

IMPLEMENTASI DATA MINING ALGORITMA APRIORI PADA SISTEM PERSEDIAAN ALAT-ALAT KESEHATAN

Kennedi Tampubolon¹⁾, Hoga Saragih²⁾, Bobby Reza³⁾

¹⁾ Dosen Tetap Prodi D3-Manajemen Informatika STMIK Budi Darma Medan. Jl. Sisingamangaraja No. 338 Sp. Limun Medan. ²⁾ Fakultas Teknik dan Ilmu Komputer – Universitas Bakrie Kampus Kuningan, Kawasan Epicentrum, Jl. HR. Rasuna Said Kav. C-22 Jakarta 12920. ³⁾ Program Pascasarjana STMIK ERESHA, Jl. H. Samali No. 51 Kalibata, 12740
Email : kenned.tampubolon@gmail.com, hogasaragih@gmail.com, bobby.reza@eresha.ac.id

ABSTRAK

Menambang data atau upaya untuk menggali informasi dan pengetahuan yang berharga pada database yang sangat besar disebut data mining atau Knowledge Discovery in Database disingkat KDD. Salah satu algoritma yang paling populer pada teknik data mining adalah algoritma Apriori. Sedangkan dalam penemuan pola kombinasi hubungan antar *item-sets* digunakan *Association Rules* (Aturan Asosiasi).

Data Mining telah diimplementasikan ke berbagai bidang, diantaranya bidang bisnis atau perdagangan, bidang pendidikan, dan telekomunikasi. Dibidang bisnis misalnya hasil implementasi data mining menggunakan algoritma Apriori dapat membantu para pebisnis dalam kebijakan pengambilan keputusan terhadap apa yang berhubungan dengan persediaan barang. Misalnya pentingnya sistem persediaan barang di suatu Apotek dan jenis barang apa yang menjadi prioritas utama yang harus di stok untuk mengantisipasi kekosongan barang. Karena minimnya stok barang dapat berpengaruh pada pelayanan konsumen dan pendapatan Apotek. Oleh sebab itu ketersediaan berbagai jenis alat-alat kesehatan di Apotek sebagai salah satu supplier alat-alat kesehatan, mutlak untuk mendukung kelancaran penyalurannya kepada konsumen, sehingga aktivitas pelayanan konsumen berjalan dengan baik.

Seiring dengan masalah diatas, data mining mampu menciptakan lingkungan bisnis yang inteligen, untuk menghadapi semakin tingginya tingkat persaingan bisnis Apotek dimasa yang akan datang.

Kata Kunci : Data Mining, Aturan Asosiasi, Algoritma Apriori, Alat-alat Kesehatan.

ABSTRACT

Extracting data or an effort to retrieve valuable knowledge and information in a large database is called data mining or Knowledge Discovery in Database or usually shortened as KDD. One of the most popular algorithm in data mining technic is Apriori Algorithm, while the discovery of "relational combination pattern among itemset used Association Rules".

Data mining has been implemented into the various fields like : business or trade, education and telecommunication. In bussiness for instance, the implementation result of data mining use 'algorithm Apriori which can give a hand to help the Businessmen make decision on supplies. For example, the necessity of supplies system in a drugstore as one of the medical stuff supplier, and to determine which product as the priority should be supplied to anticipate out of stock of supplies availability in the store, as the results will also affect to the consumer service and daily income. Medical tools are essential unit should be supplied and being an essential factor which will impact to the consumer trust to a hospital or another medical service. That is why the availability of medical tools in drugstores is completely needed to support the succes of distribution to the consumers, so the activity of medical service to consumers run thoroughly.

In this case, data mining is seen as able to build intelligent business environment as solution for competing increased competition among the drugstores in future.

Keywords : Data Mining, Association Rules, Apriori Algorithm, Medical tools.

1. PENDAHULUAN

Data Mining diartikan sebagai menambang data atau upaya untuk menggali informasi yang berharga dan berguna pada database yang sangat besar (Agrawal, R., Srikant, 1996). Hal terpenting dalam teknik data mining adalah aturan untuk menemukan pola frekuensi tinggi antar himpunan *itemset* yang disebut fungsi *Association Rules* (Aturan Asosiasi). Beberapa algoritma yang termasuk dalam Aturan Asosiasi adalah seperti *AIS Algorithm*, *Apriori*

Algorithm, *DHP Algorithm*, dan *Partition Algorithm* (Shuruti Aggarwal, 2013). Namun diantara algoritma-algoritma tersebut, penulis memilih Algoritma Apriori dalam aplikasinya pada penelitian.

Belakangan ini data mining telah diimplementasikan ke berbagai bidang, diantaranya dalam bidang bisnis atau perdagangan, bidang pendidikan, dan telekomunikasi. Dibidang bisnis misalnya hasil implementasi data mining algoritma Apriori dapat membantu para pebisnis dalam

pengambilan keputusan terhadap apa yang berhubungan dengan persediaan barang (Finn Lee S & Juan Santana. 2010). Misalnya pentingnya sistem persediaan barang di suatu Apotek dan jenis barang apa yang menjadi prioritas utama yang harus distok untuk mengantisipasi kekosongan barang. Karena minimnya stok barang dapat berpengaruh pada pelayanan konsumen dan pendapatan Apotek.

Adanya aktivitas transaksi penjualan sehari-hari pada Apotek Kelambir – 2 Medan, akan menghasilkan tumpukan data yang semakin lama semakin besar, sehingga dapat menimbulkan masalah baru. Jika hal ini dibiarkan, maka data-data transaksi tersebut akan menjadi tumpukan sampah yang merugikan karena membutuhkan media penyimpanan/database yang semakin besar. Seiring dengan kemajuan teknologi yang semakin pesat maka berkembang pula kemampuan dalam mengumpulkan, menganalisa dan mengolah data pada database. Sehingga diperlukan suatu aplikasi untuk memilah dan memilih data penting dari database.

Dalam persaingan di dunia bisnis, khususnya industri Apotek, menuntut para pengembang untuk menemukan suatu strategi jitu yang dapat meningkatkan penjualan barang. Salah satu cara mengatasinya adalah dengan tetap tersediaannya berbagai jenis alat-alat kesehatan secara kontinu digudang Apotik. Untuk mengetahui alat-alat kesehatan apa saja yang dibeli oleh para konsumen, dilakukan teknik analisis keranjang pasar yaitu analisis dari kebiasaan membeli konsumen. Penerapan Algoritma Apriori, membantu dalam membentuk kandidat kombinasi *item* yang mungkin, kemudian dilakukan pengujian apakah kombinasi tersebut memenuhi parameter *support* dan *confidence* minimum yang merupakan nilai ambang yang diberikan oleh pengguna.

Aktivitas dan transaksi di suatu organisasi pada umumnya sama, yang berbeda adalah elemen – elemen penyusunnya, misalnya organisasi pemerintahan berbeda dengan organisasi bisnis, demikian pula dengan organisasi kependidikan. Karakteristik organisasi yang berbeda demikian akan menghasilkan pemodelan bisnis yang bervariasi, dan perbedaan ini juga terlihat dalam data dan aliran datanya.

Di Apotek Kelambir – 2 Medan misalnya aktivitas transaksi dan pelayanan terhadap konsumen setiap harinya semakin lama semakin tinggi, sehingga tanpa disadari hal ini dapat menimbulkan tumpukan data yang semakin besar.

Dalam menjalankan aktivitasnya, Apotek Kelambir – 2 Medan sudah menggunakan jasa teknologi komputer sebagai alat dalam pengimputan data, pengolahan serta pencetakan/print out hasil pengolahan data berupa informasi yang di inginkan. Namun dalam pengolahan data masih menggunakan aplikasi – aplikasi yang sangat sederhana, dan cara – cara manual juga masih dilakukan terutama dalam pengecekan barang masuk dan keluar dan dalam pengarsipan data. Walaupun

hingga saat ini aktivitas pelayanan dan transaksi di Apotek Kelambir – 2 Medan belum mengalami kendala yang berarti, tentu keadaan ini suatu saat menjadi faktor penghambat dalam meningkatkan pelayanan seiring semakin banyaknya transaksi dan jenis item dan itemset transaksi yang terjadi dan tersimpan dalam kurun waktu tertentu, sehingga menyulitkan pihak apotek dalam menganalisa jenis item dan itemset barang mana yang paling diminati atau tidak diminati konsumen.

