

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Pengertian konsumen**

Konsumen adalah setiap orang pemakai barang dan jasa yang tersedia dalam masyarakat, baik bagi kepentingan diri sendiri, keluarga, orang lain, maupun makhluk hidup lainnya dan tidak untuk diperdagangkan. (Wikipedia).

Ada beberapa faktor yang mendorong kepuasan konsumen (Lupyoadi, 2001) diantaranya sebagai berikut :

1. Kualitas produk yaitu pelanggan akan merasa puas bila hasil mereka menunjukkan bahwa produk yang mereka gunakan berkualitas.
2. Kualitas pelayanan atau jasa yaitu pelanggan akan merasa puas bila mereka mendapatkan pelayanan yang baik atau sesuai dengan yang diharapkan.
3. Harga yaitu produk yang mempunyai kualitas yang sama tetapi menetapkan harga yang relatif murah akan memberikan nilai yang lebih tinggi kepada pelanggan.
4. Biaya yaitu pelanggan tidak perlu mengeluarkan biaya tambahan atau tidak perlu membuang waktu untuk mendapatkan suatu produk atau jasa cenderung puas terhadap produk atau jasa tersebut.

#### **2.2 Data Warehouse**

Data warehouse adalah data-data yang berorientasi subjek, terintegrasi, memiliki dimensi waktu, serta merupakan koleksi tetap (non-volatile), yang digunakan dalam mendukung proses pengambilan keputusan. Sedangkan data mining muncul setelah banyak dari pemilik data baik perorangan maupun organisasi mengalami penumpukan data yang telah terkumpul selama beberapa tahun, misalnya data pembelian, data penjualan, data nasabah, data transaksi, email dan sebagainya. Kemudian muncul pertanyaan dari pemilik data tersebut, apa yang harus dilakukan terhadap tumpukan data tersebut.

Data warehouse adalah database yang berisi data dari beberapa sistem operasional yang terintegrasi dan terstruktur sehingga dapat digunakan untuk

mendukung analisa dan proses pengambilan keputusan dalam bisnis. Data warehouse didesain untuk bisa melakukan query secara cepat. Informasi diturunkan dari data lain, dilakukan rolling up untuk dijadikan ringkasan, dilakukan operasi drilling down untuk mendapatkan informasi lebih detail, atau melihat pola yang menarik atau melihat trend (kecenderungan).

Ada empat tugas yang bisa dilakukan dengan adanya data warehouse, antara lain :

1. Pembuatan laporan

Pembuatan laporan merupakan salah satu kegunaan data warehouse yang paling umum dilakukan. Dengan menggunakan query sederhana didapatkan laporan perhari, perbulan, pertahun atau jangka waktu kapanpun yang ditentukan.

2. On-Line Analytical Processing (OLAP)

OLAP mendayagunakan konsep data multi dimensi dan memungkinkan para pemakai menganalisa data sampai mendetail, tanpa mengetikkan satupun perintah SQL. Hal ini dimungkinkan karena pada konsep multi dimensi, maka data yang berupa fakta yang sama bisa dilihat dengan menggunakan fungsi yang berbeda. Fasilitas lain yang ada pada software OLAP adalah fasilitas roll-up dan drill-down. Drill-down adalah kemampuan untuk melihat detail dai suatu informasi dan roll up adalah kebalikannya.

3. Data mining

Data mining merupakan proses untuk menggali pengetahuan dan informasi baru dari data yang berjumlah banyak pada data warehouse, dengan menggunakan kecerdasan buatan (Artificial Intelegence), statistic dan matematika. Data mining merupakan teknologi yang diharapkan dapatmenjembatani komunikasi antara data dan pemakainya.

4. Proses informasi executive

Data warehouse dapat membuat ringkasan informasi yang penting dengan tujuan membuat keputusan bisnis, tanpa harus menjelajahi keseluruhan data. Dengan menggunakan data warehouse segala laporan telah di ringkas dan dapat pula mengetahui segala rinciannya secara lengkap, sehingga

mempermudah proses pengambilan keputusan. Informasi dan data pada laporan data warehouse menjadi target informative bagi user.

### **2.3 Data Mining**

Data mining adalah suatu proses mengeksplorasi dan menganalisis data dalam jumlah besar baik secara otomatis maupun semi otomatis untuk mendapatkan suatu pola yang bermakna dari data (Nawawi, 2013).

