatan sedang, dan pendapatan rendah.
5. *Clustering*
Clustering lebih ke arah pengelompokan *record*, pengamatan, atau kasus dalam kelas yang memiliki kemiripan. Sebuah *cluster* adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lain dan

memiliki ketidakmiripan dengan *record-record* dalam *cluster* yang lain.

6. Asosiasi

Mengidentifikasi hubungan antara berbagai peristiwa yang terjadi pada satu waktu. Pendekatan asosiasi tersebut menekankan sebuah kelas masalah yang dicirikan dengan analisis keranjang pasar.

Algoritma-algoritma *data mining* kebanyakan berasal atau merupakan bentuk pengembangan dari algoritma-algoritma dari bidang ilmu *machine learning*, statistika, intelegensia buatan dan jaringan syaraf tiruan. Karena tidak dirancang untuk menangani data dalam ukuran yang sangat besar, sedangkan *data mining* dimaksudkan untuk menangani data dalam ukuran yang demikian, maka salah satu arah penelitian di bidang *data mining* adalah mengembangkan algoritma-algoritma tersebut agar dapat menangani data yang berukuran sangat besar.

Selain itu, integrasi teknik-teknik *data mining* ke dalam DBMS, khususnya *Object-Relation DBMS* (ORDBMS) yang merupakan teknologi DBMS terbaru, juga masih merupakan bidang penelitian yang aktif. Tujuan utama dari integrasi ini adalah untuk “melebur” algoritma *data mining* agar menjadi fungsi internal ORDBMS yang berkualitas, sehingga pengguna dapat menggunakannya sesuai dengan kebutuhan. Karena DBMS merupakan teknologi yang sudah matang, digunakan secara luas, dapat mengelola data dalam ukuran yang sangat besar, memfasilitasi kueri tabel yang mudah dengan *Structured Query Language* (SQL) dan memiliki fitur yang mengakomodasi kebutuhan pengguna ORDBMS, maka pengembangan dan peleburan algoritma *data mining* ke dalam ORDBMS yang baik adalah dengan memanfaatkan semua ini.

Penerapan *Data Mining*

Banyak penerapan yang dapat dilakukan oleh *data mining*. Apalagi ditunjang kekayaan dan keanekaragaman bidang ilmu (*artificial intelligence*, *database*, statistik, pemodelan matematika, pengolahan citra, dsb) membuat penerapan *data mining* menjadi makin luas. Bidang penerapan *data mining* dapat dilakukan diantaranya adalah :

1. Analisa Pasar dan Manajemen

Untuk analisa pasar, banyak sumber data yang dapat digunakan seperti transaksi kartu kredit, kartu anggota klub tertentu, kupon diskon, keluhan pembeli, ditambah dengan studi tentang gaya hidup publik. Beberapa kondisi yang bisa diselesaikan dengan *data mining* diantaranya :

- a. Menembak target pasar
- b. Melihat pola beli pemakai dari waktu ke waktu
- c. *Cross Market Analysis*
- d. Profil *customer*
- e. Identifikasi kebutuhan *customer*
- f. Menilai loyalitas *customer*
- g. Informasi *summary*

2. Analisa Perusahaan dan Manajemen Resiko

- a. Perencanaan keuangan dan evaluasi aset

Data mining dapat membantu melakukan analisis dan prediksi *cash flow* serta melakukan *contingent claim analysis* untuk mengevaluasi aset. Selain itu juga dapat menggunakannya untuk analisis tren.

- b. Perencanaan sumber daya (*Resource Planning*)

Dengan melihat informasi ringkas (*summary*) serta pola pembelajaran dan pemasukan dari masing-masing *resource*, dan dapat memanfaatkannya untuk melakukan *resource planning*.

- c. Persaingan (*Competition*)

1) Sekarang ini banyak perusahaan yang berupaya untuk

dapat melakukan *competitive intelligence*. *Data mining* dapat membantu memonitor pesaing-pesaing dan melihat *market direction*.

- 2) Dapat melakukan pengelompokan *customer* dan memberikan variasi harga/ layanan/ bonus untuk masing-masing grup.

3) Menyusun strategi penetapan harga di pasar yang sangat kompetitif. Hal ini diterapkan oleh perusahaan minyak REPSOL di Spanyol dalam menetapkan harga jual gas di pasaran.

3. Telekomunikasi

Sebuah perusahaan telekomunikasi menerapkan *data mining* untuk melihat dari jutaan transaksi yang masuk, transaksi mana sajakah yang masih harus ditangani secara manual (dilayani oleh orang). Tujuannya tidak lain adalah untuk menambah layanan otomatis khusus untuk transaksi-transaksi yang masih dilayani secara manual. Dengan demikian jumlah operator penerima transaksi manual tetap bisa ditekan minimal.