1.1. Identifikasi Masalah

Dari uraian permasalahan diatas berikut ini masalah yang dapat diidentifikasi oleh peneliti, antara lain :

1. Semakin tingginya aktivitas transaksi sehari-hari akan menimbulkan penumpukan data dalam database.
2. Dalam melakukan administrasi, aktivitas transaksi dan pelayanan konsumen, perusahaan sudah menggunakan teknologi komputer namun dalam pengolahan datanya masih sangat sederhana.
3. Belum adanya penelitian yang bertujuan untuk mengetahui informasi penting dalam meningkatkan penjualan dan pelayanan pada konsumen.

1.2. Ruang Lingkup Masalah

Penelitian ini membahas tentang Implementasi data mining menggunakan algoritma Apriori dengan studi kasus pada Apotik Kelambir – 2 Medan dari bulan Oktober – November 2013 dan yang menjadi populasi penelitian ini adalah data- data transaksi penjualan alat-alat kesehatan. Sedangkan dalam pengujian peneliti menggunakan program aplikasi Microsoft Excel 2007 dan Tanagra 1.4.48

1.3. Rumusan Masalah

Berdasarkan analisa penulis dari uraian diatas, maka yang menjadi rumusan masalah dari penelitian ini adalah sebagai berikut :

1. Bagaimana mengimplementasikan Data Mining pada database transaksi penjualan *item* alat - alat kesehatan.
2. Bagaimana penerapan Algoritma Apriori dalam penentuan frekuensi tinggi *item-set* untuk memprediksi persediaan barang di waktu yang akan datang.

1.4. Tujuan dan Manfaat Penelitian

Adapun tujuan dan manfaat dari penelitian ini adalah sebagai berikut:

- a. Mengimplementasikan Data Mining pada database transaksi penjualan *item* alat-alat kesehatan.
- b. Menerapkan Algoritma Apriori untuk menentukan frekuensi tinggi *itemset* untuk memprediksi persediaan barang di waktu yang akan datang.

2. LANDASAN TEORI

2.1 Tinjauan Pustaka

Penelitian yang berhubungan dengan penelitian ini adalah :

- a. Penelitian yang dilakukan oleh Ema Utami (Ema Utami, Jazi Eko Istiyanto, Suwanto Raharjo, 2007) yang menjelaskan tentang bagaimana metodologi penelitian dalam melakukan riset pada ilmu komputer.
- b. Penelitian yang dilakukan oleh Rakesh Agrawal dkk (Agrawal, R., Manilla, H., Srikant, R., Toivonen, H., & Verkamo, I., 1996) yang membahas tentang cara kerja aturan asosiasi menggunakan algoritma Apriori untuk menemukan itemset terbesar dengan menggunakan data sintetik.
- c. Penelitian yang dilakukan oleh Usama Fayyad, Gregory Piatetsky – Shapiro, dan Padhraic Smyth (Fayyad, U.M., Piatetsky – Shapiro, G., & Smyth, P., 1996) tentang Knowledge Discovery in Database (KDD) yang disebut data mining, bagaimana proses KDD, metode data mining, apa saja komponen – komponen data mining, algoritma, dan penerapannya dalam industri dan ilmu analisa data.
- d. Penelitian yang dilakukan oleh Yeong – Chyi Lee dkk (Yeong – Chyi Lee, Tzung – Pei Hong, & Wen – Yang Lin, 2005) yang membahas teknik data mining menggunakan maximum constraint.
- e. Penelitian yang dilakukan oleh N. Badal, dan Shruti Tripathi (N. Badal, & Shruti Tripathi, 2010) tentang proses data mining menggunakan algoritma VS_Apriori sebagai pengembangan algoritma Apriori.
- f. Penelitian yang dilakukan oleh Pratima Gautam (Pratima Gautam dan K. R., Pardasani, 2011) membahas tentang usulan dalam proses data mining menggunakan aturan asosiasi bermacam level dalam waktu yang bersamaan.
- g. Penelitian yang dilakukan oleh Sanjeev Rao (Sanjeev Rao, dan Priyanka Gupta, 2012) tentang pengembangan algoritma Apriori pada aturan asosiasi yaitu FP-Growth Algorithm karena membutuhkan iterasi dan waktu yang lebih singkat.
- h. Penelitian yang dilakukan oleh Bala Sundar V., T Devi dan N Saravanan (Bala Sundar V, T Devi, & N Saravanan, 2012) tentang bagaimana menerapkan algoritma K-Means pada pengklusteran data mining analisa tentang Predicting Heart Disease dan mendesain metode apa yang lebih efektif.
- i. Penelitian yang dilakukan oleh Othman Yahya (Othman Yahya, Osman Hegazy, dan Ehab Ezat, 2012) yaitu implementasi dari algoritma Apriori berbasis HADOOP-MAPREDUCE MODEL.
- j. Penelitian yang dilakukan oleh Jogi Suresh, dan T. Ramanjaneyulu (Jogi Suresh, T. Ramanjaneyulu, 2013) tentang pencarian pola

frekuensi itemset pada data mining menggunakan algoritma Apriori.

- k. Penelitian yang dilakukan oleh Shruti Aggarwal (Shruti Aggarwal, dan Ranveer Kaur, 2013) tentang studi perbandingan antara algoritma Apriori dengan algoritma versi lain seperti AIS Algorithm, DHP, dan Partition Algorithm.
- l. Penelitian yang dilakukan oleh Prof. Paresh Tanna (Prof. Paresh Tanna, dan Dr. Yogesh Ghodasara, 2013) tentang dasar-dasar implementasi pola frekuensi tinggi pada algoritma data mining.
- m. Penelitian yang dilakukan oleh Abdullah Saad Almalaise Alghamdi (Abdullah Saad Almalaise Alghamdi, 2011) membahas bagaimana implementasi algoritma FP Growth pada data mining data medis.
- n. Penelitian yang dilakukan oleh Rajneech Kumar Singh et al (Rajneech Kumar Singh, Manoj Kumar Pandey, dan Jawed Ahmed, 2013) yang menjelaskan pemakaian pendekatan *Disconnected Approach* pada algoritma Apriori.

Dari cerita jurnal tersebut sebagai studi literatur dan penelusuran ilmiah diatas, yang dikerjakan sebagai kontribusi Tesis ini adalah :

Dalam penelitian ini, peneliti telah membaca banyak hal tentang data mining, teknik data mining dengan *association rules*, dan algoritma – algoritma yang terkait dengan *association rules* serta implementasinya dalam bisnis. Peneliti ingin mengimplementasikan teknik data mining pada database barang terjual alat – alat kesehatan dengan aturan asosiasi menggunakan algoritma Apriori untuk mengetahui pola kombinasi *item* dan *itemset* frekuensi tinggi dari alat – alat kesehatan sehingga dapat dijadikan sebagai faktor pengambilan keputusan dalam memprediksi persediaan barang. Untuk itu peneliti memilih Apotek Kelambir – 2 Medan sebagai tempat riset dan pengambilan data sampel untuk diolah dan dijadikan objek dalam pengujian.

2.2 Implementasi

Implementasi adalah yang bermuara pada aktivitas, aksi, tindakan, atau adanya mekanisme suatu sistem. Implementasi bukan sekedar aktivitas, tetapi suatu kegiatan yang terencana dan untuk mencapai tujuan kegiatan. Implementasi adalah perluasan aktivitas yang saling menyesuaikan proses interaksi antara tujuan dan tindakan untuk mencapai serta memerlukan jaringan-pelaksanaan, birokrasi-yang-efektif.

(http://elib.unikom.ac.id/files/disk1/487/jbptunikom-pg-gdl-derrisepti-24335-2-babii_d-x.pdf)

2.3. Data Mining

Selain dari pengertian yang telah dijelaskan pada latar belakang, Data Mining disebut juga *Knowledge Discovery in Database (KDD)* didefinisikan sebagai ekstraksi informasi potensial,

implisit dan tidak dikenal dari sekumpulan data. Proses *Knowledge Discovery in Database* melibatkan hasil proses data mining (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami (Sri Andayani, 2010).

Ada beberapa macam pendekatan yang berbeda yang diklasifikasikan sebagai teknik pencarian informasi/pengetahuan dalam KDD. Ada pendekatan kuantitatif, seperti pendekatan probabilistik seperti logika induktif, pencarian pola, dan analisis pohon keputusan. Pendekatan yang lain meliputi deviasi, analisis kecenderungan, algoritma genetik, jaringan saraf tiruan, dan pendekatan campuran dua atau lebih dari beberapa pendekatan yang ada.