Menurut (Goela, et al, 2012), Data Mining (DM) adalah proses menganalisis data dengan berbagai perspektif dan meringkasnya kedalam informasi yang berguna, dimana informasi tersebut bisa digunakan untuk meningkatkan pendapatan, memotong biaya atau keduanya.

Dan ada juga menurut (Turban, E., J.E. Aronson dan T.P. Liang. 2005), 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.

Tan (2006) mendefinisikan data mining sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam menggambil keputusan. Salah satu teknik yang dibuat dalam data mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar bisa mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan. Kebutuhan untuk prediksi juga dapat memanfaatkan teknik ini. Dalam data mining pengelompokan juga bisa dilakukan. Tujuannya adalah agar kita dapat mengetahui pola universal data-data yang ada. Anomali data transaksi juga perlu dideteksi untuk dapat mengetahui tindak lanjut berikutnya yang dapat diambil. Semua hal tersebut bertujuan mendukung kegiatan operasional perusahaan sehingga tujuan akhir perusahaan dapat tercapai. (Prasetyo, 2012)

### 2.3.1 Proses *Data mining*

Data mining sesungguhnya merupakan salah satu rangkaian dari proses pencarian pengetahuan pada database (*Knowledge Discovery in Database / KDD*). KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data. KDD adalah keseluruhan proses non-trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti. Serangkaian proses tersebut yang memiliki tahap sebagai berikut (Tan, 2004):

1. Pembersihan data dan integrasi data (*cleaning and integration*)

Proses ini digunakan untuk membuang data yang tidak konsisten dan bersifat *noise* dari data yang terdapat di berbagai basisdata yang mungkin berbeda format maupun platform yang kemudian diintegrasikan dalam satu database *data warehouse*.

2. Seleksi dan transformasi data (*selection and transformation*)

Data yang terdapat dalam database *datawarehouse* kemudian direduksi dengan berbagai teknik. Proses reduksi diperlukan untuk mendapatkan hasil yang lebih akurat dan mengurangi waktu komputasi terutama untuk masalah dengan skala besar (*large scale problem*). Beberapa cara seleksi, antara lain:

- a. *Sampling*, adalah seleksi subset representatif dari populasi data yang besar.
- b. *Denoising*, adalah proses menghilangkan *noise* dari data yang akan ditransformasikan.
- c. *Feature extraction*, adalah proses membuka spesifikasi data yang signifikan dalam konteks tertentu. Transformasi data diperlukan sebagai tahap *pre-processing*, dimana data yang diolah siap untuk ditambah. Beberapa cara transformasi, antara lain (Santosa, 2007):
- d. *Centering*, mengurangi setiap data dengan rata-rata dari setiap atribut yang ada.
- e. *Normalisation*, membagi setiap data yang di *centering* dengan standar deviasi dari atribut bersangkutan.
- f. *Scaling*, mengubah data sehingga berada dalam skala tertentu.

### 3. Penambangan data (*data mining*)

Data-data yang telah diseleksi dan ditransformasi ditambang dengan berbagai teknik. Proses data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan fungsi-fungsi tertentu. Fungsi atau algoritma dalam data mining sangat bervariasi. Pemilihan fungsi atau algoritma yang tepat sangat bergantung pada tujuan dan proses pencarian pengetahuan secara keseluruhan.

### 4. Evaluasi pola dan presentasi pengetahuan

Tahap ini merupakan bagian dari proses pencarian pengetahuan yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Langkah terakhir KDD adalah mempresentasikan pengetahuan dalam bentuk yang mudah dipahami oleh pengguna.

## 2.3.2 Fungsi-fungsi Data Mining

Fungsi-fungsi yang umum diterapkan dalam data mining (Daniel T.Larose, 2005):

1. *Assosiation*, adalah proses untuk menemukan aturan asosiatif antara suatu kombinasi item dalam suatu waktu dimana asosiasi terdiri dari berbagai macam diantaranya adalah Asosiasi analisa Biasa, Asosiasi Kuantitatif, Asosiasi Negatif, Asosiasi Positif, Asosiasi data diskrit, Asosiasi square pattern, Asosiasi korelasi.
2. *Secuence*, hampir sama dengan *association* bedanya *seccuencen* diterapkan lebih dari satu periode.
3. *Clustering*, adalah proses pengelompokan sejumlah data/obyek kedalam kelompok-kelompok data (klaster) sehingga setiap klaster akan berisi data yang saling mirip.
4. *Classification*, adalah proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui.
5. *Regretion*, adalah proses pemetaam data dalam suatu nilai prediksi.