4. Keuangan

Financial Crimes Enforcement Network di Amerika Serikat baru-baru ini menggunakan *data mining* untuk menambang trilyunan dari berbagai subjek seperti properti, rekening bank dan transaksi keuangan lainnya untuk mendeteksi transaksi-transaksi keuangan yang mencurigakan (seperti *money laundry*), yang susah dilakukan jika menggunakan analisa standar.

5. Asuransi

Australian Health Insurance Commission menggunakan *data mining* untuk mengidentifikasi layanan kesehatan yang sebenarnya tidak perlu tetapi tetap dilakukan oleh peserta

asuransi. Hasilnya, berhasil menghemat satu juta dolar per tahunnya. Tentu saja ini tidak hanya bisa diterapkan untuk asuransi kesehatan, tetapi juga untuk berbagai jenis asuransi lainnya.

6. Olah Raga

IBM Advanced Scout menggunakan *data mining* untuk menganalisis statistik permainan NBA (jumlah *shots blocked*, *assist* dan *fouls*) dalam rangka mencapai keunggulan bersaing (*competitive advantage*) untuk tim New York Knicks dan Miami Heat.

7. Astronomi

Jet Propulsion Laboratory (JPL) di Pasadena, Clifornia dan *Palomar Observatory* berhasil menemukan 22 quasar dengan bantuan *data mining*. Hal ini merupakan salah satu kesuksesan penerapan *data mining* di bidang astronomi dan ilmu ruang angkasa.

8. Internet Web Surf-Aid

IBM Surf-Aid menggunakan algoritma *data mining* untuk mendata akses halaman web khususnya yang berkaitan dengan pemasaran guna melihat perilaku dan minat *customer* serta melihat keefektifan pemasaran melalui web.

Data mining merupakan salah satu aktivitas di bidang perangkat lunak yang dapat memberikan ROI (*return on investment*) yang tinggi. Namun demikian *data mining* tetaplah hanya alat bantu yang dapat membantu manusia untuk melihat pola, menganalisis *trend* dalam rangka mempercepat pembuatan keputusan.

Association Rule

Association rule merupakan salah satu metode yang bertujuan mencari pola yang sering muncul di antara banyak transaksi, dimana setiap transaksi terdiri dari beberapa *item* (Wandi, Hendrawan, & Mukhlason, 2012). Contoh dari aturan asosiatif dari analisa pembelian di suatu

pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersama dengan susu. Dengan pengetahuan tersebut, pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu. Karena analisis asosiasi menjadi terkenal karena aplikasinya untuk menganalisa isi keranjang belanja di pasar swalayan, analisis asosiasi juga sering disebut dengan istilah *market basket analysis*.

Association rule merupakan teknik *data mining* untuk mencari pola hubungan dalam data atau basis data. Yang paling populer adalah *market basket analysis* (MBA). Teknik MBA ini digunakan untuk mencari pelanggan yang membeli barang pada took X, akan membeli barang Y dan seterusnya (Teknik & Fakultas, 2008).

Tabel 1.
Contoh Transaksi untuk
Association Rule Analysis

Trans	Itemset
1	A, B, E
2	B, D
3	B, C
4	A, B, D
5	A, C
6	B, C
7	A, C
8	A, B, C, E
9	A, B, C

Dalam *association analysis itemset* merupakan kumpulan nol atau lebih item. Pada contoh tabel di atas diumpamakan *itemset* adalah barang yang dibeli pada took tersebut, trans adalah kode transaksi. Pada tabel tersebut dapat dilihat pada transaksi 1 yang dibeli adalah barang A, B dan E. Transaksi 2 yang dibeli adalah B dan D, demikian seterusnya. *Association analysis* akan melakukan analisis hubungan dengan aturan tertentu sering disebut sebagai *association rule*.

Sebagai contoh tabel tersebut, pembeli yang biasanya membeli barang A dan B akan membeli barang C. aturan ini dapat ditulis $\{A, B\} \rightarrow \{C\}$. Untuk memperoleh kuantifikasinya terminologi *support* dan *confidence* perlu diperhatikan. *Support* dan *confidence* dapat dihitung sebagai berikut :

$$\text{Support, } s(X \rightarrow Y) = \frac{\sigma(XUY)}{N} \quad (1)$$

$$\text{Confidence, } c(X \rightarrow Y) = \frac{\sigma(XUY)}{\sigma_X} \quad (2)$$

Di mana X dan Y sebagai *itemset* dan σ adalah *support* dari *itemset*.

Dengan rumus tersebut maka *support* yang didapat dari hubungan $s(A, B \rightarrow C)$ adalah : $\frac{2}{9} = 0,22$. Sedangkan nilai *confidence* yang diperoleh adalah $c(A, B \rightarrow C)$ adalah : $\frac{2}{4} = 0,5$. Nilai 2 pada *support* diperoleh dari *itemset* A, B, C berjumlah 2 yaitu pada trans 8 dan 9, jumlah keseluruhan trans adalah 9. Sedangkan 2 pada *confidence* sama seperti pada *support*, 4 diperoleh dari jumlah A, B yaitu pada trans 1, 4, 8 dan 9. Dari perolehan tersebut, kita dapat membuat aturan persentasi yang diinginkan untuk menentukan kepastian asosiasinya. Paling baik adalah apabila nilai *support* maupun *confidence* tinggi.

