

ANALISA ALGORITHMMA APRIORI UNTUK MENDAPATKAN POLA PEMINJAMAN BUKU PERPUSTAKAAN SMPN 3 BATANGHARI

Azwar Anas

Jurusan Manajemen, STIE – GK Muara Bulian
Jl. Gajah Mada Muara Bulian 36613
Email: aans_07@yahoo.co.id

ABSTRAK

Data yang dimiliki oleh suatu institusi merupakan salah satu aset dari institusi tersebut. Adanya kegiatan operasional sehari-hari akan semakin memperbanyak jumlah data transaksi. Jumlah data transaksi yang begitu besar justru bisa menjadi masalah bila instansi tersebut tidak bisa memanfaatkannya. Dalam penelitian ini, peneliti akan memanfaatkan algoritma apriori untuk mengelompokkan data peminjaman buku perpustakaan SMPN 3 Batanghari berdasarkan kecenderungannya yang muncul bersamaan dalam suatu kegiatan kunjungan pustaka. Dalam melakukan proses peminjaman buku, tentu saja data mentah akan diolah dengan membaginya kedalam pecahan-pecahan data yang berbeda. Diantara tabel data peminjaman buku yang diolah adalah tabel peminjaman secara umum, tabel kandidat 2-itemset, tabel kandidat 3-itemset, tabel tabular peminjaman, tabel nilai support, tabel nilai confidence dan mendapatkan pola peminjaman buku. Hasil dari penelitian ini berupa gambaran pola peminjaman buku perpustakaan SMPN 3 Batanghari. Dengan mendapatkan pengetahuan dari algoritma ini, dapat dijadikan rujukan bagi pihak sekolah dalam pengadaan dan penempatan buku-buku perpustakaan.

Kata kunci : data mining, algoritma apriori, itemset, support, confidence

ABSTRACT

Data owned by an institution is one of the assets of the institution. Their daily operational activities will further multiply the number of transaction data. The amount of transaction data so large it could be problematic if the agency can not use. In this study, researchers will utilize a priori algorithm to group data library borrowing books SMPN 3 Batanghari based tendency to appear together in a library outreach activities. In the process of borrowing books, of course, raw data will be processed by dividing it into fragments of different data. Among the data table processed borrowing books is a table of borrowing in general, candidate 2-itemset table, table candidate 3-itemsets, borrowing tabular table, the table value of the support, confidence and get the value tables lending patterns. The results of this research is a picture of a library book lending patterns SMPN 3 Batanghari. By gaining knowledge of the algorithm, can be used as a reference for the school in the procurement and placement of library books.

Keywords: data mining, a priori algorithm, itemset, support, confidenc

1. PENDAHULUAN

Data yang dimiliki oleh suatu organisasi merupakan salah satu aset dari organisasi tersebut. Adanya kegiatan operasional sehari-hari akan semakin memperbanyak jumlah data. Jumlah data yang begitu besar justru bisa menjadi masalah bila organisasi tersebut tidak bisa memanfaatkannya. Semakin banyak data, akan semakin diperlukan usaha untuk memilah data mana yang bisa diolah menjadi informasi. Jika data dibiarkan saja, maka data tersebut hanya akan menjadi sampah yang tidak berarti bagi organisasi tersebut. Oleh karena itu, diperlukan sebuah aplikasi yang mampu memilah dan memilih data yang benar sehingga bisa diperoleh informasi yang berguna bagi penggunanya.

Dalam penelitian ini, peneliti akan menjelaskan cara menggunakan aplikasi untuk mengelompokkan data pengunjung perpustakaan SMPN 3 Batanghari berdasarkan kecenderungannya yang muncul bersamaan dalam suatu kegiatan peminjaman buku pustaka menggunakan algoritma apriori.

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^[1].
2. Analisis *Market Basket* dengan Algoritma Apriori dan FP-Growth^[2].
3. Implementasi *Association Rule* terhadap Penyusunan *Layout* Makanan dan Penentuan Paket Makanan Hemat di RM Roso Echo dengan Algoritma Apriori^[3].
4. Implementasi Data Mining pada Penjualan Tiket Pesawat Menggunakan Algoritma Apriori^[4].
5. Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat^[5].

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 *confidence*.

2. TINJAUAN PUSTAKA

2.1. 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^[6]. 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^[7].