6. *Forecasting*, adalah proses pengestimasian nilai prediksi berdasarkan pola-pola di dalam sekumpulan data.
7. *Solution*, adalah proses penemuan akar masalah dan *problem solving* dari persoalan bisnis yang dihadapi atau paling tidak sebagai informasi pendukung dalam pengambilan keputusan.

#### **2.4 Assosiation Rules**

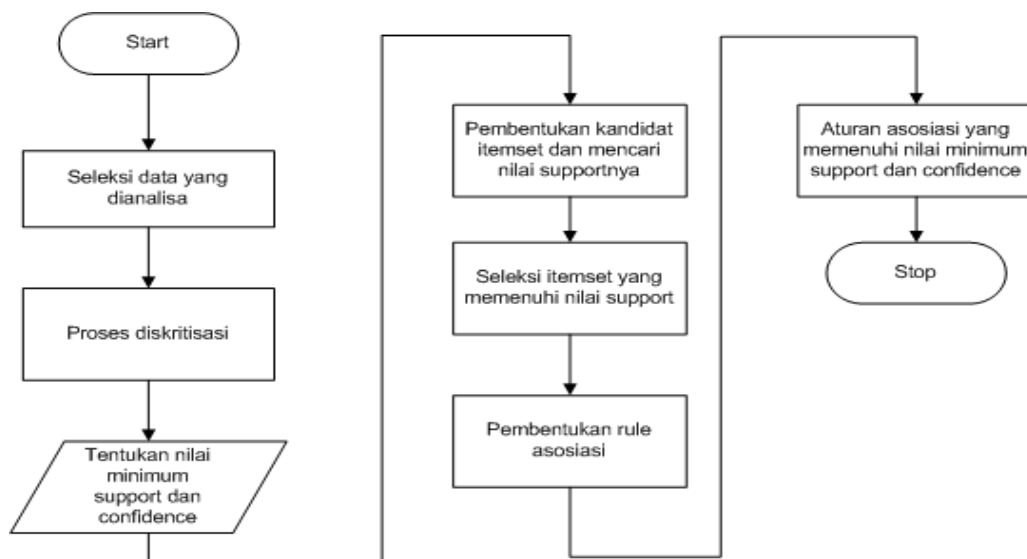
Analisis asosiasi berguna untuk menemukan hubungan penting yang tersembunyi di antara set data yang sangat besar. Hubungan yang sudah terbuka dipresentasikan dalam bentuk aturan asosiasi atau set aturan item yang sering muncul. Isu penting dalam analisis asosiasi adalah bagaimana cara menemukan pola tertentu dari data yang berjumlah sangat besar, yang membuat biaya komputasi menjadi sangat mahal. Isu penting lainnya adalah bahwa diantara pola-pola yang harus dipelajari dalam teknik analisis asosiasi, ada banyak pola yang akhirnya dibuang karena tidak termasuk dalam kriteria pola yang dicari. (Prasetyo, 2012:311).

Algoritma Apriori menggunakan pendekatan level yang baik dalam pembangkitan asosiasi, dimana setiap level berhubungan dengan sejumlah item yang dimiliki konsekuensi aturan. Awalnya, aturan dengan *confidence* yang tinggi yang mempunyai satu item diekstrak. Aturan ini kemudian digunakan untuk membangkitkan kandidat aturan yang baru. (Prasetyo, 2012:331).

Dalam algoritma apriori semakin kecil nilai minimum support, semakin banyak rule yang dihasilkan tetapi tidak semua rule yang dihasilkan valid, yang menunjukkan bahwa nilai minimum support berpengaruh pada pembentukan rule tetapi tidak banyak berpengaruh terhadap kevalidan rule. Semakin kecil nilai minimum confidence, semakin banyak rule yang dihasilkan dan besar kemungkinan rule tersebut valid, yang menunjukkan bahwa nilai minimum confidence berpengaruh pada kevalidan rule. Dari rule-rule yang memenuhi nilai confidence dihitung nilai lift atau interest factor rasionya, rule yang memiliki nilai lift rasio atau interest factor lebih dari 1 maka rule tersebut valid dan bisa dijadikan rekomendasi.