Pada dasarnya ada enam elemen yang paling esensial dalam teknik pencarian informasi/pengetahuan dalam KDD yaitu:

- 1 Mengerjakan sejumlah besar data.
- 2 Diperlukan efisiensi berkaitan dengan volume data.
- 3 Mengutamakan ketetapan/keakuratan.
- 4 Membutuhkan pemakaian bahasa tingkat tinggi.
- 5 Menggunakan beberapa bentuk dari pembelajaran otomatis.
- 6 Menghasilkan hasil yang menarik.

Perusahaan memerlukan kecerdasan bisnis untuk mengembangkan proses bisnis, memonitor waktu, biaya kualitas, dan pengendalian. Gartner Group mendefinisikan kecerdasan bisnis atau *business intelligence* (BI) sebagai berikut.

"Business intelligence is the process of transforming data into information and through discovery transforming that information into knowledge."

Dalam definisi ini disampaikan bahwa kecerdasan bisnis merupakan proses pengubahan data menjadi informasi. Dari kumpulan informasi yang ada akan diambil polanya menjadi pengetahuan. Tujuan kecerdasan bisnis adalah untuk mengubah data yang sangat banyak menjadi nilai bisnis melalui laporan analistik. (Kusrini dan Emha Taufiq Luthfi, 2009).

2.4. Defenisi Data Mining

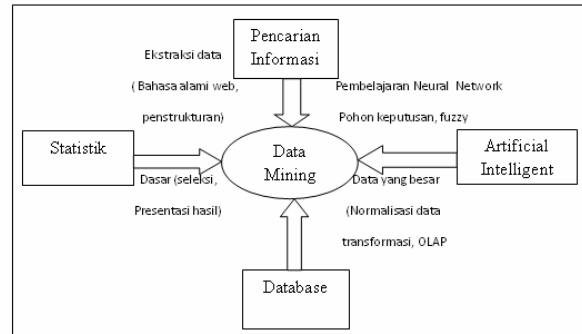
Data mining merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar/Data Warehouse (Turban, dkk. 2005)

Keluaran dari data mining bisa dipakai untuk memperbaiki pengambilan keputusan dimasa depan (Budi Santosa, 2007).

Berdasarkan defenisi-defenisi yang telah disampaikan, hal penting yang terkait dengan Data Mining adalah:

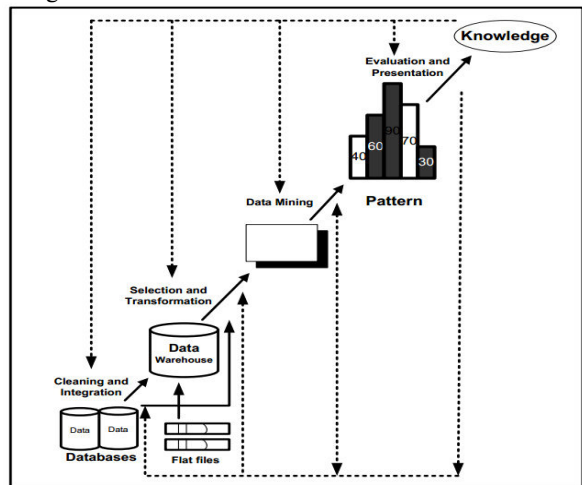
1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada.

2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang akan mungkin memberikan indikasi yang bermanfaat.



Gambar 1 : Bidang Ilmu Data Mining
Sumber: (Kusrini, Emha Taufiq Luthfi, 2009)

Proses KDD secara garis besar dapat dijelaskan sebagai berikut :



Gambar 2 : Proses Data Mining
(Sumber : Beta Noranita dan Nurdin Bahtiar, 2010)

2.5. Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Kusrini dan Emha Taufiq Luthfi, 2009):

1. Deskripsi
Terkadang peneliti dan analis secara sederhana ingin mencoba mencari data untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menentukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelesan untuk suatu pola atau kecenderungan.
2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

3. Prediksi.

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang.

Contoh prediksi bisnis dan penelitian adalah:

- Prediksi harga beras dalam tiga bulan yang akan datang.
- Prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikkan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah:

- Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau tidak.
- Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa.

5. Pengklusteran (Clustering)

Pengklusteran merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data

menjadi kelompok-kelompok yang memiliki kemiripan (*homogeny*), yang mana kemiripan dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

Contoh pengklusteran dalam bisnis dan penelitian adalah:

- Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari satu suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku financial dalam baik dan mencurigakan.
- Melakukan pengklusteran terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen dalam jumlah besar.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan *attribut* yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

Contoh asosiasi dalam bisnis dan penelitian adalah:

- Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran *upgrade* layanan yang diberikan.
- Menentukan barang dalam supermarket yang dibeli secara bersamaan dan yang tidak pernah dibeli secara bersamaan.

2.6. Fungsi dan Tugas Data Mining

Data mining menganalisis data menggunakan tool untuk menemukan pola dan aturan dalam himpunan data. Perangkat lunak bertugas untuk menemukan pola dengan mengidentifikasi aturan dan fitur pada data. Tool data mining diharapkan mampu mengenal pola ini dalam data dengan input minimal dari user (Dana Sulistiyo Kusumo et al, 2003).

2.6. 1. Association 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 (amirudin et al, 2007)

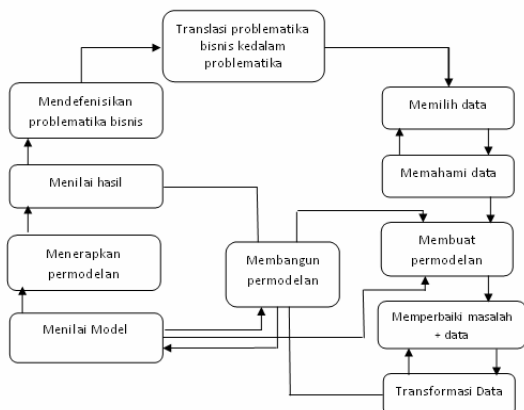
Aturan asosiasi akan menggunakan data latihan, sesuai dengan pengertian data mining, untuk menghasilkan pengetahuan. Pengetahuan untuk mengetahui *item-item* belanja yang sering dibeli secara bersamaan dalam suatu waktu. Aturan asosiasi yang berbentuk "*if...then...*" atau "jika...maka..." merupakan pengetahuan yang dihasilkan dari fungsi

Aturan Asosiasi (Seni Susanto dan Dedy Suryadi, 2010)

2.6.2. Langkah-langkah Data Mining

Ada empat tahap yang dilalui dalam Data Mining antara lain (Feen Lee & Juan Santana, 2010:37-40) :

1. Tahap pertama: Precise statement of the problem (mendefinisikan permasalahan yang ingin diketahui). Misalnya ingin mengetahui apakah seorang customer berpotensi memiliki kredit macet, atau mengidentifikasi seorang customer apakah akan pindah ke kompetitor bisnis kita, dan lain sebagainya. Setelah menemukan pertanyaan bisnis yang perlu dijawab oleh data mining, selanjutnya tentukan tipe tugas untuk menjawab pertanyaan bisnis tersebut. Tugas dasar yang menjadi dasar algoritma data mining adalah klasifikasi, regresi, segmentasi, asosiasi dan sequence analisis.
2. Tahap kedua: Initial Exploration (Mempersiapkan data yang menjadi sumber untuk data mining termasuk data “cleaning” untuk mempelajari polanya). Setelah menemukan definisi masalah, langkah berikutnya adalah mencari data yang mendukung definisi masalah. Menentukan porsi data yang digunakan men-training data mining berdasarkan algoritma data mining yang telah dibuat. Setelah persiapan data selesai dilakukan, langkah berikutnya adalah memberikan sebagian data ke dalam algoritma data mining.
3. Tahap ketiga: Model building and validation. Validasi apakah data mining memberikan prediksi yang akurat. Setelah training data selesai dilakukan, data mining tersebut perlu di-“uji” atau di-validasi keakuratannya terhadap data testing.
4. Tahap ke-empat: Deployment. Tahap ini memilih aplikasi yang tepat terhadap data mining untuk membuat prediksi.



Gambar 3 : Tahapan dalam Data Mining dan Membangun Model

Sumber: (Finn Lee & Juan Santana 2010:41)

2.6.3. Tahapan Association Rule

Analisis asosiasi dikenal juga sebagai salah satu teknik data mining yang menjadi dasar dari berbagai teknik data mining lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien (Muhammad Ikhsan et al, 2007).