2.2. Teknik – teknik *Data Mining*

Data mining dibagi menjadi beberapa teknik berdasarkan tugas yang dapat dilakukan, yaitu^[8] :

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.

2.3. Penerapan Data Mining

Banyak penerapan yang dapat dilakukan oleh *data mining*. Apalagi ditunjang kakayaan 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 Commision 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.

2.4. 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*^[9]. 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^[10].

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 $:2/9 = 0,22$. Sedangkan nilai *confidence* yang diperoleh adalah $c(A, B \rightarrow C)$ adalah $:2/4 = 0,5$. 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.

2.5. 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*.

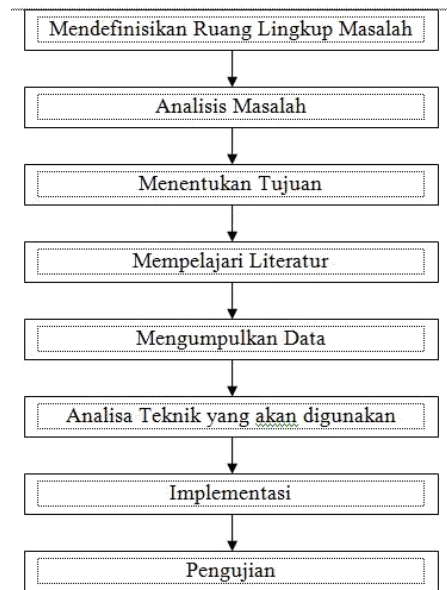
3. METODE PENELITIAN

3.1. Pendahuluan

Dalam melakukan sebuah penelitian, tentu harus berdasarkan metode penelitian yang benar sehingga dapat mempermudah jalannya penelitian tersebut. Metode penelitian merupakan kerangka kerja dalam melakukan penelitian. Dengan mengikuti kerangka kerja tersebut maka penelitian yang dilakukan akan berjalan dengan sistematis dan memberikan hasil yang baik.

3.2. Kerangka Kerja

Pada bagian ini akan diuraikan kerangka kerja penelitian, kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas.



Gambar 1. Kerangka Kerja Penelitian

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

1. Mendefinisikan 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 dengna 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 pengmatan 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 :
 - a. Data pengunjung perpustakaan berdasarkan total jumlah siswa
 - b. Data pengunjung perpustakaan berdasarkan tingkat kunjungan harian
 - c. Data pengunjung perpustakaan dari pihak guru dan staff
 - d. 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.

4. PEMBAHASAN

4.1. Analisa Asosiasi dengan Algoritma Apriori

Dalam penelitian ini, penulis hanya melakukan analisa terhadap 50 transaksi peminjaman buku. Dari 50 transaksi yang telah diamati, hanya 14 judul buku yang telah dipinjam oleh pengunjung perpustakaan.

Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menentukan aturan asosiatif antara aturan kombinasi item. Contoh dari aturan asosiatif dari analisis peminjaman buku di perpustakaan adalah mengetahui besarnya kemungkinan seorang pengunjung perpustakaan untuk meminjam buku Matematika bersamaan dengan buku IPA. Dengan pengetahuan tersebut, pengelola perpustakaan dapat mengatur penempatan bukunya.

Dalam menentukan pola peminjaman buku, akan digunakan algoritma apriori. Sehingga akan dapat dianalisa dengan mudah pola-pola pengunjung perpustakaan di SMPN 3 Batanghari. Data-data pengunjung 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 :

{Matematika, IPA} → {Agama} {*support*= 40%, *confidence* = 50% }

Artinya 50% dari transaksi peminjaman buku yang memuat buku Matematika dan IPA juga memuat buku Agama. Sementara 40% dari seluruh transaksi peminjaman yang ada memuat ketiga item tersebut.

4.2. Menentukan Variabel

Data yang akan diolah dalam menentukan pola peminjaman buku ini adalah data pengunjung. Adapun variabel-variabel data pengunjung perpustakaan SMPN 3 Batanghari adalah :

1. Nama
2. Hari/ tanggal
3. Jenis kelamin
4. Kelas
5. Judul buku 1
6. Judul buku 2
7. Judul buku 3

Dari variabel-variabel di atas, setiap pengunjung akan dibagi kedalam kelompok dengan berdasarkan judul buku yang dipinjam sama, hal ini dilakukan agar proses analisa data dapat dilakukan dengan mudah menggunakan algoritma apriori.