### 2.4.1 Tahapan *Association Rules*

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). Berikut flowchart Apriori pada gambar 2.1.



**Gambar 2.1** Flowchart Apriori

Dari gambar 2.1 dapat dijelaskan bahwa proses pertama adalah menyeleksi data sebagai data uji kemudian melakukan proses diskritisasi dimana merubah representasi atribut kategorikal menjadi biner pada setiap interval. Selanjutnya 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:

##### a. *Support* (dukungan)

*Support* merupakan suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item atau itemset dari keseluruhan transaksi.

Nilai *support* 1 *item* diperoleh dengan rumus 1 berikut :

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi mengandung A}}{\text{Jumlah transaksi}} \dots\dots\dots(2.1)$$

Nilai support 2 *item* diperoleh dengan rumus 2 berikut :

Sebagai contoh, support dari aturan "A=>B" adalah probabilitas kemunculan barang A dan B secara bersamaan.

$$\text{Support (A,B)} = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Jumlah transaksi}} \dots\dots\dots(2.2)$$

## 2. Pembentukan Aturan Asosiasi

### a. *Confidence* (tingkat kepercayaan)

Probabilitas kejadian beberapa barang dibeli bersamaan dimana salah satu produk sudah pasti dibeli. Sebagai contohnya, jika terdapat  $n$  transaksi dimana barang A dibeli dan ada  $m$  transaksi dimana barang A dan B dibeli secara bersamaan, maka confidence dari aturan *if A then B* adalah  $m/n$ .

Nilai confidence 2 item diperoleh dengan rumus 1 berikut :

$$\text{Confident (A} \rightarrow \text{B)} = \frac{\text{Support } A \cup B}{\text{Support A}} \dots\dots\dots(2.3)$$

3. Pada proses evaluasi sistem ini digunakan metode korelasi *Lift*. Dimana metode ini menguji korelasi dari rule yang telah dibentuk. Adapun korelasi *Lift* dapat dihitung dengan rumus seperti dibawah ini:

$$\text{Lift (X, Y)} = \frac{S(X \cup Y)}{S(X)P(Y)} \dots\dots\dots(2.4)$$

Dimana :

$\text{Lift (X, Y)} = \text{korelasi Lift X, Y}$

$S(X \cup Y) = \text{Jumlah kemunculan X dan Y dibagi dengan total transaksi}$

$S(X)P(Y) = \text{Jumlah kemunculan X dikali dengan jumlah kemunculan Y pada total transaksi}$

*Lift ratio* digunakan untuk mengevaluasi kuat tidaknya sebuah aturan asosiasi (Santoso, 2007). Apabila dari perhitungan tersebut menghasilkan nilai dibawah 1 maka terdapat korelasi *negatif*, untuk perhitungan yang menghasilkan



nilai diatas 1 maka terdapat korelasi *positif*. Namun apabila menghasilkan nilai sama dengan 1 maka tidak ada korelasi antara X dan Y.

Algoritma *Apriori*:

1. 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.

2. 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.

3. Tetapkan pola frekuensi tinggi.

Pola frekuensi tinggi yang memuat k item atau kitemset ditetapkan dari kandidat k-itemset yang supportnya lebih besar dari minimum *support*.

Untuk lebih memahami proses algoritmaApriori maka berikut ini akan diberikan ilustrasi penggunaan algoritma Apriori. Ilustrasi algoritma Apriori dapat dilihat pada proses berikut :

Misalkan :

TID	Itemset
1	A.Telur, C.Minyak, D.Gula
2	B.Beras, C.Minyak, E.Susu
3	A.Telur, B.Beras, C.Minyak, E.Susu
4	B.Beras, E.Susu

Keterangan : A : Telur B : Beras C : Minyak D :Gula E : Susu

Misalkan diinginkan minsup : 50% (2 dari 4 transaksi)

