

BAB II

TINJAUAN PUSTAKA

2.1 Teori Dasar

2.1.1 Analisis sistem

Menurut Mulyadi, analisis sistem adalah sebuah tahapan dalam pengembangan sistem yang akan menghasilkan berbagai dokumen yang menyajikan rencana pekerjaan yang akan dilaksanakan untuk mengembangkan sistem tersebut (Al Fatta, 2015).

Menurut McLeod yang diterjemahkan oleh Teguh, analisis sistem adalah penelitian atas sistem yang telah ada dengan tujuan untuk merancang sistem yang baru atau diperbarui (Hendarti & Haryanto, 2009).

Berdasarkan pendapat-pendapat diatas dapat disimpulkan bahwa analisis sistem adalah proses penelitian terhadap sistem yang sudah ada sebelumnya dengan tujuan untuk merancang sistem yang baru.

2.1.1.1 Langkah-langkah dasar analisis sistem

Menurut (Al Fatta, 2015) ada beberapa langkah-langkah dasar dari analisis sistem sebagai berikut :

1. *Identify*, yaitu mengidentifikasi masalah.
2. *Understand*, yaitu memahami kerja sistem yang ada.
3. *Analyze*, yaitu menganalisa sistem.

2.1.1.2 Tahap tahap analisis sistem

Menurut McLeod (Hendarti & Haryanto, 2009) tahap-tahap analisis sistem adalah sebagai berikut:

1. Mengumpulkan penelitian sistem.
2. Mengorganisasikan tim proyek.
3. Mendefinisikan kebutuhan sistem informasi.
4. Mendefinisikan kriteria kinerja sistem.
5. Menyiapkan usul rancangan.
6. Menyetujui atau menolak rancangan sistem.

2.1.2 Sistem

Menurut Sutabri, sistem adalah suatu kumpulan atau himpunan dari unsur, komponen, atau variabel yang terorganisasi, saling berinteraksi, saling tergantung satu sama lain dan terpadu (Ermatita, 2016).

Menurut sugiyanto *et al*, sistem adalah suatu susunan yang teratur dari kegiatan-kegiatan yang saling bergantung dan prosedur-prosedur yang saling

berhubungan yang melaksanakan dan mempermudah kegiatan-kegiatan utama organisasi (sugiyanto, sukadi, 2013).

Berdasarkan pendapat-pendapat diatas dapat disimpulkan bahwa sistem adalah kumpulan dari berbagai kegiatan yang saling berhubungan dalam melaksanakan suatu kegiatan organisasi untuk mencapai suatu tujuan.

2.1.2.1 Karakteristik Sistem

Menurut Sutabri (Ermatita, 2016), karakteristik dari sistem adalah sebagai berikut:

1. Komponen Sistem (*Components*)

Suatu sistem terdiri dari sejumlah komponen yang saling berinteraksi yang bekerja sama membentuk satu kesatuan. Komponen-komponen sistem tersebut dapat berupa suatu bentuk subsistem.

2. Batasan Sistem (*Boundary*)

Ruang lingkup sistem adalah daerah yang membatasi antara sistem dengan sistem lainnya atau sistem dengan lingkungan luarnya. Batasan sistem ini memungkinkan suatu sistem dipandang sebagai satu kesatuan yang tidak dapat dipisah-pisahkan.

3. Lingkungan Luar Sistem (*Environtment*)

Lingkungan luar sistem adalah bentuk apapun yang ada di luar ruang lingkup atau batasan sistem yang mempengaruhi operasi sistem tersebut.

4. Penghubung Sistem (*Interface*)

Penghubung sistem atau *interface* adalah media yang menghubungkan sistem dengan subsistem yang lain.

5. Masukan Sistem (*Input*)

Energi yang dimasukkan ke dalam sistem disebut masukan sistem, yang dapat berupa pemeliharaan (*maintenance input*) dan sinyal (*signal input*).

6. Keluaran Sistem (*Output*)

Keluaran ini adalah masukan bagi subsistem yang lain, seperti contoh sistem informasi, keluaran yang dihasilkan adalah informasi dimana informasi ini dapat digunakan sebagai masukan untuk pengambilan keputusan atau hal-hal lain yang adalah *input* bagi subsistem lain.

7. Pengolah Sistem (*Process*)

Suatu sistem dapat mempunyai suatu proses yang akan mengubah masukan menjadi keluaran.