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap :

1. Analisa pola frekuensi tinggi

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

$Support(A) =$

$$\frac{JumlahTransaksiMenganandungA}{TotalTransaksi}$$

aksi

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

$Support(A,B) = \frac{JumlahTransaksiMenganandungAB}{TotalTransaksi}$

$$\frac{JumlahTransaksiMenganandungAB}{TotalTransaksi}$$

aksi

2. Pembentukan Aturan Asosiasi

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

$Confidence = P(B|A) =$

$$\frac{JumlahTransaksiMenganandungAB}{JumlahTransaksiMenganandungA}$$

aksi

2.7. Langkah-Langkah Proses Aturan Asosiasi

Proses Aturan Asosiasi terdiri dari beberapa tahap sebagai berikut (Eko Wahyu Tyas d, 2008).

1. Sistem men-scan database untuk mendapat kandidat *1-itemset* (himpunan item yang terdiri dari 1 item) dan menghitung nilai supportnya. Kemudian nilai *supportnya* tersebut dibandingkan dengan *minimum support* yang telah ditentukan, jika nilainya lebih besar atau sama dengan *minimum support* maka *itemset* tersebut termasuk dalam *large itemset*.
2. *Itemset* yang tidak termasuk dalam *large itemset* tidak diikuti dalam iterasi selanjutnya (di *prune*).
3. Pada iterasi kedua sistem akan menggunakan hasil *large itemset* pada iterasi pertama (L_1) untuk membentuk kandidat *itemset* kedua (L_2). Pada iterasi selanjutnya sistem akan menggunakan hasil *large itemset* pada iterasi selanjutnya akan menggunakan hasil *large itemset* pada iterasi sebelumnya (L_{k-1}) untuk membentuk kandidat *itemset* berikut (L_k). Sistem akan menggabungkan

- (join) L_{k-1} dengan L_{k-1} untuk mendapatkan L_k , seperti pada iterasi sebelumnya sistem akan menghapus (*prune*) kombinasi itemset yang tidak termasuk dalam large itemset.
- Setelah dilakukan operasi *join*, maka pasangan itemset baru hasil proses *join* tersebut dihitung supportnya.
 - Proses pembentuk kandidat yang terdiri dari proses *join* dan *prune* akan terus dilakukan hingga himpunan kandidat itemsetnya null, atau sudah tidak ada lagi kandidat yang akan dibentuk.
 - Setelah itu, dari hasil *frequent itemset* tersebut dibentuk *association rule* yang memenuhi nilai *support* dan *confidence* yang telah ditentukan.
 - Pada pembentukan *association rule*, nilai yang sama dianggap sebagai satu nilai.
 - Association rule* yang terbentuk harus memenuhi nilai minimum yang telah ditentukan.
 - Untuk setiap *large itemset* L , kita cari himpunan bagian L yang tidak kosong. Untuk setiap himpunan bagian tersebut, dihasilkan *rule* dengan bentuk $aB(L-a)$ jika *supportnya* (L) dan *supportnya* (a) lebih besar dari minimum *support*.

2.8. Algoritma Apriori

Algoritma apriori adalah suatu algoritma dasar yang diusulkan oleh Agrawal & Srikant pada tahun 1994 untuk menentukan Frequent itemsets untuk aturan asosiasi Boolean. Algoritma Apriori termasuk jenis *Aturan Asosiasi* pada data mining. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut *affinity analysis* atau *market basket analysis*. Analisis asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan suatu kombinasi *item*. Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (*frequent pattern mining*). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolok ukur, yaitu : *support* dan *confidence*. *Support* (nilai penunjang) adalah persentase kombinasi *item* tersebut dalam database, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antara-item dalam aturan asosiasi. (Kusrini, Emha Taufiq Luthfi, 2009 : 149)

Algoritma apriori dibagi menjadi beberapa tahap yang disebut narasi atau pass (Devi dinda setiawan, 2009).

- Pembentukan kandidat itemset.
Kandidat k-itemset dibentuk dari kombinasi (k-1)-itemset yang didapat dari iterasi sebelumnya. Satu cara dari algoritma apriori adalah pemangkasan kandidat k-itemset yang subsetnya berisi k-1 item tidak termasuk dalam pola frekuensi tinggi dengan panjang k-1.
- Penghitungan support dari tiap kandidat k-itemset.
Support dari tiap kandidat k-itemset didapat dengan menscan database untuk menghitung jumlah transaksi yang memuat semua item didalam kandidat k-itemset tersebut. Ini adalah

juga ciri dari algoritma apriori dimana diperlukan penghitungan dengan cara seluruh database sebanyak k-itemset terpanjang.

- Tetapkan pola frekuensi tinggi.
Pola frekuensi tinggi yang memuat k item atau k-itemset ditetapkan dari kandidat k-itemset yang supportnya lebih besar dari minimum *support*.
- Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan.

Untuk lebih memahami proses algoritma Apriori maka berikut ini akan diberikan ilustrasi penggunaan algoritma Apriori. Dengan menggunakan database pada gambar 4 dan mengasumsikan minimum *support* adalah 2 transaksi.

Database		\hat{C}_1		L_1	
TID	Items	TID	Set-of-items	Itemset	Support
100	1 3 4	100	{(1),(3),(4)}	{1}	2
200	2 3 5	200	{(2),(3),(5)}	{2}	3
300	1 2 3 5	300	{(1),(2),(3),(5)}	{3}	3
400	2 5	400	{(2),(5)}	{5}	3

\hat{C}_2		\hat{C}_2		L_2	
Itemset	Support	TID	Set-of-items	Itemset	Support
{1 2}	1	100	{(1 3)}	{1 3}	2
{1 3}	2	200	{(2 3),(2 5),(3 5)}	{2 3}	2
{1 5}	1	300	{(1 2),(1 3),(1 5),(2 3),(2 5),(3 5)}	{2 5}	3
{2 3}	2	400	{(2 5)}	{3 5}	2
{2 5}	3				
{3 5}	2				

\hat{C}_3		\hat{C}_3		L_3	
itemset	Support	TID	Set-of-items	Itemset	Support
{2 3 5}	2	200	{(2 3 5)}	{2 3 5}	2
		300	{(2 3 5)}		

Gambar 4 : Ilustrasi Algoritma Apriori
(Sumber : Agrawal, R., Srikant, 1996)

2.9. Apotek dan Alat-Alat Kesehatan

Pengertian Apotek menurut (Kepmenkes RI). No. 1332/M ENKES/SK/X/2002, Apotek adalah suatu tempat tertentu, tempat di lakukan pekerjaan kefarmasian penyaluran perbekalan farmasi kepada masyarakat. Yang di maksud pekerjaan kefarmasian di antaranya pengadaan obat, penyimpanan obat, pembuatan sediaan obat, peracikan, penyaluran dan penyerahan perbekalan farmasi serta memberikan informasi kepada masyarakat mengenai perbekalan kefarmasian yang terdiri dari obat, bahan obat, obat tradisional, alat-alat kesehatan, dan kosmetik. (<http://farmasismk.blogspot.com/2011/11/pengertian-apotek.html>)

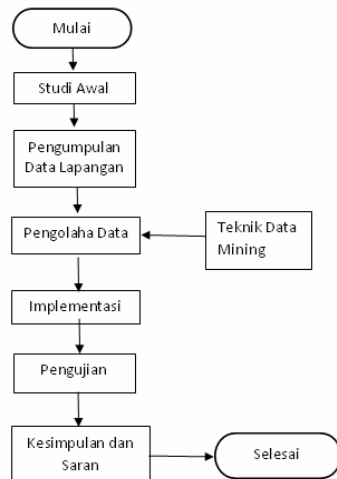
3. METODOLOGI PENELITIAN

3.1. Kerangka Pemikiran.

Kerangka pemikiran dari penelitian ini dilakukan dengan beberapa tahap bahwa pada tahapan pertama adalah pendefinisian rumusan masalah. Kemudian pada tahapan yang kedua melakukan analisis dan desain menggunakan teknik data mining untuk membantu memecahkan rumusan masalah yang ada. Tahap ketiga adalah metode pengumpulan data dimana data di ambil sebagai kepentingan penelitian dengan menggunakan sampel. Di tahap keempat merupakan pencarian subjek penelitian, dimana peneliti mendapatkan data yang di butuhkan dalam pencapaian tujuannya. Hasil yang di peroleh dari penelitian ini adalah penentuan pola kombinasi *itemset frekuensi tinggi* (*frequent patternt*)

yaitu yang memperoleh *support* dan *confidence* tertinggi. Sedangkan tahap terakhir adalah implementasi dan pengujian menggunakan sistem operasi aplikasi software tanagra 1.4, dan database dengan microsoft excel 2007.