Langkah 1:

$$L1=\{\text{large 1-itemset}\}$$

Itemset	Support
A	50%
B	75%
C	75%
D	25%
E	75%

Langkah 2: Mencari kandidat itemset untuk L2:

2.1 : Gabungkan itemset pada L1 { A B, A C, A D, A E, B C, B D, B E, C D, C E, D E}

2.2 : Hapus yang tidak ada dalam itemset

Itemset { B D, DE} dihapus karena tidak ada dalam itemset

Langkah 3 :

Hitung support dari setiap kandidat itemset

Itemset	Support
A B	25 %
A C	50 %
A D	25 %
A E	25%
B C	50%
B E	75%
C D	25%
C E	50%

Langkah 4:

L2 { large 2-itemset }

Itemset	Support
A C	50 %
B C	50%
B E	75%
C E	50%

Langkah 5 : Ulangi langkah 2-4

5.1 :Gabungkan itemset pada L2 & L2:

Itemset	Hasil Gabungan (3 itemset)
A C + B C	A C B

A C + B E	A C B, A C E, A B E
A C + C E	A C E
B C + B E	B C E
B C + C E	B C E
B E + C E	B C E

5. 2 : Hapus yang tidak ada dalam itemset : { A C E }

Langkah 6 : Hitung support dari setiap kandidat itemset L3

Itemset	Support
A B C	25 %
A B E	25 %
B C E	50 %

Langkah 7 : L3 { large 3-itemset } { B C E }

Langkah 8 : STOP karena sudah tidak ada lagi kandidat untuk 4-itemset.

Dari hasil – hasil diatas hasil akhir sebagai berikut:

L1		L2		L3	
A	50%	A C	50%	B C E	50%
B	75%	B C	50%		
C	75%	B E	75%		
D	25%	C E	50%		
E	75%				

Untuk mencari aturan asosiasi diperlukan juga minconf

Misal minconf : 75 %, aturan asosiasi yang mungkin terbentuk:

Aturan ( $X \rightarrow Y$ )	Sup( $X \cup Y$ )	Sup(X)	Confidence
B C $\rightarrow$ E	50%	50%	100%
B E $\rightarrow$ C	50%	75%	66.67%

$C \rightarrow E \rightarrow B$	50%	50%	100%
$A \rightarrow C$	50%	50%	100 %
$C \rightarrow A$	50%	75%	66.67%
$B \rightarrow C$	50%	75%	66.67%
$C \rightarrow B$	50%	75%	66.67%
$B \rightarrow E$	75%	75%	100%
$E \rightarrow B$	75%	75%	100%
$C \rightarrow E$	50%	75%	66.67%
$E \rightarrow C$	50%	75%	66.67%

## 2.5 Analisis Dasar Diskritisasi

Diskritisasi adalah metode dasar pendekatan untuk mengatur keberlanjutan sebuah atribut. Pendekatan tersebut dilakukan dengan cara mengubah representasi atribut kategorikal menjadi biner pada setiap interval. Nilai untuk sebuah item adalah 1 jika item tersebut ada dalam atribut, atau 0 jika tidak ada dalam atribut. (Kumar, Tan, & Steinbach, 2006).

Beberapa data memiliki format yang kompleks dimana terdapat penggabungan tipe data numerik dan diskrit. Data diskrit lebih dekat ke tingkat pengetahuan representasi manusia dan terkadang lebih efisien. (Yul, 2010). Diskritisasi merupakan peralihan dari data kontinu menjadi data diskrit (kategori) (Gorunescu, 2011). Diskritisasi juga dapat digunakan untuk mengubah atribut dengan nilai numerik menjadi nilai nominal / kategorikal (Sullivan, 2014). Terdapat dua bentuk diskritisasi yaitu diskritisasi yang tidak diawasi (unsupervised) dan diskritisasi yang diawasi (supervised). Diskritisasi yang tidak diawasi diperuntukkan untuk tugas data mining clustering sedangkan diskritisasi yang diawasi diperuntukkan untuk tugas data mining regresi dan klasifikasi (Yul, 2010). Untuk contoh dasar proses diskritisasi dapat dilihat pada **Tabel 2.1**.