4.3. 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 :

$$\text{Support}(A) = \frac{\text{JumlahtransaksimengandungA}}{\text{Jumlahtransaksi}}$$

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

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

$$Support(A,B) = \frac{\sum \text{Jumlahtransaksimengandung}A \text{ dan } B}{\sum \text{Jumlahtransaksi}}$$

4.4. Melakukan Proses

Berdasarkan variabel-variabel data pengunjung di atas, maka format data dapat dilihat pada tabel berikut ini :

Tabel 3. Data Pengunjung Perpustakaan

No	Nama Pengunjung	Hari/Tanggal	JK	Kelas	Judul Buku 1	Judul Buku 2	Judul Buku 3
1	Finny Wahyuni Putri	Senin/ 7 Nop 2011	P	VII A	TIK	PKn	Penjaskes
2	Peggy Pratama	Senin/ 7 Nop 2011	L	VII A	PKn	MTK	-
3	M. Jhordy	Senin/ 7 Nop 2011	L	VII B	IPA	Penjaskes	-
4	Ade Putri	Senin/ 7 Nop 2011	P	VII H	Agama	PKn	TIK
5	Muhardical Alfayedh	Selasa/ 8 Nop 2011	L	VII H	IPA	TIK	Agama
6	Albujadi Krishma	Senin/ 7 Nop 2011	L	VIII C	TIK	-	-
7	Fikhri Sulthan	Rabu/ 9 Nop 2011	L	VIII D	IPA	MTK	TIK
8	Habib Herman	Senin/ 7 Nop 2011	L	VIII A	TIK	Agama	Penjaskes
9	Yori Karenia	Senin/ 7 Nop 2011	P	VIII A	IPA	TIK	Agama
10	Meilani Putri	Senin/ 7 Nop 2011	P	VIII D	TIK	Agama	-
11	Sandi Kurnia	Senin/ 7 Nop 2011	L	VIII A	Penjaskes	-	-
12	Isadora Azaria	Senin/ 7 Nop 2011	P	IX A	B. Inggris	MTK	IPA
13	Bayu Syam Putra	Senin/ 7 Nop 2011	L	VIII E	Penjaskes	-	-
14	Ardika Ramadhanil	Senin/ 7 Nop 2011	L	VIII E	MTK	B. Indonesia	-
15	Ririn Rilda	Senin/ 7 Nop 2011	L	VIII F	IPS	-	-

16	Delvinda	Senin/ 7 Nop 2011	P	IX D	MTK	B. Inggris	B. Indonesia
17	Deha Dwi Andari	Senin/ 7 Nop 2011	P	IX G	MTK	B. Indonesia	B. Inggris
18	Elvira Andeszuri	Senin/ 7 Nop 2011	P	IX G	IPA	B. Inggris	TIK
19	Suqra Hamda	Senin/ 7 Nop 2011	L	IX B	Penjaskes	TIK	-
20	Aghnia Fila	Senin/ 7 Nop 2011	P	VII C	Agama	TIK	-

Setelah data pengunjung diperoleh, selanjutnya adalah mengelompokkan pengunjung berdasarkan judul buku yang dipinjam. Namun sebelum hal ini dilakukan, agar mempermudah dalam analisa, maka dibuatlah tabel petunjuk buku berdasarkan judul buku masing-masing.

Tabel 4. Transaksi Peminjaman Buku

Pengunjung	Buku yang dipinjam
1	TIK, PKn, Penjaskes
2	PKn, MTK
3	IPA, Penjaskes
4	Agama, PKn, TIK
5	IPA, TIK, Agama
6	TIK
7	IPA, MTK, TIK
8	TIK, Agama, Penjaskes
9	IPA, TIK, Agama
10	TIK, Agama
11	Penjaskes
12	B. Inggris, MTK, IPA
13	Penjaskes
14	MTK, B. Indonesia
15	IPS
16	MTK, B. Inggris, B. Indonesia
17	MTK, B. Indonesia, B. Inggris
18	IPA, B. Inggris, TIK
19	Penjaskes, TIK
20	Agama, TIK

Data transaksional di atas lalu direpresentasikan dalam bentuk seperti terlihat pada tabel 5 berikut ini :

Tabel 5. Representasi Data Peminjaman Buku

Pengunjung	Buku yang dipinjam
1	TIK
1	PKn
1	Penjaskes
2	PKn
2	MTK
3	IPA
3	Penjaskes
4	Agama
4	PKn