Algoritma Apriori

Apriori adalah suatu algoritma untuk melakukan pencarian *frequent itemset* untuk mendapatkan *association rules*. Sesuai dengan namanya, algoritma ini menggunakan *prior knowledge* mengenai *frequent itemset properties* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Apriori menggunakan pendekatan secara *iterative* yang disebut juga sebagai *level-wise search* di mana *k-itemset* digunakan untuk mencari $(k+1)$ -*itemset*^[11].

Algoritma apriori menggunakan *knowledge* mengenai *frequent itemset*

yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Pola frekuensi tinggi adalah pola-pola item dalam *database* yang memiliki frekuensi atau *support* di atas ambang batas tertentu yang disebut istilah minimum *support*. Pola frekuensi tinggi ini digunakan untuk menyusun aturan asosiatif dan juga beberapa teknik *data mining* lainnya.

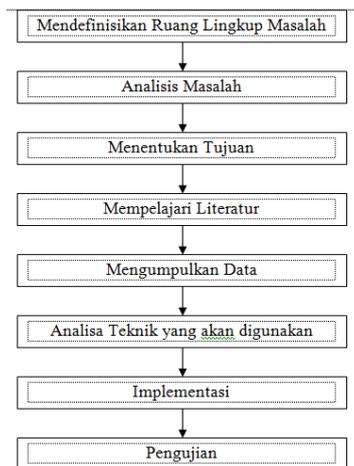
Prinsip dari algoritma apriori adalah

1. Kumpulkan jumlah item tunggal, dapatkan item besar
2. Dapatkan *candidate pairs*, hitung \rightarrow *large pairs* dari item-item
3. Dapatkan *candidate triplets*, hitung \rightarrow *large triplets* dari item-item dan seterusnya
4. Sebagai petunjuk : setiap *subset* dari sebuah *frequent itemset* harus menjadi *frequent*

Dua proses utama dalam algoritma apriori merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*. Walaupun algoritma apriori mudah untuk dipahami dan diimplementasikan dibandingkan algoritma yang lainnya yang memang diterapkan untuk proses *association rule*, akan tetapi algoritma apriori juga memiliki kekurangan yaitu untuk melakukan pencarian *frequent itemset*, algoritma apriori harus melakukan *scanning database* berulang kali untuk setiap kombinasi item. Hal tersebut menyebabkan banyaknya waktu yang dibutuhkan untuk melakukan *scanning database*. Selain itu dibutuhkan *generate candidate* yang besar untuk mendapatkan kombinasi item dari *database*

METODE PENELITIAN

Kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas. Adapun kerangka kerja penelitian ini dapat digambarkan pada gambar 1.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja pada gambar 1. maka masing-masing langkahnya dapat diuraikan seperti berikut ini :

1. Ruang Lingkup Masalah

Ruang lingkup masalah yang akan diteliti harus ditentukan terlebih dahulu, karena tanpa mampu menentukan serta mendefinisikan batasan masalah yang akan diteliti, maka tidak akan pernah didapat suatu solusi yang terbaik dari masalah tersebut. Jadi langkah pertama ini adalah langkah awal yang terpenting dalam penulisan ini.

2. Analisis Masalah

Langkah analisis masalah adalah untuk dapat memahami masalah yang telah ditentukan ruang lingkup atau batasannya. Dengan menganalisa masalah yang telah ditentukan tersebut, maka diharapkan masalah dapat dipahami dengan baik. Pada bagian ini digambarkan proses untuk menentukan pola peminjaman buku perpustakaan dengan menggunakan algoritma apriori. Disamping itu juga diuraikan tentang kebutuhan data untuk menganalisa pola peminjaman buku perpustakaan. Dari hasil analisa tersebut, maka algoritma apriori hendaknya mampu menentukan pola peminjaman buku perpustakaan dengan baik.

3. Penentuan Tujuan

Berdasarkan pemahaman dari masalah, maka ditentukan tujuan yang akan dicapai dari penulisan ini. Pada tujuan ini ditentukan target yang dicapai, terutama yang dapat mengatasi masalah-masalah yang ada.

4. Mempelajari Literatur

Untuk mencapai tujuan, maka dipelajari berbagai literatur yang relevan dengan masalah yang diteliti. Kemudian literatur-literatur yang dipelajari tersebut diseleksi untuk dapat ditentukan literatur-literatur mana yang akan digunakan dalam penelitian.

Literatur diambil dari internet, yang berupa artikel dan jurnal ilmiah tentang *data mining*, *association rule* dan algoritma apriori, serta bahan bacaan lain yang mendukung penelitian.