8. Sasaran Sistem (*Objective*)

Suatu sistem memiliki tujuan dan sasaran yang pasti dan bersifat *deterministic*.

2.1.3 Penjualan

Menurut Basu swasta, penjualan adalah suatu sistem keseluruhan dari kegiatan usaha yang ditujukan untuk merencanakan, menentukan harga, mempromosikan dan mendistribusikan barang, jasa, ide kepada pasar sasaran agar dapat mencapai tujuan organisasi (sugiyanto, sukadi, 2013).

Menurut Winardi, penjualan adalah berkumpulnya seorang pembeli dan penjual dengan tujuan melaksanakan tukar menukar barang dan jasa berdasarkan pertimbangan yang berharga misalnya pertimbangan uang (Anthony et al., 2017).

Berdasarkan pendapat-pendapat diatas dapat disimpulkan bahwa, penjualan adalah kegiatan usaha dengan mempertemukan pembeli dan penjual untuk menjual barang dagang kepasar dengan melaksanakan tukar menukar barang guna memperoleh keuntungan agar dapat mencapai tujuan organisasi.

2.1.3.1 Klasifikasi Penjualan

Menurut Martin S, Colleran G (Anthony et al., 2017), bila diidentifikasi berdasarkan perusahaannya maka penjualan dapat dikategorikan menjadi beberapa jenis, yaitu:

1. Penjualan langsung dimana penjualan ini adalah dengan cara mengambil barang dari *supplier* kemudian secara langsung dikirim ke *customer*.
2. Penjualan stok gudang dimana penjualan ini adalah dengan cara menjual barang dari stok yang ada di gudang.
3. Penjualan kombinasi adalah penjualan dimana dengan mengambil sebagian barang dari *supplier* serta sebagian dari stok yang ada di gudang.

Dilihat dari pengertian tersebut maka suatu proses penjualan haruslah memberikan keuntungan bagi suatu usaha yang sedang berlangsung, karena itu diperlukan sebuah sistem informasi penjualan sehingga proses penjualan yang dilakukan dapat berjalan dengan efektif serta efisien.

Menurut Basu swastha dalam Anthony et al., 2017, jika melihat dari jenis dan bentuknya penjualan dapat dibedakan sebagai berikut:

1. *Trade selling* yaitu penjualan yang dapat terjadi jika antara produsen dan pedagang mengizinkan pengecer untuk memperbaiki distribusi produk mereka.
2. *Missionary selling*, yaitu proses untuk meningkatkan penjualan dengan cara mendorong para pembeli agar membeli barang-barang dari penyalur.
3. *Technical selling* yaitu proses untuk meningkatkan penjualan dengan cara memberikan saran serta nasihat kepada para pembeli barang dan jasa.
4. *New businiess selling* adalah usaha untuk membuka transaksi baru dengan calon pembeli seperti yang dilakukan perusahaan asuransi.
5. *Responsive selling* adalah dimana setiap tenaga penjual dapat memberikan reaksi terhadap permintaan serta pembeli melalui *route driving and retailing*.

Menurut Basu swastha dalam Anthony et al., 2017, penjualan dapat dibedakan berdasarkan bentuknya antara lain:

1. Penjualan Tunai/Cash adalah penjualan bersifat *cash and carry* yang mana penjualan dilakukan setelah terdapat kesepakatan harga antara penjual dengan pembeli, lalu pembeli dapat membayar secara langsung dan barang dapat langsung dimiliki.

2. Penjualan Kredit/*Non-Cash* yaitu penjualan *non-cash*, dengan memberikan tenggat waktu tertentu, biasanya diatas satu bulan.
3. Penjualan Tender adalah penjualan yang dilaksanakan melalui prosedur tender yang digunakan untuk memenuhi permintaan pihak pembeli yang membuka tender.
4. Penjualan Ekspor yaitu penjualan yang dilaksanakan dengan pihak pembeli dimana mengimpor barang dari luar negeri yang biasanya dengan menggunakan *letter of credit*.
5. Penjualan Konsinyasi yaitu penjualan barang dengan cara menitipkan kepada pembeli yang juga berperan sebagai penjual dimana jika barang yang dititipkan tersebut tidak terjual maka barang akan dikembalikan lagi kepada penjual.
6. Penjualan Grosir yaitu penjualan yang dilakukan tidak langsung kepada pembeli, tetapi melalui perantara pedagang.

