

Hak Cipta Diindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB II

### LANDASAN TEORI

#### 2.1 *Data Mining*

Pada bagian ini akan menjelaskan tentang *data mining*, yang meliputi pengertian, proses, dan pengelompokan pada *data mining*

##### 2.1.1 *Pengertian Data Mining*

*Data mining* adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam kumpulan data. *Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar. (Kusrini *et al*, 2010 dalam Pane, 2013)

Selain definisi di atas ada definisi lain juga seperti, “*data mining* adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi (*database*) yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui. *Data mining* sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu *data mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, *database* dan statistik” (Wirdasari *et al*, 2011).

Berdasarkan beberapa definisi tersebut dapat ditarik kesimpulan bahwa *data mining* adalah suatu teknik untuk menemukan suatu informasi yang tersembunyi dalam kumpulan data yang sangat besar sehingga mampu mengidentifikasi informasi yang berguna yang sebelumnya tidak diketahui untuk membuat keputusan yang kritis, terutama dalam strategi.

Hak Cipta Dilindungi Undang-Undang

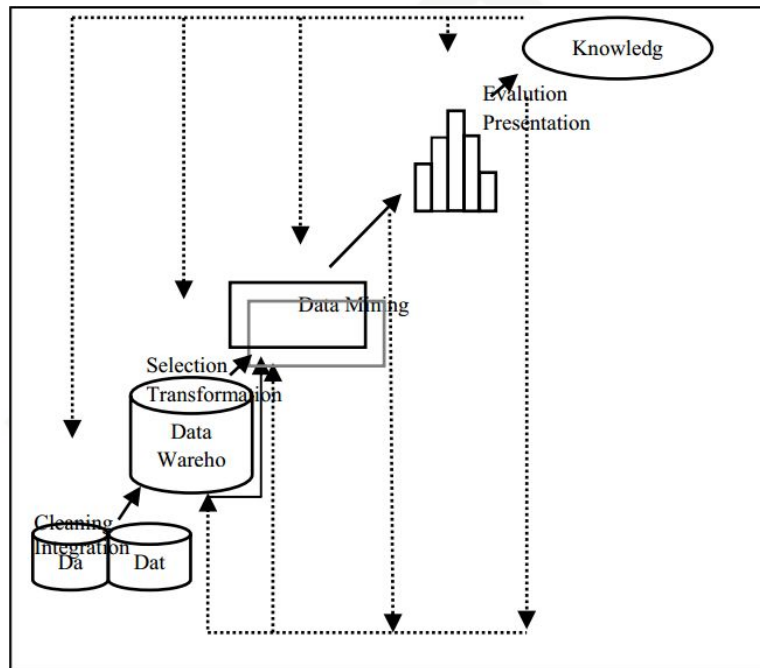
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 2.1.2 Proses Data Mining

Istilah *data mining* sering disebut juga *Knowledge Discovery in Database* (KDD). *Knowledge Discovery in Database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. (Kusrini *et al*, 2009 dalam Budiono *et al*, 2014).



Gambar 2.1 Proses Data Mining

(Sumber : Noranita *et al*, 2012 dalam Buulolo, 2013)

Tahap-tahap data mining ada 6 yaitu :

- 1) Pembersihan data (data cleaning)

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi

#### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

performasi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

#### 2) Integrasi data (data integration)

Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

#### 3) Seleksi Data (Data Selection)

Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

#### 4) Transformasi data (Data Transformation)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

#### 5) Proses mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

#### 6) Evaluasi pola (pattern evaluation)

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

#### 7) Presentasi pengetahuan (knowledge presentation)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining (Han, 2006 dalam R Siburian, 2014).

### 2.1.3 Pengelompokan *Data Mining*

*Data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat di lakukan, yaitu (Kusrini *et al*, 2009 dalam Buulolo, 2013).

#### 1) Deskripsi

Terkadang peneliti dan analisis 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.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 2) Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik dari pada kearah 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 hasik akan ada dimasa mendatang. Contoh prediksi bisnis dan penelitian adalah:

- a. Prediksi harga beras dalam tiga bulan yang akan datang.
- b. Prediksi persentasi 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:

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

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 5) Pengklusteran

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:

- a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari satu suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku financial dalam baik dan mencurigakan.
- c. 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:

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

Hak Cipta Dilindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Diarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 2.2 Association Rules Mining

Association rules mining adalah suatu prosedur untuk mencari hubungan antar item suatu dataset yang telah ditentukan . Association rules mining mencari dan menemukan hubungan antar item yang ada pada suatu dataset. Penerapan data mining dengan aturan asosiasi bertujuan menemukan informasi item-item yang saling berhubungan dalam bentuk aturan/rule. Aturan asosiasi adalah teknik data mining untuk menemukan aturan asosiasi antar suatu kombinasi item. (Han, J, 2006 dalam Anggreani *et al*, 2013).