5. Pengumpulan Data dan Informasi

Dalam pengumpulan data dilakukan observasi yaitu pengamatan secara langsung di tempat penelitian sehingga permasalahan yang ada dapat diketahui dengan jelas. Kemudian dilakukan *interview* yang bertujuan untuk mendapatkan informasi atau data yang dibutuhkan. Selain itu juga dilakukan studi kepustakaan yaitu dengan membaca buku-buku yang menunjang dalam melakukan analisa terhadap data dan informasi yang didapat.

Adapun data-data yang diperlukan dalam penelitian ini adalah:

- Data pengunjung perpustakaan berdasarkan total jumlah siswa
- Data pengunjung perpustakaan berdasarkan tingkat kunjungan harian
- Data pengunjung perpustakaan dari pihak guru dan staff
- Data buku-buku, majalah, jurnal dan artikel koleksi perpustakaan lainnya.

6. Analisa Teknik yang digunakan
Bagian ini bertujuan untuk menganalisis dan memahami teknik yang akan digunakan dalam pengolahan data yang telah diperoleh dari tempat penelitian, terutama pada proses, yang mana teknik yang digunakan untuk pengolahan data dengan algoritma apriori.
7. Implementasi
Setelah dilakukan analisa teknik yang digunakan, maka dilakukan implementasi terhadap *software* dalam hal ini adalah *software data mining Orange*. Tentu saja penerapan analisa pada *software* tersebut menggunakan algoritma apriori. Jika penerapan sistem sudah berjalan dengan lancar dan akurat, maka sistem dapat diimplementasikan sesuai dengan kebutuhan.
8. Pengujian
Pengujian dilakukan untuk mendapatkan hasil yang sesuai dengan melakukan perbandingan antara perhitungan manual dan komputerisasi, berikut ini langkah-langkah pengujian yang akan dilakukan :
 - a. Melakukan pengujian algoritma apriori dengan data kepustakaan untuk menentukan pola pengunjung perpustakaan secara manual
 - b. Setelah dilakukan perhitungan manual, selanjutnya dilakukan perhitungan menggunakan *software data mining Orange* dengan memasukkan data-data kepustakaan dan diproses menggunakan algoritma apriori.
 - c. Tahap terakhir adalah melakukan perbandingan antara perhitungan manual dengan komputerisasi, jika hasil yang dicapai sama atau hampir sama, maka data yang diproses dan teknik yang digunakan telah benar

HASIL DAN PEMBAHASAN

Analisa Asosiasi dengan Algoritma Apriori

Dalam penelitian ini, penulis hanya melakukan analisa terhadap 50 konsumen yang melakukan transaksi pembelian barang. Dari 50 transaksi yang telah diamati, penulis hanya mengamati konsumen yang membeli maksimal 3 jenis barang.

Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menentukan aturan asosiatif antara aturan kombinasi item. Contoh dari aturan asosiatif dari analisis pembelian barang adalah mengetahui besarnya kemungkinan seorang konsumen untuk membeli Gula bersamaan dengan Kopi. Dengan pengetahuan tersebut, pengelola pedagang dapat mengatur penempatan barang dagangannya.

Dalam menentukan perilaku konsumen dalam melakukan pembelian barang, akan digunakan algoritma apriori untuk menganalisisnya. Data-data perdagangan yang telah dikumpulkan akan diolah menggunakan prinsip kerja algoritma apriori yang mampu memecahkan data bersifat *frequent itemset* atau mencari pola berulang.

Aturan asosiasi biasanya dinyatakan dalam bentuk :

$$\{Gula, Kopi\} \rightarrow \{Garam\} \{support = 40\%, confidence = 50\% \}$$

Artinya 50% dari transaksi pembelian Gula dan Kopi juga memuat Garam. Sementara 40% dari seluruh transaksi peminjaman yang ada memuat ketiga *item* tersebut.

Menentukan Variabel

Data yang akan diolah dalam menentukan perilaku konsumen ini adalah data pembelian barang. Adapun variabel-variabel data pembelian barang yang akan diolah adalah:

1. Pembeli
2. Tanggal
3. Barang 1
4. Barang 2
5. Barang 3

Dari variabel-variabel di atas, setiap konsumen akan dibagi kedalam beberapa kelompok berdasarkan jenis barang yang sama, hal ini dilakukan agar proses analisa data dapat dilakukan dengan mudah menggunakan algoritma apriori.

Analisis Frekuensi Tinggi

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

$$Support(A) =$$

$$\frac{\text{Jumlah transaksi mengandung } A}{\text{Jumlah transaksi}}$$

Sedangkan nilai *support* dari 2 item diperoleh dari rumus berikut :

$$Support(A,B) = P(A \cap B)$$

$$Support(A,B) =$$

$$\frac{\sum \text{Jumlah transaksi mengandung } A \text{ dan } B}{\sum \text{Jumlah transaksi}}$$

Melakukan Proses

Berdasarkan variabel-variabel data konsumen di atas, maka format data dapat dilihat pada tabel berikut ini.