2.1.4 Algoritma apriori

Menurut Heroe Santoso *et al*, algoritma apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule* (Santoso, Hariyadi, & Prayitno, 2016).

Menurut Kennedi Tampubolon *et al*, 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* (Tampubolon et al., 2013).

Berdasarkan pendapat-pendapat diatas dapat disimpulkan bahwa algoritma apriori adalah algoritma dasar untuk melakukan pencarian *frequent itemsets* untuk aturan asosiasi dalam menemukan suatu kombinasi item dengan menggunakan data mining.

2.1.4.1 Tahapan algoritma

Menurut Devi dinda setiawan (Tampubolon et al., 2013) algoritma apriori dibagi menjadi beberapa tahap yang disebut narasi atau pass.

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 k- itemset di tetapkan dari kandidat k- itemset yang supportnya lebih besar dari minimum support.

4. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan.

2.1.4.2 Analisis asosiasi

Menurut Amirudin *et al* dalam Tampubolon et al., 2013 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.

Menurut Seni Susanto dan Dedy Suryadi dalam Tampubolon et al., 2013 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...” adalah pengetahuan yang dihasilkan dari fungsi *Aturan Asosiasi* (Tampubolon et al., 2013).

2.1.4.3 Konsep aturan asosiasi

Aturan asosiasi data mining dulu diperkenalkan oleh Agarwal dan digunakan untuk analisis keranjang pasar. Tugas aturan asosiasi data mining adalah untuk menghasilkan semua aturan asosiasi yang memiliki *support* dan *confidence* lebih besar daripada yang ditetapkan pengguna *minimum support* (minsup) dan *minimum confidence* (minconf) masing-masing (Muralidhar & Pattabiraman, 2015).

Menurut (Santoso et al., 2016) metodologi dasar analisis asosiasi dibagi menjadi dua bagian sebagai berikut.

1. Analisis Pola Frekuensi Tinggi

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

$$\text{Support A} = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total transaksi}} \quad \text{Rumus 2.1 Nilai support 1 item}$$

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

$$\text{Support (A,B)} = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \quad \text{Rumus 2.2 Nilai support 2 item}$$

2. Pembentukan Aturan Asosiasi

Pembentukan Aturan Asosiasi Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi 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.

$$\text{Confidence } P(B|A) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} \quad \text{Rumus 2.3 Aturan asosiasi}$$

2.1.4.4 Langkah-langkah Proses Aturan Asosiasi

Menurut Kennedi Tampubolon proses perhitungan *association rule* terdiri dari beberapa tahap adalah sebagai berikut:

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 (L1) untuk membentuk kandidat *itemset* kedua (L2). 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 bertikut (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*.

4. Setelah dilakukan operasi *join*, maka pasangan *itemset* baru hasil proses *join* tersebut dihitung *supportnya*.
5. 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.
6. Setelah itu, dari hasil *frequent itemset* tersebut dibentuk *association rule* yang memenuhi nilai *support* dan *confidence* yang telah ditentukan.
7. Pada pembentukan *association rule*, nilai yang sama dianggap sebagai satu nilai.
8. *Assosiation rule* yang terbentuk harus memenuhi nilai minimum yang telah ditentukan.
9. 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* (Lingga, 2016).

2.1.5 Sistem pendukung keputusan

Menurut Keen dan Morton, sistem pendukung keputusan adalah pasangan intelektual dari sumber daya manusia dengan kemampuan komputer untuk memperbaiki keputusan, yaitu sistem pendukung keputusan berbasis komputer bagi pembuat keputusan manajemen yang menghadapi masalah semiterstruktur (Ananta & Winiarti, 2013).

Menurut Linny Oktavianny, sistem pendukung keputusan adalah suatu sistem interaktif yang mendukung keputusan dalam proses pengambilan keputusan melalui alternatif-alternatif yang diperoleh dari hasil pengolahan data, informasi dan rancangan model (Rohayani, 2013).

Berdasarkan pendapat-pendapat diatas dapat disimpulkan bahwa sistem pendukung keputusan adalah pasangan intelektual dari sumber daya manusia dalam memproses data dengan menggunakan keahlian komputer dalam menangani suatu masalah guna membantu manajer dalam mengambil suatu keputusan.