4	TIK
5	IPA
5	TIK
5	Agama
6	TIK
7	IPA
7	TIK
7	MTK
8	TIK
8	Agama
8	Penjaskes
9	IPA
9	TIK
9	Agama
10	TIK
10	Agama
11	Penjaskes
12	B. Inggris
12	MTK
12	IPA
13	Penjaskes
14	MTK
15	B. Indonesia
16	MTK
16	B. Inggris
16	B. Indonesia
17	MTK
17	B. Indonesia
17	B. Inggris
18	IPA
18	B. Inggris
18	TIK
19	Penjaskes
19	TIK
20	Agama
20	TIK

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 6 berikut ini menunjukkan calon *2-itemset* dari data peminjaman yang memenuhi ketentuan *minimum support* dan *minimum confidence* berdasarkan pada tabel 5.

Tabel 6. Calon *2-itemset*

Kombinasi	Jumlah
PKn, MTK	1
IPA, Penjaskes	1
TIK, Agama	2
MTK, B. Indonesia	1
Penjaskes, TIK	1

Dari tabel 6 di atas, dapat disimpulkan jika ditetapkan nilai dari *threshold* di mana $\Phi = 2$, maka : $F_2 = \{(PKn, MTK), (IPA, Penjaskes), (TIK, Agama), (MTK, B. Indonesia), (Penjaskes, TIK)\}$.

Tabel 7. Calon 3-itemset

Kombinasi	Jumlah
Agama, TIK, Penjaskes	3
TIK, IPA, Agama	3
TIK, Agama, Penjaskes	3
IPA, TIK, Agama	3

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 7. Dengan demikian $F_3 = \{(Agama, TIK, Penjaskes), (TIK, IPA, Agama), (TIK, IPA, Penjaskes), (IPA, TIK, Agama)\}$.

4.5. 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 meminjam buku Agama dan TIK, maka akan meminjam buku Penjaskes. Untuk mendapatkan nilai *confidence* dari aturan ini maka masukkan nilai total transaksi peminjaman yang mengandung buku Agama dan TIK lalu dibagi dengan total transaksi peminjaman yang mengandung buku Agama dan TIK. Rujukan dari rumus ini adalah pada tabel 7 dengan 3-itemset dan tabel 6 dengan 2-itemset.

Sementara untuk mendapatkan nilai *support* dari jika meminjam buku MTK, maka akan meminjam buku IPA merupakan hasil dari perhitungan total transaksi mengandung peminjaman buku MTK dan IPA lalu dibagi dengan total transaksi peminjaman.

Aturan asosiasi yang terbentuk berdasarkan *minimum support* 5% dan *minimum confidence* 10% dapat dilihat pada tabel 8 berikut.

Tabel 8. Aturan Asosiasi 2-itemset

Aturan	Support	Confidence
Jika meminjam MTK, maka akan meminjam IPA	7%	50%
Jika meminjam Agama, maka akan meminjam Penjaskes	8%	100%
Jika meminjam TIK, maka akan meminjam Agama	11%	75%
Jika meminjam TIK, maka akan meminjam IPA	11%	37%
Jika meminjam TIK, maka akan meminjam Penjaskes	11%	37%
Jika meminjam Penjaskes, maka akan meminjam IPS	7%	50%
Jika meminjam IPA, maka akan meminjam Agama	8%	33%
Jika meminjam IPA, maka akan meminjam TIK	11%	55%
Jika meminjam B. Inggris, maka akan meminjam TIK	8%	50%

Dari tabel 8 di atas dapat disimpulkan nilai *support* dan *confidence* dengan acuan 2-itemset yang memiliki nilai tertinggi adalah kombinasi antara buku Agama dan buku Penjaskes dengan total *support X confidence* adalah 80%.

Sementara aturan asosiasi dengan 3-itemset yang memenuhi kriteria *minimum support* dan *minimum confidence* dapat dilihat pada tabel 9 berikut.