Tabel 2.
Data Pembelian Barang

Pem beli	Tang gal	Barang 1	Barang 2	Barang 3
1	1-8-16	Gula	Kopi	Garam
2	1-8-16	Beras		
3	1-8-16	Rokok	Teh	
4	1-8-16	Sampo	Pena	Cuka
5	1-8-16	Garam	Beras	
6	1-8-16	Cuka	Buku	Permen
7	1-8-16	Roti	Gula	Pena
8	1-8-16	Teh	Beras	
9	1-8-16	Garam	Clear	
10	1-8-16	Silet	Cuka	Sendok
11	1-8-16	Sendok	Clear	
12	1-8-16	Rinso	Kopi	Vape
13	1-8-16	Rokok	Piring	
14	1-8-16	Clear	Kopi	
15	1-8-16	Gula	Tepung	Roti

16	1-8-16	Vape	Roti	Tisu
17	1-8-16	Clear		
18	1-8-16	Cuka	Pena	Silet
19	1-8-16	Rexona		
20	1-8-16	Sampo	Gula	Teh
21	1-8-16	Garam	Clear	
22	1-8-16	Silet	Cuka	Sendok
23	1-8-16	Sendok	Clear	
24	1-8-16	Rinso	Kopi	Vape
25	1-8-16	Rokok	Piring	
26	2-8-16	Clear	Kopi	
27	2-8-16	Gula	Tepung	Roti
28	2-8-16	Vape	Roti	Tisu
29	2-8-16	Clear		
30	2-8-16	Garam	Clear	
31	2-8-16	Silet	Cuka	Sendok
32	2-8-16	Sendok	Clear	
33	2-8-16	Rinso	Kopi	Vape
34	2-8-16	Rokok	Piring	
35	2-8-16	Clear	Kopi	
36	2-8-16	Gula	Tepung	Roti
37	2-8-16	Vape	Roti	Tisu
38	2-8-16	Clear		
39	2-8-16	Cuka	Pena	Silet
40	2-8-16	Rexona		
41	2-8-16	Silet	Cuka	Sendok
42	2-8-16	Sendok	Clear	
43	2-8-16	Rinso	Kopi	Vape
44	2-8-16	Rokok	Piring	
45	2-8-16	Clear	Kopi	
46	2-8-16	Gula	Tepung	Roti
47	2-8-16	Vape	Roti	Tisu
48	2-8-16	Clear		
49	2-8-16	Cuka	Pena	Silet
50	2-8-16	Rexona		

Setelah data konsumen diperoleh, selanjutnya adalah mengelompokkan konsumen berdasarkan jenis barang yang dibeli sebagaimana tabel berikut.

Tabel 3.
Data Pembelian Barang per Konsumen

Pembeli	Barang
1	Gula, Kopi, Garam
2	Beras
3	Rokok, Teh
4	Sampo, Pena, Cuka
5	Garam, Beras
6	Cuka, Buku, Permen
7	Roti, Gula, Pena
8	Teh, Beras
9	Garam, Clear
10	Silet, Cuko, Sendok
11	Sendok, Clear
12	Rinso, Kopi, Vape
13	Rokok, Piring
14	Clear, Kopi
15	Gula, Tepung, Roti
16	Vape, Roti, Tisu
17	Clear
18	Cuka, Pena, Silet
19	Rexona
20	Sampo, Gula, Teh
21	Garam, Clear
22	Silet, Cuka, Sendok
23	Sendok, Clear
24	Rinso, Kopi, Vape
25	Rokok, Piring
26	Clear, Kopi
27	Gula, Tepung, Roti
28	Vape, Roti, Tisu
29	Clear
30	Garam, Clear
31	Silet, Cuka, Sendok
32	Sendok, Clear
33	Rinso, Kopi, Vape
34	Rokok, Piring
35	Clear, Kopi
36	Gula, Tepung, Roti
37	Vape, Roti, Tisu
38	Clear
39	Cuka, Pena, Silet
40	Rexona
41	Silet, Cuka, Sendok
42	Sendok, Clear
43	Rinso, Kopi, Vape
44	Rokok, Piring
45	Clear, Kopi
46	Gula, Tepung, Roti
47	Vape, Roti, Tisu
48	Clear
49	Cuka, Pena, Silet
50	Rexona

Data transaksional di atas lalu direpresentasikan dalam bentuk seperti terlihat pada tabel berikut ini.