Langkah - langkah di dalam kerangka pemikiran tersebut dapat di gambarkan dalam bentuk flowchart berikut ini :



Gambar 5 : Flowchart penelitian.

3.2. Uraian Flowchart Penelitian

Berikut ini langkah – langkah dari flowchart penelitian yang sudah di buat.

1. Studi Awal

Langkah awal dari penelitian ini adalah dengan mencari dan mempelajari masalah yang akan di teliti. Kemudian menentukan ruang lingkup masalah, latar belakang masalah, dan mempelajari beberapa literatur yang berkaitan dengan permasalahan dan bagaimana mencari solusi dari masalah tersebut. Untuk mencapai tujuan yang akan ditentukan, maka perlu dipelajari beberapa literatur-literatur yang digunakan. Kemudian literatur-literatur yang dipelajari tersebut diseleksi untuk dapat ditentukan literatur mana yang akan digunakan dalam penelitian.

Melalui studi literatur, dipelajari teori-teori yang berhubungan dengan *data mining*, *association rule*, serta algoritma apriori yang akan dibahas. Teori-teori tersebut diatas berdasarkan sumber berupa buku, jurnal, paper, maupun situs internet yang berhubungan dengan *data mining*.

2. Pengumpulan Data

Metode yang digunakan dalam proses pengumpulan data dalam penelitian ini adalah sebagai berikut:

3. Wawancara

Wawancara diperlukan oleh penulis guna menanyakan secara langsung tentang apa saja masalah yang sering dihadapi oleh pihak Apotek khususnya tentang persediaan alat-alat kesehatan karena data tersebut dapat digunakan sebagai

salah satu sumber pertimbangan dalam pengambilan keputusan.

4. Pengamatan/Observasi

Sebagai metode ilmiah observasi dapat diartikan sebagai pengamatan. Jadi observasi merupakan suatu penyelidikan yang dilakukan secara sistematis dan sengaja dilakukan dengan menggunakan alat indera terutama mata terhadap kejadian yang sedang berlangsung dan dapat dianalisa pada waktu kejadian itu terjadi. Dalam penelitian ini penulis melakukan pengamatan terhadap objek data dengan tujuan mendapatkan korelasi antara persediaan dengan kebutuhan, permintaan dan lain-lain.

5. Dokumentasi

Metode dokumentasi adalah mencari data mengenai hal-hal atau variabel berupa catatan buku, surat, transkrip, majalah, prasasti, notulen, rapat, lengger, agenda dan sebagainya baik yang berupa *file* atau dokumen. Dalam hal ini penulis melakukan pengumpulan dokumentasi berupa dokumentasi *file* data atau data faktur barang keluar untuk digunakan sebagai bahan implementasi dan uji coba.

6. Pengolahan Data

Pada tahap pengolahan data terlebih dahulu melakukan identifikasi masalah yang ada dan sering dihadapi oleh pihak ApotE, untuk kemudian mendeskripsikan masalah-masalah tersebut untuk diperoleh solusinya.

Tahap selanjutnya dilakukan analisa masalah, dengan tujuan agar penulis mengetahui dan memperoleh gambaran yang jelas bagaimana bentuk penyelesaian dan algoritma apa yang dapat digunakan untuk penyelesaiannya. Selanjutnya penulis menggunakan teknik data mining dengan algoritma Apriori untuk mendapatkan hasil sebagai tujuan yang akan dicapai oleh peneliti yang kemudian dapat dijadikan pihak Apotek sebagai pengetahuan dalam meningkatkan penjualan alat-alat kesehatan kepada konsumen.

7. Implementasi

Pada tahap ini untuk memudahkan penulis dalam pembuktian hasil analisa yang dilakukan, maka penulis menggunakan suatu *tools* atau aplikasi data mining yang telah ada. Sedangkan *database* yang digunakan untuk menyimpan data-data yang ada adalah penulis menggunakan *database* Microsoft Excel.

8. Pengujian

Pada tahap ini penulis melakukan pengujian dengan menggunakan *tools* atau aplikasi data mining dengan menghubungkan dengan *database* yang telah diisi atau berisi data-data yang akan diuji.

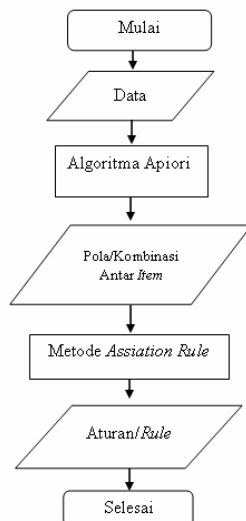
Mekanisme pengujian sistem dimulai dengan mempersiapkan *hardware* dan *software* yang dibutuhkan untuk menjalankan sistem yang telah ada. Sistem yang telah tersedia merupakan suatu

aplikasi *data mining* yang pada dasarnya telah siap untuk dijalankan dikomputer.

4. HASIL DAN PEMBAHASAN

4.1 Analisa

Sistem informasi persediaan alat-alat kesehatan di Apotek Kelambir – 2 Medan sudah terkomputerisasi artinya semua yang berhubungan dengan aktifitas penjualan dan pembelian menggunakan komputer yang berbasis jaringan dengan database terpusat kedalam satu server. Sistem informasi persediaan barang di Apotek Kelambir – 2 Medan terdiri atas bagi-bagian data obat dan alat-alat kesehatan, gudang farmasi, pelayanan resep obat dan penjualan. Ditampilkan gudang farmasi terdiri atas bagian persediaan, laporan pembelian, laporan penjualan sedangkan data alat-alat kesehatan keluar terdiri atas *attribute* nomor, tanggal, nomor slip, nama, jumlah barang, harga, dan total harga dan setiap transaksi penjualan minimal terdiri atas satu jenis alat-alat kesehatan. Untuk mendapat analisa data maka data penjualan di *export* kedalam database Microsoft excel karena database Microsoft excel bersifat *spreadsheet* sehingga sangat mendukung dalam analisa data. Microsoft excel dalam analisis data digunakan sebagai database karena Microsoft excel sangat mendukung beberapa aplikasi data mining, dimana aplikasi (perangkat lunak) data mining digunakan sebagai tempat uji coba atau implementasi. Data mining ialah proses menambang data untuk mendapatkan ilmu pengetahuan atau informasi penting dari database khususnya database penjualan alat-alat kesehatan.



Gambar 6 : Flowchart Pengolahan Data Mining

4.2. Pembahasan

4.2.1 Analisis Data

Berhubungan dengan studi kasus pada Apotik Kelambir – 2 Medan, dapat dilakukan analisis data. Analisis data dilakukan khusus pada penjualan (data alat kesehatan) dengan tujuan untuk menemukan

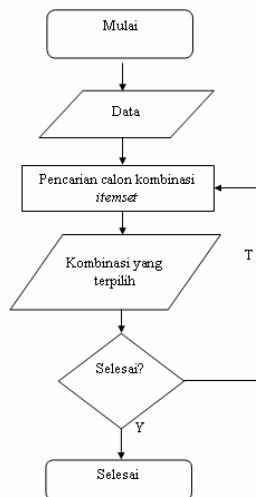
pola penjualan alat-alat kesehatan dan hubungan antar *item* jenis alat-alat kesehatan didalam transaksi. Berikut ini adalah tabel 1 sebagai sampel data dan sebagian terlampir yang akan dijadikan untuk analisa dan juga untuk uji coba.