2.1.5.1 Jenis-jenis sistem pendukung keputusan

Menurut (Yulianti, Sari, & Hayadi, 2012) ditinjau dari tingkat teknologinya sistem pendukung keputusan dibagi menjadi tiga, yaitu:

1. SPK spesifik

SPK spesifik bertujuan membantu memecahkan suatu masalah dengan karakteristik tertentu, misalnya SPK penentuan harga satuan barang.

2. Pembangkit SPK

Suatu software yang khusus digunakan untuk membangun dan mengembangkan SPK, pembangkit SPK akan memudahkan perancang dalam membangun SPK spesifik.

3. Perlengkapan SPK

Berupa software dan hardware yang digunakan atau mendukung pembangunan SPK spesifik maupun pembangkit SPK.

2.1.5.2 Komponen-komponen sistem pendukung keputusan

Adapun komponen-komponen dari sistem pendukung keputusan adalah sebagai berikut:

1. Manajemen data

Mencakup database yang mendukung data yang relevan dan diatur oleh sistem yang disebut *Database Management System (DBMS)*.

2. Manajemen model

Manajemen model adalah paket perangkat lunak yang memasukkan model-model finansial, *statistic*, ilmu manajemen, atau model kuantitatif yang lain yang menyediakan kemampuan analisis sistem dan *management software* yang terkait.

3. Antar muka Pengguna

Media interaksi antara sistem dengan pengguna, sehingga pengguna dapat berkomunikasi dan memberikan perintah pada SPK melalui subsistem ini.

4. Subsistem Berbasis Pengetahuan

Subsistem yang dapat mendukung subsistem lain atau bertindak sebagai komponen yang berdiri sendiri (Purnama, 2014).

2.1.5.3 Tujuan sistem pendukung keputusan

Sistem pendukung keputusan mempunyai tiga tujuan yang akan dicapai yaitu:

1. Membantu manajer membuat keputusan untuk memecahkan masalah semiterstruktur.
2. Mendukung penilaian manajer bukan mencoba menggantikannya.
3. Meningkatkan efektifitas pengambilan keputusan manajer dari pada efisiensinya (Iswandy, 2015).

2.1.5.4 Tahap pengambilan keputusan

Menurut Simon (Purnama, 2014), proses pengambilan keputusan meliputi tiga tahapan utama yaitu tahap inteligensi, desain dan pemilihan, namun kemudian ditambahkan dengan tahap keempat yaitu tahap implementasi. Keempat tahap tersebut dapat dijelaskan sebagai berikut:

1. Tahap penelurusan (*intelligence*)

Tahap penelurusan adalah tahap pendefinisian masalah serta identifikasi informasi yang dibutuhkan yang berkaitan dengan persoalan yang dihadapi serta keputusan yang akan diambil. Langkah ini sangat penting karena sebelum suatu tindakan diambil, tentunya persoalan yang dihadapi harus dirumuskan secara jelas terlebih dahulu.

2. Perancangan (*design*)

Perancangan adalah tahap analisis dalam mencari kaitan atau merumuskan alternatif-alternatif pemecahan masalah. Setelah permasalahan dirumuskan dengan baik maka, tahap berikutnya adalah merancang atau membangun model pemecahan masalahnya dan menyusun berbagai alternatif pemecahan masalah.

3. Pemilihan (*choice*)

Dengan mengacu pada rumusan tujuan serta hasil yang diharapkan, selanjutnya manajemen memilih alternatif solusi yang diperkirakan paling sesuai. Pemilihan alternatif ini akan mudah dilakukan kalau hasil yang diinginkan teratur atau memiliki nilai kuantitas tertentu.

4. Implementasi (*implementation*)

Implementasi adalah tahap pelaksanaan dari keputusan yang telah diambil. Pada tahap ini perlu disusun serangkaian tindakan yang terencana, sehingga hasil keputusan dapat dipantau dan disesuaikan bila diperlukan perbaikan-perbaikan.

2.1.5.5 Karakteristik sistem pendukung keputusan

Menurut Turban (Rohayani, 2013) karakteristik dari sistem pendukung keputusan adalah dibedakan sebagai berikut:

1. Sistem pendukung keputusan memberikan dukungan bagi pengambil keputusan pada situasi semi terstruktur dan tak terstruktur dengan memadukan pertimbangan manusia dan informasi terkomputerisasi.