Tabel 4.
Representasi Data Peminjaman Buku

Pembeli	Barang
1	Gula
1	Kopi
1	Garam
2	Beras
3	Rokok
3	Teh
4	Sampo
4	Pena
4	Cuka
5	Garam
5	Beras
6	Cuka
6	Buku
6	Permen
7	Roti
7	Gula
7	Pena
8	Teh
8	Beras
9	Garam
9	Clear
10	Silet
10	Cuka
10	Sendok
11	Sendok
11	Clear
12	Rinso
12	Kopi
12	Vape
13	Rokok
13	Piring
14	Clear
14	Kopi
15	Gula
15	Tepung
15	Roti
16	Vape
16	Roti
16	Tisu
17	Clear
18	Cuka
18	Pena
18	Silet
19	Rexona
20	Sampo
20	Gula
20	Teh
21	Garam
21	Clear

Pembeli	Barang
22	Silet
22	Cuka
22	Sendok
23	Sendok
23	Clear
24	Rinso
24	Kopi
24	Vape
25	Rokok
25	Piring
26	Clear
26	Kopi
27	Gula
27	Tepung
27	Roti
28	Vape
28	Roti
28	Tisu
29	Clear
30	Garam
30	Clear
31	Silet
31	Cuka
31	Sendok
32	Sendok
32	Clear
33	Rinso
33	Kopi
33	vape
34	Rokok
34	Piring
35	Clear
35	Kopi
36	Gula
36	tepung
36	Roti
37	vape
37	Roti
37	Tisu
38	Clear
39	cuka
39	Pena
39	Silet
40	Rexona
41	Silet
41	Cuka
41	Sendok
42	Sendok
42	Clear
43	Rinso
43	kopi
43	Vape
44	Rokok
44	Piring
45	Clear

Pembeli	Barang
45	Kopi
46	Gula
46	Tepung
46	Roti
47	Vape
47	Roti
47	Tisu
48	Clear
49	Cuka
49	Pena
49	Silet
50	Rexona

Pada pengujian ini, penulis menetapkan nilai *minimum support* sebesar 5% dan *minimum confidence* sebesar 10%. Sehingga data yang akan diproses oleh *software Orange* hanyalah *rule-rule* yang memenuhi kriteria di atas. Tabel berikut ini menunjukkan calon *2-itemset* dari data pembelian barang yang memenuhi ketentuan *minimum support* dan *minimum confidence* berdasarkan pada Tabel 4.

Tabel 5. Calon 2-itemset

Kombinasi	Jumlah
Rokok, Teh	1
Garam, Beras	1
Teh, Beras	1
Garam, Clear	2
Sendok, Clear	3
Rokok, Piring	3
Clear, Kopi	3

Dari Tabel 5 di atas, dapat disimpulkan jika ditetapkan nilai dari *threshold* di mana $\Phi = 2$, maka:

$F_2 = \{(Rokok, Teh), (Garam, Beras), (Teh, Beras), (Garam, Clear), (Sendok, Clear), (Rokok, Piring), (Clear, Kopi)\}$.

Tabel 6. Calon 3-itemset

Kombinasi	Jumlah
Gula, Kopi, Garam	1
Sampo, Pena, Cuka	1
Silet, Cuka, Sendok	4
Rinso, Kopi, Vape	4
Gula, Tepung, Roti	4
Vape, Roti, Tisu	4
Cuka, Pena, Silet	4
Sampo, Gula, Teh	1

Kombinasi dari *itemset* dalam F_2 dapat kita gabungkan menjadi calon *3-itemset*. *Itemset-itemset* dari F_2 yang dapat digabungkan adalah *itemset-itemset* yang memiliki kesamaan dalam $k-1$ item pertama. Calon *3-itemset* yang dapat dibentuk dari F_2 tampak pada Tabel 5. Dengan demikian $F_3 = \{(Gula, Kopi, Garam), (Sampo, Pena, Cuka), (Silet, Cuka, Sendok), (Rinso, Kopi, Vape), (Gula, Tepung, Roti), (Vape, Roti, Tisu), (Cuka, Pena, Silet), (Sampo, Gula, Teh)\}$.

Pembentukan Aturan Asosiasi

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

$$Confidence = P(B | A) =$$

Jika membeli Gula dan Kopi, maka juga membeli Garam. Untuk mendapatkan nilai *support* dari aturan ini maka masukkan nilai total transaksi pembelian yang mengandung Gula, Kopi dan Garam lalu dibagi dengan total transaksi pembelian. Rujukan dari rumus ini adalah pada tabel 7 dengan *3-itemset* dan tabel 6 dengan *2-itemset*. Sementara untuk mendapatkan nilai *confidence* dari aturan ini adalah dengan menghitung total transaksi mengandung pembelian Gula, Kopi dan Garam lalu dibagi dengan total transaksi pembelian yang mengandung Gula dan Kopi.

Aturan asosiasi yang terbentuk berdasarkan *minimum support* 5% dan *minimum confidence* 10% dapat dilihat pada Tabel 7 berikut.