Tabel 1 : Daftar Real Penjualan Alat-alat Kesehatan

NO	No.Faktur	Nama Barang
1.	011/k/II/13	a. Infuset b. Abocat
2.	012/k/III/13	a. Abocat b. Medicep
3.	017/k/VIII/13	a. Leukoplast b. Termometer c. Stick Gula
4.	019/k/X/13	a. Masker b. Handscound c. Stick Gula d. Stick Colestrol e. Urinal
5.	020/k/I/13	a. Stick Gula b. Stick Asam Urat
6.	023/k/V/13	a. Nald b. Dermafix c. Urinal
7.	028/k/IX/13	a. Stick Pan b. Dermafix
8.	031/k/I/13	a. Abocat b. Kapas c. Dermafix
9.	032/k/II/13	a. Lumatul b. Leukoplast c. Stick Gula
10.	043/k/III/13	a. Stick Colestrol b. Stick Gula c. Stick Asam Urat
11.	056/k/X/13	a. Stick Gula b. Stick Asam Urat c. Stick Colestrol
12.	057/k/I/13	a. Camelux b. Masker c. Kasa d. Kapas
13.	072/k/VII/13	a. Verban b. Kapas c. Infuset d. Kasa
14.	086/k/I/13	a. Breast Pump b. Stick Pan c. Buli - buli Panas
15.	089/k/IV/13	a. Nald b. Abocat c. Infuset
16.	090/k/V/13	a. Sput b. Masker c. Kapas
17.	091/k/VI/13	a. Handscound b. Stick Gula c. Stick Asam
18.	096/k/I/13	a. Accu Cek Strip b. Tensi Meter c. Alkohol Swab
19.	111/k/VI/13	a. Hipapix b. Urine Bag c. Abocat d. Kasa
20.	114/k/IX/13	a. Tes Hamul b. Kotak P3k c. Nald d. Buli - buli Panas e. Masker
21.	118/k/III/13	a. Leukoplast b. Adcare c. Hipapix
22.	122/k/VII/13	a. Camelux b. Alkohol Swab c. Kapas
23.	124/k/IX/13	a. Dermafix b. Leukoplast c. Verban d. Kasa
24.	138/k/III/13	a. Stick Gula b. Stick Asam Urat c. Stick Colestrol
25.	142/k/VII/13	a. Termometer b. Leukoplast c. Tes Hamul

Tahap dalam menganalisa data dengan algoritma Apriori pada penjualan (data alat-alat kesehatan)

dimulai dengan menyeleksi dan membersihkan data data yang akan dianalisis, kemudian mencari semua jenis item nama alat-alat kesehatan yang ada didalam *list* transaksi penjualan, selanjut mencari jumlah setiap item yang ada pada setiap transaksi penjualan (alat-alat kesehatan). Sesuai dengan *support* yang telah ditentukan maka terseleksi lah beberapa item data yang memenuhi minimal *support* sesuai dengan jumlah item barang didalam transaksi, itu disebut pembentukan kombinasi satu item. Tahap selanjutnya dari *item-item* alat-alat kesehatan yang telah terseleksi dibentuk lah kombinasi dua *item*, maka terbentuk lah beberapa item data dengan kombinasi 2 item yang berbeda, dengan *support* yang ditentukan maka terseleksi lah beberapa data dua *item*, ini disebut pembentukan kombinasi dua *item*. Demikian seterusnya sampai kombinasi batas maksimal *item* transaksi.



Gambar 7 : Flowchart algoritma Apriori

4.2.2. Analisa Pola Frekuensi Tinggi

Sebelum dilakukan pencarian pola dari data transaksi terlebih dulu, dicari semua nama jenis item alat-alat kesehatan yang ada didalam transaksi seperti pada tabel 1 sekaligus menentukan support peritem jenis alat-alat kesehatan, dimana tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam database, nilai support sebuah item diperoleh dengan rumus berikut:

$$Support(A) = \frac{Jumlah_{transaksi_yang_mengandung\ A}}{\sum Transaksi}$$

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

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

$$Support(A, B) =$$

$$\frac{\sum Transaksi_yang_mengandung_A_dan_B}{\sum Transaksi}$$

Berikut merupakan tabel dari semua jenis *itemset* alat-alat kesehatan didalam transaksi penjualan (data barang keluar), seperti yang ditunjukkan pada tabel 2 berikut ini:

Tabel 2 : Daftar Jenis *items* Alat-alat Kesehatan

No.	Nama item	Support	Support (%)
1.	Abocat	5	20%
2.	Accu Cek Strip	1	4%
3.	Adesare	1	4%
4.	Alkohol Swab	2	8%
5.	Breast Pump	1	4%
6.	Boli - Boli Panas	2	8%
7.	Cannexur	2	8%
8.	Demasfir	3	12%
9.	Handsound	2	8%
10.	Hipapir	2	8%
11.	Influset	3	12%
12.	Kapas	5	20%
13.	Kasa	4	16%
14.	Kotak P3k	1	4%
15.	Leukoplast	5	20%
16.	Lumabul	1	4%
17.	Masker	4	16%
18.	Medicrep	1	4%
19.	Nald	3	12%
20.	Sprit	1	4%
21.	Stick Asam Urat	5	20%
22.	Stick Colestol	4	16%
23.	Stick Gula	8	32%
24.	Stick Pan	2	8%
25.	Tensi Meter	1	4%
26.	Termometer	2	8%
27.	Tes Hamil	2	8%
28.	Urnal	2	8%
29.	Urine Bag	1	4%
30.	Verban	2	8%

Data diatas menggambarkan bentuk data 1 *item* yang terdiri atas *attribute item* sebagai nama *item* jenis semua alat-alat kesehatan yang ada didalam transaksi, support yaitu jumlah setiap item yang ada disemua transaksi, sedangkan *support(%)* adalah adalah presentasi jumlah *item* yang ada didalam transaksi, yang didapat dari jumlah item dibagi jumlah semua transaksi yang akan dianalisis di kali seratur persen. Sedangkan Tabel 3 adalah *item* data yang terpilih dengan minimal support adalah 16 persen(%). Seperti yang terlihat pada tabel berikut ini:

Tabel 3 : Daftar Jenis *items* Alat-alat Kesehatan dengan *support* yang telah ditentukan

No.	Nama item	Support	Support (%)
1.	Abocat	5	20%
2.	Kapas	5	20%
3.	Kasa	4	16%
4.	Leukoplast	5	20%
5.	Masker	4	16%
6.	Stick Asam Urat	5	20%
7.	Stick Colestol	4	16%
8.	Stick Gula	8	32%

Tabel diatas merupakan data *item* jenis alat-alat kesehatan yang terseleksi atau terpilih sesuai dengan *support* yang telah ditentukan. Dimana data diatas akan digunakan untuk membentuk pola atau *kombinasi item* dan juga digunakan untuk menentukan *support* dan *confidence* pada pembahasan selanjutnya.

4.2.3. Pembentukan Pola Kombinasi dua *items*

Pembentukan pola frekuensi dua *item*, dibentuk dari *items-items* jenis alat – alat kesehatan yang memenuhi *support* minimal yaitu dengan cara

mengkombinasi semua *items* kedalam dua kombinasi, hasil dari kombinasi dua *items* seperti pada tabel 4 berikut ini:

Tabel 4 : Daftar Calon Pola Kombinasi Dua *itemset*

No.	Nama item	Support	Support (%)
1.	Abocat, kapas	1	4%
2.	Abocat, kasa	1	4%
3.	Abocat, Leukoplast	0	0
4.	Abocat, Masker	0	0
5.	Abocat, Stick Asam Urat	0	0
6.	Abocat, Stick Colestol	0	0
7.	Abocat, Stick Gula	0	0
8.	Kapas, kasa	2	8%
9.	Kapas, Leukoplast	0	0
10.	Kapas, Masker	1	4%
11.	Kapas, Stick Asam Urat	0	0
12.	Kapas, Stick Colestol	0	0
13.	Kapas, Stick Gula	0	0
14.	Kasa, Leukoplast	1	4%
15.	Kasa, Masker	1	4%
16.	Kasa, Stick Asam Urat	0	0
17.	Kasa, Stick Colestol	0	0
18.	Kasa, Stick Gula	0	0
19.	Leukoplast, Masker	0	0
20.	Leukoplast, Stick Asam Urat	0	0
21.	Leukoplast, Stick Colestol	0	0
22.	Leukoplast, Stick Gula	2	8%
23.	Masker, Stick Asam Urat	0	0
24.	Masker, Stick Colestol	1	4%
25.	Masker, Stick Gula	1	4%
26.	Stick Asam Urat, Stick Colestol	3	12%
27.	Stick Asam Urat, Stick Gula	4	16%
28.	Stick Colestol, Stick gula	4	16%

Data diatas merupakan calon kombinasi dua *item* yang merupakan hasil dari semua kombinasi semua jenis *item*. Dengan menetapkan *support* minimal sama dengan 16% persen, maka data diatas terseleksi atau terpilih, seperti pada tabel 5 berikut ini:

Tabel 5 : Daftar Pola kombinasi dua *items* yang memenuhi *support* minimal

No.	Nama item	Support	Support (%)
1.	Stick Asam Urat, Stick Gula	5	20%
2.	Stick Colestol, Stick gula	4	16%

Data diatas adalah kombinasi pola dua *item* data yang terpilih dengan *support* yang telah ditentukan, terlihat data kombinasi jenis **Stick Asam Urat, Stick Gula** dan **Stick Colestol, Stick gula** memiliki *support* yang terbanyak, itu menandakan bahwa kombinasi dua *item* data tersebut paling banyak didalam transaksi.