Metode ini dapat membantu mengenali pola-pola tertentu di dalam kumpulan data yang besar. Dalam association rules, suatu kelompok item dinamakan itemset. Support dari itemset X adalah persentase transaksi di D yang mengandung X, biasa ditulis dengan  $\text{supp}(X)$ . Pencarian association rules dilakukan melalui dua tahap yaitu pencarian frequent itemset dan penyusunan rules. Jika support suatu itemset lebih besar atau sama dengan minimum support  $\sigma$ , maka itemset tersebut dapat dikatakan sebagai frequent itemset atau frequent pattern; yang tidak memenuhi dinamakan infrequent. Confidence adalah nilai ukuran seberapa besar valid tidaknya suatu association rules. Confidence suatu rule R ( $X \Rightarrow Y$ ) adalah proporsi dari semua transaksi yang mengandung baik X maupun Y dengan yang mengandung X, biasa ditulis sebagai  $\text{conf}(R)$ . Sebuah association rule dengan confidence sama atau lebih besar dari minimum confidence  $\gamma$  dapat dikatakan sebagai valid association rule (Agrawal *et al*, 1994 dalam Ruldeviyani *et al*, 2008).

## 2.3 Pengertian Algoritma Apriori

Algoritma Apriori adalah suatu algoritma dasar yang diusulkan oleh Agrawal & Srikant pada tahun 1994 untuk penentuan frequent itemsets untuk aturan asosiasi boolean. Algoritma a priori termasuk jenis aturan asosiasi pada data mining. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut affinity analysis atau market basket analysis (Sari, 2013)

Untuk penerapan algoritma Apriori, secara umum dibutuhkan struktur data untuk menyimpan candidate frequent itemset untuk suatu iterasi ke k dan untuk

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

menyimpan frequent itemset yang dihasilkan. Ketika membaca tiap item dari seluruh transaksi, selain mendapatkan item-item baru juga dilakukan perhitungan nilai support item-item yang sudah ditemukan, sehingga untuk mendapatkan candidate 1-itemset beserta nilai support-nya cukup membutuhkan satu kali pembacaan data (Ruldeviyani *et al*, 2008).

### 2.3.1 Kelebihan Algoritma Apriori

Algoritma Apriori yang akan digunakan pada penelitian ini, mempunyai beberapa kelebihan (Santoso, 2003):

- a. Menggunakan format data *vertical tid-list* yang mengasosiasikan itemset dengan transaksi yang terjadi.
- b. Menggunakan pendekatan apriori untuk mencari maksimal *frequent* itemset.
- c. Membutuhkan hanya sedikit pembacaan database, dan meminimalkan penggunaan I/O.

### 2.3.2 Tahap-Tahap Algoritma Apriori

Algoritma apriori dibagi menjadi beberapa tahap yang disebut narasi atau pass (Setiawan, 2009 dalam Buulolo, 2013):

- 1) Pembentukan kandidat itemset, kandidat k-itemset dibentuk dari kombinasi (k-1) itemset yang didapat dari iterasi sebelumnya. Satu cara dari algoritma apriori adalah adanya pemangkasan kandidat k-itemset yang subset-nya yang berisi k- 1 item tidak termasuk dalam pola frekuensi tinggi dengan panjang k-1.
- 2) Penghitungan support dari tiap kandidat kitemset. 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.



Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- 3) 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.
- 4) Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Bila tidak, maka k ditambah satu dan kembali bagian 1

**2.3.3 Analisis Pola Frekuensi Tinggi (Support)**

Mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam basis data. Nilai support sebuah item diperoleh dengan menggunakan rumus berikut:

$$Support(A) = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} \dots\dots\dots(2.1)$$

Nilai support dari 2 item diperoleh dengan menggunakan rumus:

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

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

(Santoso, B, 2007 dalam Yanto *et al*, 2015).