Tabel 7. Aturan Asosiasi 2-itemset

Aturan	Supp	Conf
Jika membeli Rokok, maka membeli Teh.	2%	50%
Jika membeli Garam, maka membeli Beras.	2%	50%
Jika membeli Teh, maka membeli Beras.	2%	33%

Aturan	Supp	Conf
Jika membeli Garam, maka membeli Clear.	4%	50%
Jika membeli Sendok, maka membeli Clear.	6%	42%
Jika membeli Rokok, maka membeli Piring.	6%	75%
Jika membeli Clear, maka membeli Kopi.	6%	25%

Dari Tabel 7 di atas dapat disimpulkan nilai *support* dan *confidence* dengan acuan *2-itemset* yang memiliki nilai tertinggi adalah kombinasi antara Rokok dan Piring dengan *support* 6% dan *confidence* 75%.

Sementara aturan asosiasi dengan *3-itemset* yang memenuhi kriteria *minimum support* dan *minimum confidence* dapat dilihat pada Tabel 8 berikut.

Tabel 8. Aturan Asosiasi 3-itemset

Aturan	Supp	Conf
Jika membeli Gula dan Kopi, maka membeli Garam.	2%	100%
Jika membeli Sampo dan Pena, maka membeli Cuka.	2%	100%
Jika membeli Silet dan Cuka, maka membeli Sendok.	8%	50%
Jika membeli Rinso dan Kopi, maka membeli Vape.	8%	100%
Jika membeli Gula dan Tepung, maka membeli Roti.	8%	100%
Jika membeli Vape dan Pena, maka membeli Tisu.	8%	100%
Jika membeli Cuka dan Pena, maka membeli Silet.	8%	80%

Dari Tabel 8 di atas diperoleh 5 aturan asosiasi yang memenuhi syarat *minimum support* dan *minimum confidence*. Dapat pula disimpulkan bahwa semua kombinasi pembelian Rinso-Kopi-Vape, Gula-Tepung-Roti dan Vape-Pena-Tisu memiliki tingkat kemungkinan pembelian secara bersamaan sama tinggi.

Hasil Pengujian pada Software

Pengujian terhadap hasil analisa, sangat penting dilakukan untuk menentukan dan memastikan apakah hasil analisa tersebut benar atau tidak. Software

yang penulis gunakan dalam pengujian ini adalah *Orange*.

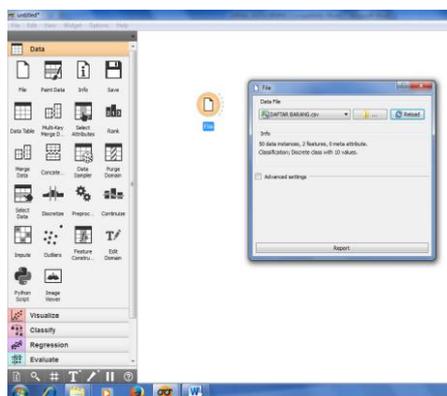
Tahapan dalam pengujian pada *Orange* adalah sebagai berikut.

- a. Buatlah semua daftar transaksi belanja dalam *Office Excel* dan simpan dengan format CSV (*comma delimited*).
- b. Aktifkan *software Orange* sehingga tampil gambar berikut.



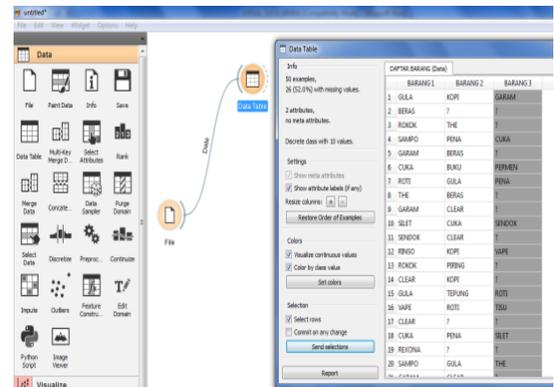
Gambar 2. Area Kerja pada *Orange*

- c. Setelah jendela *Orange* terbuka, klik menu *File*, klik kanan lalu pilih *Open* untuk membuka file daftar transaksi belanja pada *office Excel* seperti gambar berikut.



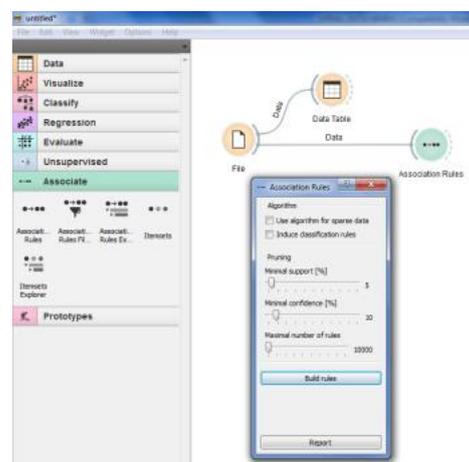
Gambar 3. Proses Pemanggilan Data

- d. Langkah berikutnya adalah melakukan pengecekan apakah data terbaca pada *Orange*, caranya dengan mengklik menu *Data Table* tepat di bawah menu *File*, hubungkan menu *File* dan *Data Table* lalu klik kanan dan pilih *Open* pada *Data Table* untuk melihat tabel belanja seperti gambar berikut.