4.2.3 Pembentukan Pola Aturan Asosiasi

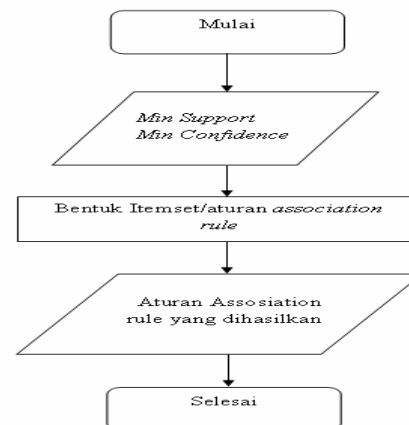
Setelah semua pola frekuensi tinggi ditemukan, baru la dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiasi A ke B. Nilai *confidence* dari aturan A ke B diperoleh dengan rumus:

$$Confidence = P(B | A) = \frac{\sum \text{Transaksi_mengandung_A_dan_B}}{\sum \text{Transaksi_yang_mengandung_A}}$$

Proses mencari jumlah kombinasi dan kuatnya hubungan antara satu item dengan item yang

lain dalam satu kombinasi disebut metode *association rule*. pembentukan *association rule* adalah menganalisis pola frekuensi tinggi, tahap ini mencari kombinasi yang memenuhi syarat minimum dari *support* dalam database., Pembentukan aturan *association rule*, dengan mencari nilai *confidence*. Dimana *support* adalah jumlah dari kombinasi antara suatu *item* dengan *item* yang lain sedangkan *confidence* adalah nilai yang mendefinisikan kuat tidaknya hubungan antara *item-item* tersebut.

Berikut ini adalah flowchat pembentukan Association Rule seperti yang terlihat pada gambar 8 berikut ini:



Gambar 8 : Flowchart Association Rule

Dari tabel 5 yaitu tabel pola kombinasi dua *item*, dapat dilihat besarnya nilai *support* dan *confidence* dari calon aturan asosiasi seperti tampak pada tabel 6 berikut ini:

Tabel 6 : Daftar Calon Aturan Asosiasi dari tabel 5

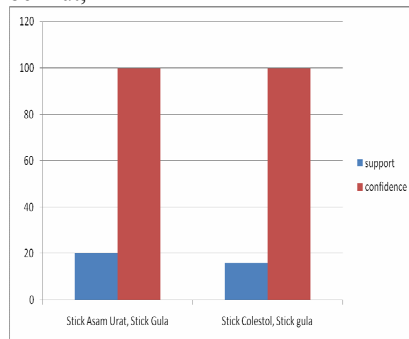
No.	Kombinasi Item (merek)	Confidence	
1.	Jika membeli Stick Asam Urat, maka akan membeli Stick Gula	5/5	100%
2.	Jika membeli Stick Gula, maka akan membeli Stick Asam Urat	5/8	62,5%
3.	Jika membeli Stick Colestol, maka akan membeli Stick gula	4/4	100%
4.	Jika membeli Stick gula, maka akan membeli Stick Colestol	4/8	50%

Dengan menetapkan nilai *confidence* minimum adalah 70 persen(%), maka aturan yang bisa terbentuk adalah aturan seperti pada tabel 7 berikut ini:

Tabel 7 : Daftar Aturan Asosiasi yang terbentuk dari tabel 5

No.	Kombinasi Item (merek)	support	Confidence
1.	Jika membeli Stick Asam Urat, maka akan membeli Stick Gula	20%	100%
2.	Jika membeli Stick Colestol, maka akan membeli Stick gula	16%	100%

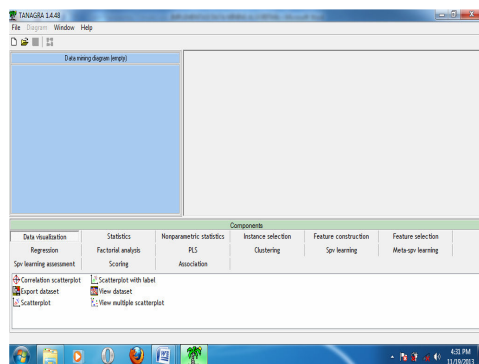
Berdasarkan hasil yang diperoleh berupa data statistik seperti yang tertera pada tabel 7 diatas, maka dapat dinyatakan dalam bentuk diagram sebagai berikut;



Gambar 9 : Grafik korelasi antara Support dan Confidence

4.3. Pengujian

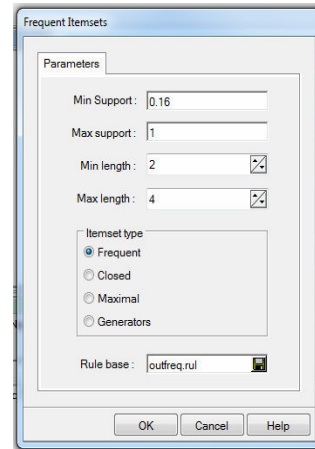
Untuk membuktikan data-data yang telah dihasilkan berupa pola hubungan kombinasi antar *items* dan *rules-rules* asosiasi sesuai dengan Algoritma Apriori maka perlu dilakukan pengujian dengan menggunakan suatu aplikasi. Aplikasi yang digunakan adalah Tanagra versi 1.4 sebagai berikut:



Gambar 10 : Tampilan Utama Tanagra

Tampilan diatas adalah merupakan tampilan utama aplikasi Tanagra. Tampilan bagian atas horizontal adalah merupakan menu utama program yang terdiri atas menu *File*, menu *Diagram*, menu *Component*, menu *Windows* dan menu *Help*. Jendela sebelah kanan merupakan tempat komponen yang akan diproses hasilnya, jendela bagian kiri yang lebar merupakan tempat hasil dari proses, sedangkan bagian bawah adalah merupakan tempat komponen berbagai jenis algoritma, termasuk algoritma apriori. Secara umum, cara menggunakan aplikasi Tanagra dimulai dengan mengklik menu *File*, *New*, muncul jendela untuk *load database*. *Load database* supaya masuk kesistem, klik menu *icon define status* dan atur parameternya, masukkan komponen *frequent itemsets* dan *Apriori* dan masukkan parameternya, kemudian *execute*, setelah *view* untuk melihat hasilnya.

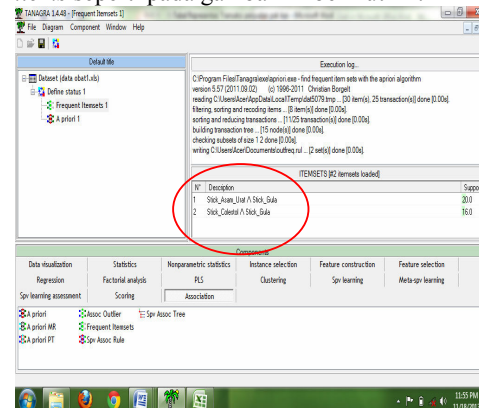
4.3.1. Pengujian Untuk Menghasilkan Pola Kombinasi Itemsets



Gambar 11 : Jendela *Frequent Itemsets*

Pengujian pertama yang dilakukan dengan aplikasi Tanagra adalah dengan menghitung frekuensi *items* atau *frequent itemsets*. Pengujian yang dilakukan pertama adalah pengujian untuk menghasilkan jenis *item* dan pola kombinasi dua *items*. Dengan parameters adalah minimal *support* sama dengan 16 persen (%), maximal *support* sama dengan 100 persen (%), minimal *length* sama dengan satu, maksimal *length* sama dengan dua, dan *itemset type* sama dengan *frequent*.

Setelah dilakukan penentuan nilai parameter, maka dihasilkan jenis *item* dan pola kombinasi dua *items* seperti pada gambar 12 berikut ini:



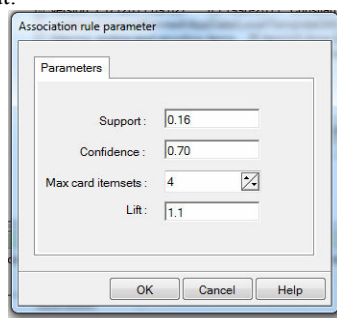
Gambar 12 : Hasil Pengujian dengan Satu *item* dan Pola Kombinasi dua *itemsets*

Tampilan pengujian diatas yaitu pengujian untuk menghasilkan jenis *item* dan pola kombinasi dua *itemset*. Terlihat ada dua *attribute* yang digunakan, adalah *Description* dan *Support*. *Description* menggambarkan jenis *items* dan bentuk pola kombinasi dua *items*, Sedangkan *support* adalah nilai presentasi banyak jumlah jenis *items*, dan jumlah persentasi banyak *items* dan pola kombinasi dua *items* didalam transaksi.