2. Dukungan untuk semua level manajerial, dari eksekutif puncak sampai manajer line.
3. Dukungan untuk individu dan kelompok.
4. Dukungan untuk keputusan independen dan sekuensial.
5. Dukungan di semua fase proses pengambilan keputusan, yaitu *intelligence, design, choice, dan implementation*.
6. Dukungan di berbagai proses dan gaya yang berbeda-beda.
7. Adaptivitas sepanjang waktu.
8. Mudah untuk digunakan *user*.
9. Peningkatan efektivitas dari pengambilan keputusan daripada efisiensi.
10. Kontrol penuh oleh pengambil terhadap semua langkah proses pengambilan keputusan.
11. Pengguna akhir bisa mengembangkan dan memodifikasi sendiri sistem sederhana.
12. Biasanya model-model digunakan untuk menganalisis situasi pengambilan keputusan.
13. Akses disediakan untuk berbagai sumber daya, format, dan tipe, mulai dari sistem informasi sampai sistem berorientasi objek.

14. Dapat digunakan sebagai *standalone* oleh seorang pengambil keputusan pada satu lokasi atau didistribusikan di suatu organisasi secara keseluruhan dan di beberapa organisasi sepanjang rantai persediaan.

2.1.6 Data Mining

Menurut Heroe Santoso *et al*, data mining atau sering juga disebut sebagai *knowledge discovery in database* (KDD) yaitu kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santoso et al., 2016).

Menurut Chen *et al*, data mining adalah proses menemukan pengetahuan yang menarik dari sejumlah besar data yang disimpan di database, gudang data, atau repositori informasi lainnya (Chen et al., 2015).

Berdasarkan pendapat-pendapat diatas dapat disimpulkan bahwa data mining adalah suatu proses menemukan pola dengan cara pengumpulan pemakaian data yang besar yang digunakan untuk menguraikan pengetahuan dalam database.

2.1.6.1 Empat tugas utama Data Mining

Menurut (Rodiyansyah, 2015) terdapat empat tugas utama data mining yaitu:

1. *Predictive modelling*

Predictive modelling digunakan untuk membangun sebuah model untuk target variabel sebagai fungsi dari *explanatory* variabel. *Explanatory* variabel dalam hal ini adalah semua atribut yang digunakan untuk

melakukan prediksi, sedangkan variable target adalah atribut yang akan diprediksi nilainya. *Predictive modelling* dibagi menjadi dua tipe yaitu, *Classification* yang digunakan untuk memprediksi nilai dari target variabel yang diskrit dan regresi yang digunakan untuk memprediksi nilai dari target variabel yang kontinyu.

2. *Association analysis*

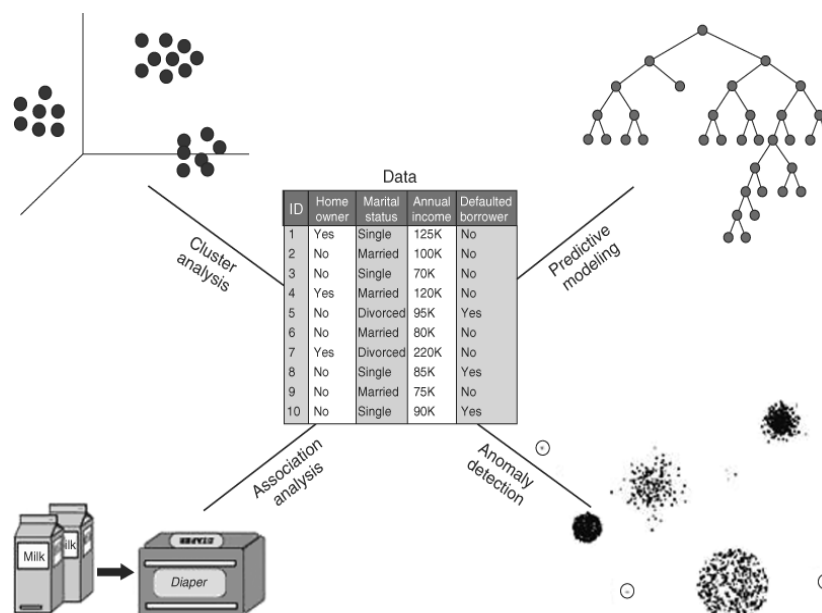
Association analysis adalah penemuan *association rule* yang menunjukkan pola-pola yang sering muncul dalam data. Terdapat nilai *support* dan *confidence* yang dapat menunjukkan seberapa besar suatu rule dapat dipercaya. *Support* adalah ukuran dimana seberapa besar tingkat dominasi suatu item atau itemset terhadap keseluruhan transaksi, sedangkan *confidence* adalah ukuran yang menunjukkan hubungan antara dua item secara *conditional*. *Association analysis* digunakan untuk menemukan aturan-aturan asosiasi yang memperlihatkan kondisi-kondisi nilai atribut yang sering muncul secara bersamaan dalam sebuah himpunan data.