**2.3.4 Pembentukan Aturan Asosiasi (Confidence)**

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

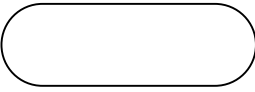
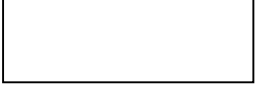
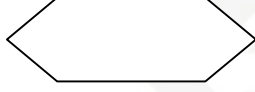
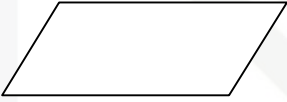
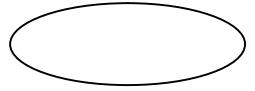




$$Confidence = P(B|A) = \frac{\Sigma \text{Transaksi mengandung A dan B}}{\Sigma \text{Transaksi mengandung A}} \dots\dots\dots(2.3)$$

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan Support × Confidence. Aturan diambil sebanyak n aturan yang memiliki hasil terbesar (Santoso, B., 2007 dalam Robi Yanto *et al*, 2015).

## 2.4 Flowchart

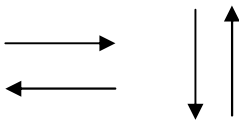

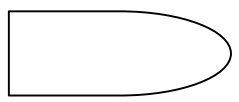
*Flow chart* adalah suatu diagram alir yang menggunakan simbol atau tanda untuk menyelesaikan masalah. Adapun simbol-simbol yang sering digunakan dalam *flow chart* (Siallagan, 2009 dalam Ramadhan, 2014) yaitu:

Tabel 2.1. Simbol-simbol *Flowchart*

Simbol Flow Chart	Fungsi Flow Chart
	Terminal menyatakan awal atau akhir dari suatu algoritma.
	Menyatakan Proses.
	Proses yang terdefinisi atau sub program.
	Persiapan yang digunakan untuk memberi nilai awal suatu besaran.
	Menyatakan masukan dan keluaran ( <i>input/output</i> )
	Menyatakan penyambung ke simbol lain dalam satu halaman.
	Menyatakan penyambung ke halaman lainnya.
	Menyatakan pencetakan (dokumen) pada kertas.
	Menyatakan <i>decision</i> (keputusan) yang digunakan untuk penyeleksian kondisi di dalam program.
	Menyatakan media penyimpanan drum magnetik.
	Menyatakan input/output menggunakan disket.
	Menyatakan input/output dari kartu plong.

Hak Cipta Diindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

	Menyatakan arah alir pekerjaan (proses).
	Multidocument (banyak dokumen)
	Delay (penundaan atau kelambatan).

(Sumber: Siallagan, 2009 dalam Ramadhan, 2014)

## 2.5 Basis Data (*Database*)

Menurut Fathansyah (1999), basis data terdiri atas dua kata, yaitu Basis dan Data. Basis dapat diartikan sebagai markas atau gudang, tempat bersarang/berkumpul. Sedangkan data adalah representasi fakta dunia nyata yang mewakili suatu objek seperti manusia (siswa, pegawai, pembeli, pelanggan), barang, hewan, peristiwa, konsep, keadaan, dan sebagainya.

Abdul Kadir (2003) mengatakan bahwa basis data (*database*) adalah suatu pengorganisasian sekumpulan data yang saling terkait sehingga memudahkan aktivitas untuk memperoleh informasi. Basis data dimaksudkan untuk mengatasi problem pada sistem yang memakai pendekatan berbasis berkas.

Secara lebih lengkap, tujuan basis data adalah sebagai berikut :

- 1) Kecepatan dan kemudahan (*speed*)
- 2) Efisiensi ruang penyimpanan (*space*)
- 3) Keakuratan (*accuracy*)
- 4) Ketersediaan (*availability*)
- 5) Kelengkapan (*completeness*)
- 6) Keamanan (*security*)
- 7) Kebersamaan pemakai (*shareability*)

## 2.6 MySQL

MySQL adalah sebuah aplikasi *Relational Database Management Server* (RDBMS) yang sangat cepat dan kokoh serta bersifat *open source*. MySQL

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

merupakan salah satu jenis *database server* yang banyak digunakan di dunia maya, yang menggunakan *SQL* sebagai bahasa dasar untuk mengakses database. MySQL dapat digunakan pada berbagai *platform* sistem operasi. Keunggulan MySQL dalam mengolah database adalah:

- 1) Kecepatan
- 2) Mudah digunakan
- 3) *Open Source*
- 4) Kapabilitas
- 5) Biaya murah
- 6) Keamanan
- 7) Lintas *Platform*
- 8) Minim ”bug”