4.3.2. Pengujian Untuk Menghasilkan Rules

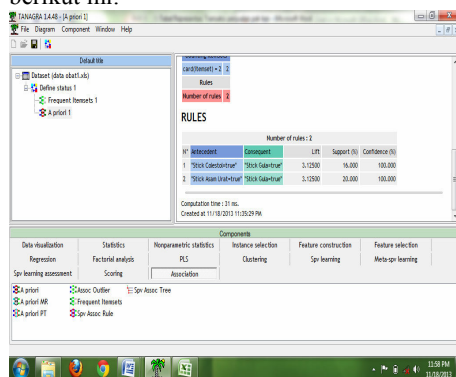
Setelah melakukan pengujian untuk menghasilkan jenis *items* dan pola kombinasi, mulai dari kombinasi dua *items* sampai dengan pola kombinasi lima *items*, maka pengujian yang dilakukan selanjutnya adalah pengujian untuk menghasilkan *rules-rules*, dimana *rules-rules* tersebut terbentuk dari pola kombinasi *items* pada pengujian sebelumnya. Pengujian yang dilakukan pertama adalah pengujian untuk menghasilkan *rules-rules* yang terbentuk dari pola kombinasi dua *items*.

Parameters yang harus diisi adalah *support* sama dengan 16 persen (%), *Confidence* sama dengan 70 persen (%), maksimal *card itemsets* sama dengan dua. Jendela *Association rules* adalah seperti pada gambar 13 berikut:



Gambar 13 : Jendela *Association Rules Parameters*

Setelah dilakukan penentuan nilai parameters, maka *rules-rules* yang dihasilkan dari parameters diatas adalah seperti pada gambar 14 berikut ini:



Gambar 14 : Pengujian dengan *rules* dari pola kombinasi dua *itemsets*

Pada pengujian diatas yaitu pengujian untuk menghasilkan *rules* dari pola kombinasi dua *items*. *Rules* diatas terdiri atas *Antecedent*, *consequent*, *lift*, *support*(%), *Confidence* (%). *Antecedent* adalah bentuk kondisi dari pada *rules*, *consequent* adalah bentuk pernyataan dari pada *rules*, *lift* adalah menunjukkan adanya tingkat kekuatan *rules* kejadian acak dari *antecedent* dan *consequent* berdasarkan pada *support* masing-masing, *support* adalah persentasi kombinasi *items* tersebut, sedangkan *confidence* adalah kuatnya hubungan antar *items* dalam aturan asosiasi.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan pembahasan yang telah dilakukan dengan algoritma *Apriori* dan dilakukannya pengujian dengan aplikasi *Tanagra* maka penulis menarik beberapa kesimpulan yang penting.

Adapun kesimpulan-kesimpulan tersebut adalah sebagai berikut:

1. Data Mining dapat di implementasikan dengan menggunakan Database penjualan alat-alat kesehatan karena dapat menemukan kecenderungan pola kombinasi *itemsets* sehingga dapat dijadikan sebagai informasi yang sangat berharga dalam pengambilan keputusan untuk mempersiapkan stok jenis barang apa yang diperlukan kemudian.
2. Penerapan Algoritma *Apriori* pada teknik Data Mining sangat efisien dan dapat mempercepat proses pembentukan kecenderungan pola kombinasi itemset hasil penjualan alat-alat kesehatan di Apotek Kelambir-2 Medan, yaitu dengan support dan confidence tertinggi adalah *Stick Asam Urat - Stick Gula* dan *Stick Colestrol - Stick Gula*.

5.2 Saran

Untuk kepentingan lebih lanjut dari penulisan tesis ini maka penulis memberikan beberapa saran sebagai berikut :

1. Data mining dengan Algoritma *Apriori* memiliki kelemahan karena harus melakukan *scan* database setiap kali iterasi, sehingga untuk database yang sangat besar membutuhkan waktu yang lama.
2. Penerapan algoritma *Apriori* sangat praktis namun perlu dilakukan perbandingan dengan algoritma lain, untuk menguji sejauh mana *Algoritma Apriori* masih dapat diandalkan untuk memproses dan menemukan pola hubungan (asosiasi) antar *item* pada database berskala besar.
3. Dalam penggunaan algoritma *Apriori* di perlukan pemahaman tentang aturan Asosiasi dan Data Mining dan cara menjalankan aplikasi software *Tanagra* 1.4.48.

DAFTAR PUSTAKA

- [1] Jiawei Han, Micheline Kamber. 2006 "Data Mining ; Concepts and Techniques", Morgan Kaufmann Publishers, Champaign ; CA9411.
- [2] Kusriani dan Luthfi, E. T., 2009. *Algoritma Data Mining*. Yogyakarta : Penerbit Andi.
- [3] Anand Rajaraman, Jure Leskovec, and Jeffrey D. Ullman. 2013. "Mining of Massive Datasets". Stanford Univ., CS345A.
- [4] Ramon A. Mata Toledo, Pailine K. Cushman. 2007. *Dasar-dasar Database Relasional*. Jakarta : Erlangga.

- [5] Finn Lee S & Juan Santana. 2010. *Data Mining : Meramalkan Bisnis Perusahaan*. Jakarta : Penerbit PT. Elex Media Komputindo.
- [6] Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., and Verkamo, I. 1996. Fast Discovery of Association Rules, in AKDDM, AAAI/MIT Press, 307-328.
- [7] Fayyad, U.M., Piatetsky-Shapiro, G., and Smyth, P. 1996. *Knowledge Discovery and Data Mining : Towards a Unifying Framework*. In Proceedings of KDD-96 (AKDDM), AAAI/MIT Press, pp. 82-88.
- [8] Yeong-Chyi Lee, Tzung-Pei Hong, and Wen-Yang Lin. *Mining Association rules with multiple minimum support using maximum constraints*. International Journal of Approximate Reasoning 40 (2005) 44-45.
- [9] N. Badal, and Shuruti Tripathi. *Frequent Data Itemset Mining Using VS_Apriori Algorithm*. International Journal on Computer Science and Engineering (IJCE)-VOL.02, No. 04, 2010.
- [10] Pratima Gautam, and K. R. Pardasani. *Efficient Method for Multiple-Level Association Rules in Large Databases*. Journal of Emerging Trends in Computing and Information Sciences. VOL.2 No.12. December 2011.
- [11] Abdullah Saad Almalaise Alghamdi., *Efficient Implementation of FP Growth Algoritma-Data Mining on Medical Data*. International Journal of Computer Science and Network Security, VOL.11 No.12, December 2011.
- [12] Sanjeev Rao, and Priyanka Gupta. *Implementing Improved Algorithm Over Apriori Data Mining Association Rule Algorithm*. International Journal of Computer Science and Technology (IJCT)-VOL.3 Issue 1, Jan-March 2012.
- [13] OTHMAN YAHYA, OSMAN HEGAZY, and EHAB EZAT. *An Efficient Implementation of Apriori Algorithm Based on HADOOP-MAPREDUCE Model*. International Journal of Reviews in Computing (IJRIC)-VOL.12, December 2012.
- [14] Jogi. Suresh, and T. Ramanjaneyulu. *Mining Frequent Itemsets Using Apriori Algorithm*. International Journal of Computer Trends and Technology (IJCTT)-VOL.4 Issue4-April 2013.
- [15] Shuruti Aggarwal, and Ranveer Kaur. *Comparative Study of Various Improved Versions of Apriori Algorithm*. International Journal of Engineering Trends and Technology (IJETT)-VOL.4 Issue4-April 2013.
- [16] Prof. Paresh Tama., and Dr. Yogesh Ghodasara. *Foundation for Frequent Pattern Mining Algorithms' Implementation*. International Journal of Computer Trends and Technology (IJCTI)-VOL.4 Issue 7 - July 2013.
- [17] Ema Utami, Jazi Eko Istiyanto, dan Suwanto Raharjo. *METODOLOGI PENELITIAN PADA ILMU KOMPUTER*. Seminar Nasional

Teknologi 2007 (SNT 2007) Yogyakarta, 24 No'vember 2007. ISSN : 1978 – 9777

AUTHOR BIOGRAPHIES



Kennedi Tampubolon, lahir di Tapanuli Selatan tanggal 22 April 1967. Pada tahun 1993 menyelesaikan pendidikan S1 di Universitas Sumatera Utara jurusan Matematika, saat ini sedang menyusun tesis Magister Komputer di STMIK ERESHA School Of IT Jakarta. Saat ini juga aktif sebagai dosen tetap pada STMIK Budi