3. *Cluster analysis*

Clustering digunakan untuk menganalisa objek dari data tanpa memeriksa kelas label yang diketahui. Label-label kelas dilibatkan didalam data *training* karena belum diketahui sebelumnya. *Clustering* adalah proses mengelompokkan sekumpulan objek yang sangat mirip.

4. *Anomaly detection/outlier mining*

Sebuah database dapat mengandung data objek yang tidak sesuai atau menyimpang dari model data. Data objek ini disebut *outlier*, banyak metode data mining yang menghilangkan *outlier* ini, padahal pada beberapa aplikasi seperti *fraud detection*, kejadian yang jarang terjadi tersebut justru lebih menarik untuk dianalisa daripada kejadian yang sering terjadi. Analisa dari *outlier* data tersebut disebut sebagai *outlier mining*. *Outlier mining* juga sering disebut dengan *anomaly detection* yang adalah metode pendeteksian suatu data dimana tujuannya adalah menemukan objek yang berbeda dari sebagian besar objek lain. *Anomaly* dapat dideteksi dengan menggunakan uji statistik yang menerapkan model distribusi atau probabilitas untuk data.



Gambar 2.1 Empat Tugas Data Mining

Sumber : (Rodiyansyah, 2015)

2.1.6.2 Karakteristik Data Mining

Menurut (Santoso et al., 2016) beberapa karakteristik data mining sebagai berikut:

1. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
2. Data mining biasa menggunakan data yang sangat besar.
3. Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

2.1.6.3 Tahapan Data Mining

Menurut (Santoso et al., 2016) ada beberapa tahapan data mining sebagai berikut:

1. Pembersihan data (*data cleaning*)

Pembersihan data adalah proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.

2. Integrasi data (*data integration*)

Integrasi data adalah penggabungan data dari berbagai database kedalam satu database baru.

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.

4. *Data transformation*

Data diubah atau digabung kedalam format yang sesuai untuk diproses dalam data mining.

5. *Proses mining*

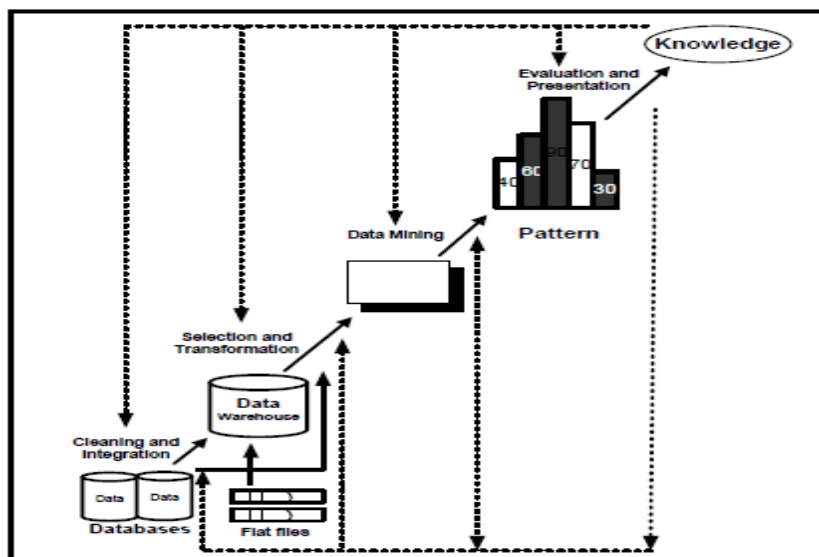
Proses *mining* adalah suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. *Evaluasi pola (pattern evaluation)*

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan bermanfaat.

7. *Persentasi pengetahuan*

Persentasi pengetahuan adalah visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.



Gambar 2.2 Tahap-Tahap Data Mining

Sumber : (Santoso et al., 2016)

2.1.6.4 Pengelompokan Data Mining

Menurut Kusrini dan Emha Taufiq Luthfid (Tampubolon et al., 2013), data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu:

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 penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variable 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 hasil akan ada dimasa mendatang, contoh prediksi bisnis dan penelitian adalah:

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

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

5. Pengklusteran (*Clustering*)

Pengklusteran adalah 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.