**Tabel 2.1** Proses Dasar Diskritisasi

Gender	Male	Female
Male	1	0
Female	0	1
Female	0	1
Male	1	0
Male	1	0
Female	0	1

## 2.6 Penelitian Sebelumnya

Pada penelitian sebelumnya yang menggunakan metode Apriori adalah penelitian yang berjudul “*Analisis profil pasien rawat inap berdasarkan latar belakang atribut pasien dengan menggunakan metode apriori (association rules)*” oleh M. Hussein Al Habsyi. Adapun data yang diambil dalam penelitian ini adalah data pasien rawat inap rumah sakit Muhammadiyah Gresik Bulan Mei – Agustus 2014. Atribut yang digunakan adalah Jenis Kelamin, alamat, status pernikahan, sumber pembiayaan, pendidikan terakhir, pekerjaan dan umur dan outputnya adalah berupa informasi tentang asosiasi latar belakang pasien rawat inap sebagai strategi promosi oleh pihak rumah sakit Muhammadiyah Gresik. Hasil menunjukkan bahwa tingkat akurasi Rule yang mempunyai tingkat confidence tertinggi pada masing-masing percobaan dengan rule 3 itemset adalah {Perempuan, Ibu Rumah Tangga} → {Menikah} dengan nilai *confidence* sebesar 94.87% dan 92.50%. sedangkan rule 2 itemset tertinggi adalah {Ibu Rumah Tangga} → {Perempuan} dan {Ibu Rumah Tangga} → {Menikah} dengan nilai *confidence* sebesar 95.24% dan 92.86%. Sedangkan Ada 6 rule pada pengujian periode bulan Mei – Agustus 2014 dengan minsup 25% dan minconf 80% dan 19 rule dengan minsup 15% dan 80% dan memiliki korelasi atau hubungan yang *positif* dari pasangan variabel pasien rawat inap dengan nilai korelasi  $\geq 1$  sedangkan pada pengujian dengan minsup 30% dan minconf 70% terdapat 8 rule yang masing-masing memiliki 5 korelasi atau hubungan yang *positif* dan terdapat 3 korelasi atau hubungan *negatif* dengan nilai korelasi  $\leq 1$ . Rule yang memiliki

korelasi negatif maka tidak dapat dijadikan sebagai rekomendasi strategi promosi pasien.

Penelitian selanjutnya yang dikemukakan oleh Rohman Dijaya (2013) dengan judul *APLIKASI ANALISA POLA DIAGNOSA PENYAKIT PADA REKAM MEDIS ELEKTRONIK DENGAN METODE APRIORI* pembentukan confidence pada asosiasi dari pasangan obat dapat dibentuk dengan metode apriori melalui query mysql yang dapat membentuk pasangan obat yang sering muncul secara bersamaan adalah dengan pasangan obat 85,888 dan 1720 atau Gliseril Guayakalot tablet 100 mg, Klorfen M 4 mg dan pamset Tab 500 mg. hasil keluaran output perangkat lunak berupa kaidah asosiasi yang didasarkan pada nilai minimum support yaitu yaitu nilai prosentase minimal yang menunjukkan seberapa besar tingkat dominasi sumber obat (item) atau kumpulan item (itemset) terhadap keseluruhan data transaksi yang dianalisa dari nilai minimum confidence.

Penelitian lain dilakukan oleh mohammad iljas mengenai *Rancang bangun perangkat lunak analisa keranjang belanja pasar dengan metode Apriori*. Dalam penelitiannya analisa keranjang pasar dengan metode apriori sebagai salah satu teknik data mining digunakan untuk menggali pola kecenderungan kemunculan barang secara bersamaan yang dilakukan oleh para pembeli. Dengan data transaksi yang besar maka sql dapat digunakan sebagai solusi yang tepat untuk meningkatkan kecepatan proses pembentukan frequen itemset, pembentukan kandidat 3 itemset dari 2 itemset, perhitungan nilai support dan confidence serta dapat melakukan pembentukan kaidah asosiasi dari 2-itemset dan 3-itemset. Berdasarkan grafik analisa data transaksi order detail pada database northwide SQL yang telah diubah ke database MySQL, nilai support tertinggi yang dapat digunakan untuk membentuk 2 itemset sebesar 1% dengan kemunculan sebanyak 8 transaksi berbanding terhadap seluruh transaksi yang berjumlah 330 transaksi.