Gambar 4. Daftar Belanja pada *Orange*

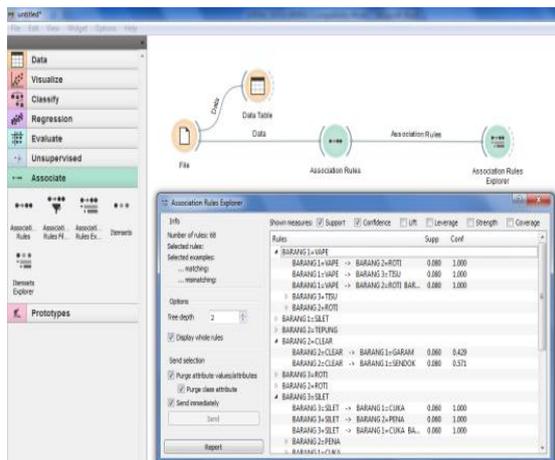
- e. Langkah berikutnya adalah memasukkan menu *Association Rules*, hubungkan menu *File* dengan *Association Rules*, klik kanan *Open*, lalu atur ukuran *Minimum Support 5%* & *Minimum Confidence 10%* seperti gambar berikut.



Gambar 5. Pengaturan *Support* dan *Confidence*

- f. Selanjutnya masukkan menu *Association Rules Explorer*, hubungkan menu *Association Rules* dengan *Association Rules Explorer*, klik kanan lalu pilih *Open* untuk melihat *Rules*

yang memenuhi *Minimum Support* dan *Minimum Confidence* seperti gambar berikut.



Gambar 6. Hasil Akhir Rule Asosiasi.

SIMPULAN

Dari uraian di atas dapat ditarik kesimpulan bahwa Algoritma Apriori yang diproses dengan *software Orange* lebih efektif dan fleksibel digunakan untuk menentukan perilaku konsumen dalam berbelanja. Sistem yang dibangun dapat membantu dalam penempatan barang-barang sesuai dengan kombinasi yang paling sering muncul. Rule-rule yang dihasilkan dapat digunakan sebagai referensi dalam pengadaan barang sesuai kebutuhan konsumen.

UCAPAN TERIMA KASIH

Penelitian ini dapat dilaksanakan dengan bantuan banyak pihak, untuk itu diucapkan terima kasih yang tak terhingga kepada : STIE-GK Muara Bulian-Jambi yang telah memberikan motivasi dan pendanaan dalam menyusun jurnal ini. Tak lupa ucapan terima kasih juga khusus untuk istri dan anakku, semoga jurnal ini bermanfaat.

DAFTAR PUSTAKA

- Emha Taufiq Luthfi. 2009. "Penerapan Data Mining Algoritma Asosiasi Untuk Meningkatkan Penjualan". Seminar, 10(1), 1–21.
- Erwin. 2009. "Analisis Market Basket dengan Algoritma". Jurnal Generic, 4, 26–30.
- Pane, D. K. 2013. "Implementasi Data Mining pada Penjualan Produk Elektronik dengan Algoritma Apriori (Studi Kasus: Kreditplus)". Pelita Informatika Budi Darma, volume : I, 25–29.
- Siregar, S. R. 2014. Implementasi Data Mining Pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori (Studi Kasus: Jumbo Travel Medan), 152–156.
- Studi, P., Informasi, S., Ilmu, F., Universitas, K., & Kuning, L. 2016. Penggunaan Algoritma Apriori Data Mining untuk Mengetahui Tingkat Kesetiaan Konsumen (Brand Loyalty) terhadap Merek Kendaraan Bermotor (Studi Kasus Dealer Honda Rumbai)", x(x), 44–52.
- Tampubolon, K., Saragih, H., Reza, B., Epicentrum .2013. "Implementasi Data Mining Algoritma Apriori pada Sistem Persediaan Alat-Alat Kesehatan", Informasi Dan Teknologi Ilmiah, 93–106.
- Teknik, J., & Fakultas, E. 2008. "Prediksi Mata Kuliah Pilihan dengan Aturan Asosiasi Widodo", Science, 2008, 21–23.
- Wandi, N., Hendrawan, R. A., & Mukhlason, A. 2012. "Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalan Association Rule Menggunakan Algoritma Apriori", Jurnal Teknik ITS, 1, 1–5.

- Widiati, E., & Evita Dewi, K. 2014. "Implementasi Association Rule Terhadap Penyusunan Layout Makanan Dan Penentuan Paket Makanan Hemat Di Rm Roso Echo Dengan Algoritma Apriori", *Jurnal Ilmiah Komputer Dan Informatika (KOMPUTA)*, 96(2), 2089–9033.
- Yanto, R., & Khoiriah, R. 2015. "Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat", 2, 102–113