2.1.6.5 Langkah-langkah Data Mining

Menurut Hanter (Rinaldi, 2015), ada beberapa langkah dalam proses data mining sebagai berikut:

1. *Identity the business problem*

Identity the business problem adalah tahapan mencari pokok permasalahan perusahaan dan menganalisis apa saja data-data yang dibutuhkan untuk mengidentifikasi masalah tersebut.

2. *Mine the data for actionable information*

Mine the data for actionable information adalah tahapan dalam menjabarkan data-data yang didapatkan untuk menyelesaikan masalah.

3. *Take the action*

Take the action adalah tahapan rekomendasi dalam mengambil keputusan berdasarkan penjabaran data yang ada atau berupa tindakan yang harus dilakukan perusahaan berdasarkan analisis data yang ada.

4. *Measure results*

Measure results adalah tahapan dalam mengukur hasil dari penyelesaian masalah yang ada, apakah sesuai dengan yang diinginkan atau tidak.

2.2 Penelitian Terdahulu

Penelitian yang berhubungan dengan penelitian ini adalah:

1. Penelitian yang dilakukan oleh Usep Tatang Suryadi dan Hendra Purnama pada tahun 2014 dengan judul sistem pendukung keputusan penentuan kualitas padi dengan menggunakan algoritma apriori, memperoleh hasil penelitian bahwa perangkat lunak yang dipakai menggunakan metode apriori dimana data diurutkan dan ditabulasi sehingga dapat dihitung nilai support dan confidencenya.
2. Penelitian yang dilakukan oleh Feng Chen, Pan Deng, Jiaufu Wan, Daqiang Zhang, Athanasios V, Vasilakos and Xiaohui Rong pada tahun 2015 dengan judul *Data Mining for the Internet of Things : Literature Review and Challenges* memperoleh hasil penelitian bahwa data mining melibatkan penemuan pola baru, menarik, dan berpotensi berguna dari

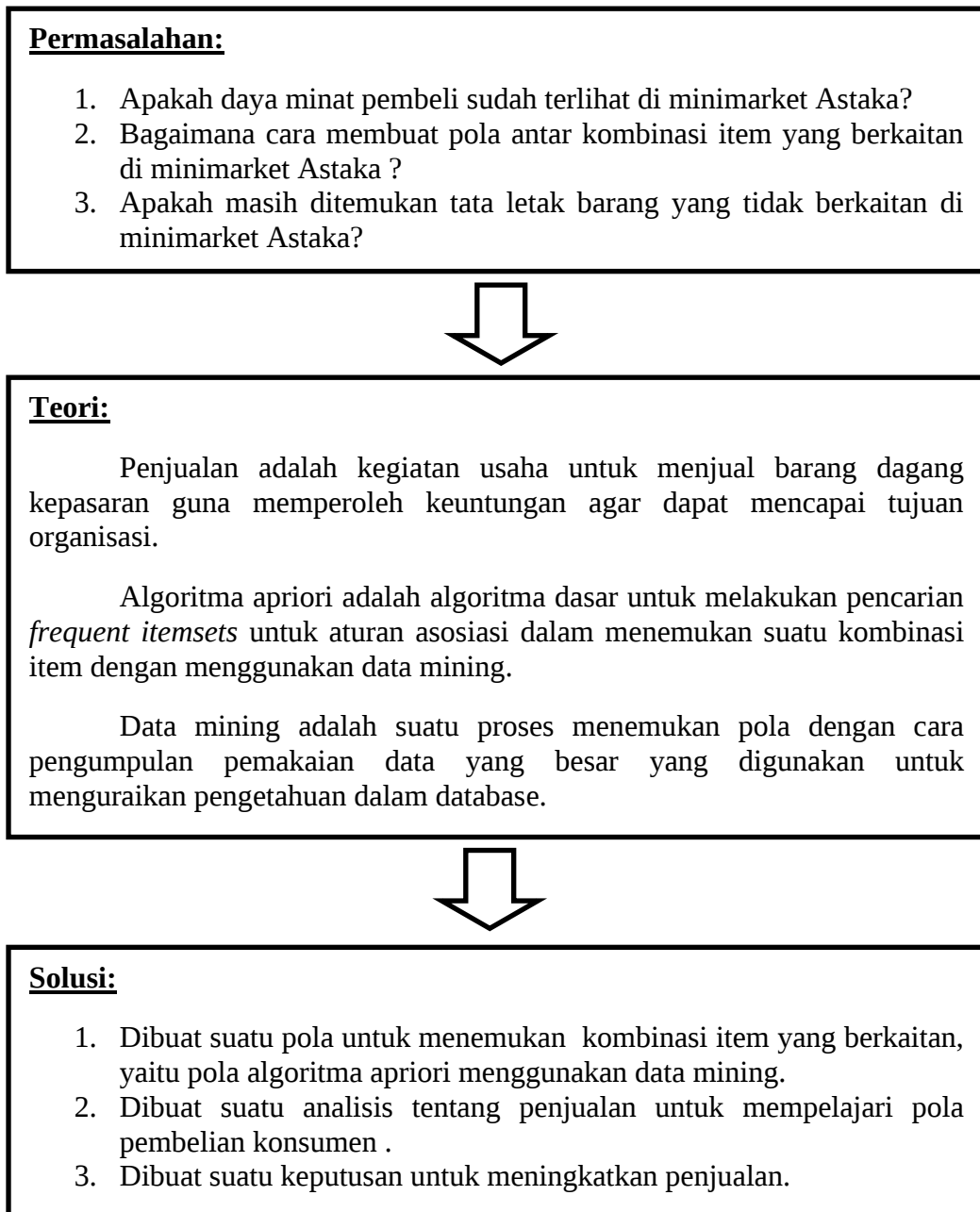
data dan menerapkan algoritma untuk mengekstrak informasi tersembunyi.