Tabel 9. Aturan Asosiasi 3-itemset

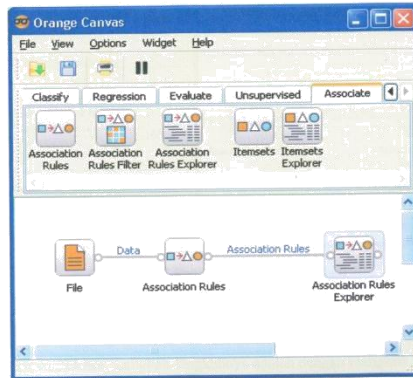
Aturan	Support	Confidence
Jika meminjam Agama dan TIK, maka akan meminjam Penjaskes	6%	100%
Jika meminjam TIK dan IPA, maka akan meminjam Agama	6%	100%
Jika meminjam TIK dan Agama, maka akan meminjam Penjaskes	6%	100%
Jika meminjam IPA dan TIK, maka akan meminjam Agama	6%	100%

Dari tabel 9 di atas diperoleh 4 aturan asosiasi yang memenuhi syarat *minimum support* dan *minimum confidence*. Dapat pula disimpulkan bahwa semua kombinasi peminjaman buku Agama, TIK dan Penjaskes

di perpustakaan SMPN 3 Batanghari memiliki tingkat kemungkinan peminjaman secara bersamaan sama tinggi.

4.6. Hasil Pengujian pada Software

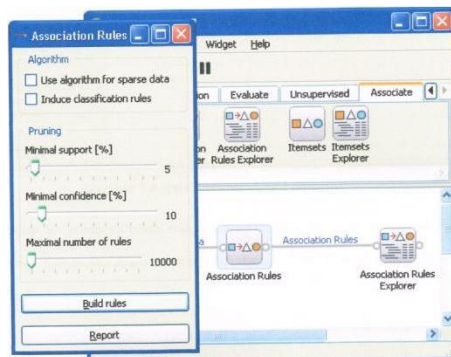
Pengujian terhadap hasil analisa, sangat penting dilakukan untuk menentukan dan memastikan apakah hasil analisa tersebut benar atau tidak. Software yang penulis gunakan dalam pengujian ini adalah Orange.



Gambar 2. Proses Rule Asosiasi

Gambar di atas memperlihatkan hubungan antara file yang telah diakses dan rule asosiasi yang akan dicari. Data yang telah diakses akan diproses oleh menu *Association Rules* dan akan ditampilkan melalui menu *Association Rules Explorer*.

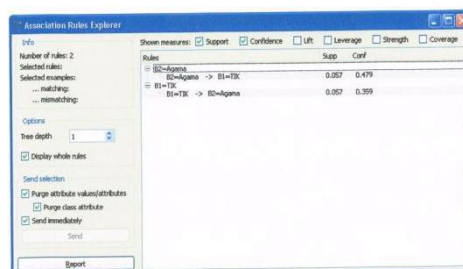
Setelah dihubungkan antara file dan rule asosiasi. Lakukan pengaturan nilai *minimum support* dan *minimum confidence* dalam penelitian ini penulis menetapkan 5 % untuk *minimum support* dan 10 % *minimum confidence*.



Gambar 3. Pengaturan Nilai Minimum Support dan Confidence

Gambar 3 memperlihatkan halaman pengaturan *minimum support* dan *minimum confidence*. Langkah-langkahnya adalah dengan mengklik kanan menu *Association Rules* Open sehingga tampil sebagaimana gambar di atas.

Untuk melihat rule-rule yang terbentuk, klik kanan *Association Rules Explorer* dan pilih *Open*, maka rule-rule yang memenuhi kriteria *minim support* dan *minimum confidence* yang telah ditetapkan akan ditampilkan sebagaimana gambar 4 berikut.



Gambar 4. Rule Asosiasi yang Terbentuk

Gambar 4 memperlihatkan *rule-rule* asosiasi yang terbentuk setelah *software Orange* memproses database yang telah diakses. Rule asosiasi yang terbentuk ialah :

- a. Jika meminjam buku Agama, maka akan meminjam buku TIK dengan *support* 5 % dan *confidence* 50 %
- b. Jika meminjam buku TIK, maka akan meminjam buku Agama dengan *support* 5 % dan *confidence* 40 %.

Dari rule di atas, ternyata buku Agama dan TIK menjadi buku yang paling sering dipinjam secara bersamaan oleh pengunjung perpustakaan SMPN 3 Batanghari.