### 2.7 *Hypertext Pre-Processor (PHP)*

PHP singkatan dari Hypertext Preprocessor yang digunakan sebagai bahasa script *server-side* yang cukup handal, yang akan disatukan dengan *HTML* dan berada di *server*. Artinya, sintaks dan perintah yang diberikan akan sepenuhnya dijalankan di *server* sebelum dikirim ke komputer klien. Seluruh aplikasi yang berbasis *web* dapat dibuat menggunakan *PHP*. Salah satu kelebihan *PHP* adalah kemampuan untuk dapat melakukan koneksi dengan berbagai database, seperti MySQL, PostgreSQL, Access, dll. Selain itu *PHP* juga bersifat *open source*.

### 2.8 *Blackbox Testing*

Teknik pengujian *black-box* berfokus pada domain informasi dari perangkat lunak, dengan melakukan *test case* dengan mempartisi domain input dari suatu program dengan cara yang memberikan cakupan pengujian yang mendalam.

*Black-box* Merupakan metode pengujian perangkat lunak yang dilakukan hanya mengamati hasil eksekusi data uji dan memeriksa fungsional dari perangkat lunak. Cakupan pengujian yang dilakukan pada *blackbox testing* adalah perih





pengujian *interface* dan *form validation*. Pengujian *interface* adalah pengujian yang dilakukan secara langsung terhadap desain interface yang dibuat pada system. Tujuan yang diharapkan dalam melakukan *blackbox testing* adalah dapat membuat desain yang sesuai dengan kebutuhan perusahaan.

## 2.9 Obat

Pengertian obat adalah bahan yang digunakan untuk mengurangi, menghilangkan, atau menyembuhkan penyakit (Salim, 2013 dalam Ramadhan, 2014).

## 2.10 Apotek

Pengertian Apotek adalah tempat pengambilan atau pembelian obat berdasarkan resep dokter (Salim, 2013 dalam Ramadhan, 2014).

Menurut PP No.51 Tahun 2009, apotek adalah saran pelayanan kefarmasian tempat dilakukan praktik kefarmasian oleh apoteker. Pelayanan kefarmasian adalah suatu layanan langsung dan bertanggung jawab kepada pasien yang berkaitan dengan sediaan farmasi dengan maksud mencapai hasil yang pasti untuk meningkatkan mutu kehidupan pasien. Pekerjaan kefarmasian yang dilakukan meliputi pembuatan termasuk pengendalian mutu sediaan farmasi, pengamanan, pengadaan, penyimpanan, dan pendistribusian atau penyaluran obat, pengelolaan obat, pelayanan obat atau resep dokter, pelayanan informasi obat, serta pengembangan obat, bahan obat, obat tradisional dan kosmetika (Supriyati, 2014).

## 2.11 Riset Terdahulu

Pemakaian algoritma apriori telah digunakan pada peneliti sebelumnya yaitu (Buulolo, 2013) melakukan Implementasi Algoritma Apriori Pada Sistem Persediaan Obat Studi Kasus: Apotek Rumah Sakit Estomihi Medan dengan pengujian aplikasi Tanagra menghasilkan pola dan kombinasi *itemset* dan *rules* sebagai pengetahuan dan informasi penting dari data penjualan (data obat keluar). Hasil dari penelitian tersebut menemukan pola kombinasi paling tinggi dengan *support* 22% dan *confidence* 60%.

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
    - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
    - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
  2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Penelitian terkait lainnya adalah Aplikasi *Data Mining* Analisa Data Transaksi Penjualan Obat Menggunakan Algoritma *Apriori* Studi Kasus di Apotek Setya Sehat Semarang, yang dilakukan oleh Anggraini *et al*, 2013) Pada pengujian data yang digunakan adalah data transaksi penjualan obat bulan Januari, Februari, dan Maret tahun 2012, Berdasarkan pengujian yang dilakukan hasil dari aplikasi *data mining* yang dibangun menggunakan proses sekuensial linier dengan baha pmrograman PHP dan basis data MySQL. Nilai minimum support 6 dan minimum confidance 40% merupakan nilai yang dapat menghasilkan aturan paling kuat pada bulan Februari(1 aturan) dan Maret (2 aturan), sedangkan nilai minimum support 6 dan minimum confidance 50% menghasilkan aturan paling kuat pada bulan Januari (3 aturan).

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengummumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.