3. Penelitian ini dilakukan oleh Kennedi Tampubolon, Hoga Saragih, Bobby Reza pada tahun 2013 dengan judul Implementasi data mining algoritma apriori pada sistem persediaan alat-alat kesehatan, memperoleh hasil penelitian bahwa data mining dapat diimplementasikan 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.
4. Penelitian ini dilakukan oleh Sandi Fajar Rodiyansyah pada tahun 2015 dengan judul algoritma apriori untuk analisis keranjang belanja pada data transaksi penjualan memperoleh hasil penelitian bahwa pemanfaatan data mining untuk analisis asosiasi data transaksi dapat membantu manajemen perusahaan perdagangan untuk menentukan pola keterkaitan kemunculan barang dalam transaksi penjualan, yang pada akhirnya dapat digunakan untuk menyusun strategi penjualan.
5. Penelitian ini dilakukan oleh Heroe Santoso, I Putu Hariyadi, Prayitno pada tahun 2016 dengan judul data mining analisa pola pembelian produk dengan menggunakan metode algoritma apriori memperoleh hasil penelitian bahwa *Data Mining* dapat diimplementasikan dengan menggunakan *database* penjualan produk barang untuk dapat

menemukan kecenderungan pola kombinasi *itemsets* sehingga dapat dijadikan sebagai informasi untuk mengetahui perilaku konsumen dalam membeli produk barang secara bersamaan, sebagai *alternative* alat bantu keputusan dalam menentukan penempatan barang di area yang saling berdekatan sesuai perilaku konsumen dalam membeli barang secara bersamaan, membantu untuk mengetahui produk barang yang jarang dibeli konsumen dan sebagai alat *alternative* dalam meningkatkan strategi pemasaran dengan cara membuat diskon barang tertentu yang jarang di beli untuk menarik minat beli konsumen.

2.3 Kerangka Pemikiran

Kerangka pemikiran adalah cara berfikir atau gambaran yang dibuat peneliti tentang bagaimana cara peneliti dalam mendefenisikan teori yang telah dibuat, berikut model konseptual yang digunakan peneliti dalam mendefinisikan teori yang telah dibuat



Gambar 2.3 Kerangka pemikiran

2.4 Hipotesis penelitian ini

Hipotesis penelitian ini adalah suatu pertanyaan yang bersifat sementara atau dugaan sementara, pendapat yang mungkin benar dan mungkin salah yang masih membutuhkan jawaban.

Berdasarkan uraian kerangka pemikiran diatas, maka hipotesis penelitian ini yang disajikan peneliti adalah sebagai berikut:

1. Diduga terdapat pola kombinasi item yang berkaitan.
2. Diduga pemilik usaha mudah untuk menghasilkan pengetahuan mengenai pola pembelian konsumen.