5. PENUTUP

5.1. Kesimpulan

Dari uraian pada bab-bab yang sudah dibahas sebelumnya dapat ditarik kesimpulan :

1. Algoritma Apriori yang diproses dengan *software Orange* lebih efektif dan fleksibel digunakan untuk menentukan pola pengunjung perpustakaan
2. Pemilihan variabel yang digunakan sangat mempengaruhi *rule* atau *knowledge* yang dihasilkan
3. Sistem yang dibangun dapat membantu dalam menentukan pola pengunjung perpustakaan yang umumnya masih dilakukan secara manual
4. Algoritma Apriori dengan *rule* asosiasinya dapat memberikan informasi efektif untuk menggambarkan proses yang terkait dengan pola pengunjung perpustakaan
5. *Rule-rule* yang dihasilkan dapat digunakan sebagai referensi dalam pengadaan buku-buku koleksi di perpustakaan SMPN 3 Batanghari.

5.2. Saran

Pengembangan sistem ini adalah dari sistem manual atau berdasarkan basis pengetahuan ke sistem komputerisasi yang dibangun yang dibangun atas dasar pertimbangan permasalahan-permasalahan yang timbul dari sistem yang lama. Berikut ini adalah saran yang dapat diperlihatkan untuk masa yang akan datang.

1. Pada penelitian ini, penulis mencoba salah satu teknik yang digunakan dalam menentukan pola pengunjung perpustakaan yaitu Algoritma Apriori. Untuk mendapatkan hasil pola yang baik dapat digunakan beberapa atau penggabungan beberapa teknik prediksi. Oleh sebab itu perlunya dilakukan penelitian lanjut untuk membandingkan hasil pola dan menentukan teori mana yang menghasilkan prediksi yang terbaik.
2. Untuk mendapatkan hasil pola pengunjung yang baik diperlukan sumber data yang lengkap. Oleh sebab itu diharapkan pada proses penyimpanan data perpustakaan sudah dilakukan secara elektronik dengan demikian akan memudahkan dalam pencarian data yang dibutuhkan untuk melakukan proses penemuan *knowledge* seperti untuk menentukan pola pengunjung perpustakaan.
3. Walaupun aplikasi ini sangat mudah untuk dijalankan, namun perlu diasosiasikan cara-cara pengoperasiannya.

DAFTAR PUSTAKA

- [1]. Emha T.L. (2009). Penerapan Data Mining Algoritma Asosiasi untuk Meningkatkan Penjualan. Jurnal DASI.
- [2]. Erwin. (2009). Analisis *MarketBasket* dengan Algoritma *Apriori* dan *FP-Growth*. Jurnal Generic.
- [3]. Elsa.W & Kania.E.D. (2014).Implementasi *Association Rule* terhadap Penyusunan *Layout* Makanan dan Penentuan Paket Makanan Hemat di RM Roso Echo dengan Algoritma Apriori. Jurnal KOMPUTA.
- [4]. Sri R.S. (2014). Implementasi Data Mining pada Penjualan Tiket Pesawat menggunakan Algoritma Apriori. Jurnal Pelita Informatika Budi Darma. VII
- [5]. Robi.Y & Riri.K. (2015).Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat. Citec Journal. 2
- [6]. Wirdah. C. (2016).Penggunaan Algoritma Apriori Data Mining untuk Mengetahui Tingkat Kesetiaan Konsumen (*Brand Loyalty*) terhadap Merek Kendaraan Bermotor (Studi Kasus Dealer Honda Rumbai).Jurnal Teknologi Informasi dan Komunikasi Digital Zone.7
- [7]. Dewi.K.P. (2013). Implementasi Data Mining pada Penjualan Produk Elektronik dengan Algoritma Apriori (Studi Kasus: Kredit Plus). Jurnal Pelita Informatika Budi Darma. IV

- [8]. Kennedy. T, Hoga.S & Bobby.R. (2013).Implementasi Data Mining Algoritma Apriori pada Sistem Persediaan Alat-alat Kesehatan.Jurnal Informasi dan Teknologi Ilmiah.1
- [9]. Nugroho W, Rully. A.H & Ahmad.M. (2012).Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalian *Association Rule* menggunakan Algoritma Apriori (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur).Jurnal Teknik ITS.1
- [10]. Widodo. (2008, Mei).*Prediksi Mata Kuliah Pilihan dengan Aturan Asosiasi*.Makalah disajikan dalam Konferensi dan Temu Nasional Teknologi Informasi dan Komunikasi untuk Indonesia. Jakarta: Universitas Negeri Jakarta.