BAB 2

TINJAUAN PUSTAKA

2.1 Tinjauan Tempat Penelitian

2.1.1 Sejarah Perusahaan

PT. Trimitra Tunas Sakti adalah sebuah perusahaan yang sudah berdiri sejak tahun 2006 merupakan perusahaan yang bergerak dalam bidang penjualan produk – produk yang berasal dari Indosat, produk yang dijual seperti kartu perdana atau stater pack (SP), pulsa isi ulang elektronik dan voucher fisik . PT. Trimitra Tunas Sakti ini memiliki dua wilayah yaitu di Karawang dan Jawa Timur.

Pada saat ini kantor cabang yang dimiliki oleh PT. Trimitra Tunas Sakti adalah yang terdapat di wilayah Karawang dan wilayah Jawa Timur diantaranya adalah :

1. Renggas Dengklok (Karawang).
2. Jombang (Jawa Timur).
3. Lamongan (Jawa Timur).
4. Madiun (Jawa Timur).
5. Magetan (Jawa Timur).
6. Ponorogo / Pacitan (Jawa Timur).
7. Nganjuk (Jawa Timur).

Produk – produk Indosat yang didistribusikan oleh PT. Trimitra Tunas Sakti ini adalah :

1. Kartu perdana isi ulang Indosat (Mentari dan IM3).
2. Kartu Internet isi ulang Indosat.
3. Voucher isi ulang fisik Indosat.
4. Voucher isi ulang elektrik Indosat.

2.1.2 Logo Perusahaan

Logo yang dimiliki oleh PT. Trimitra Tunas Sakti terdiri dari beberapa segitiga yang membentuk seperti layang – layang dan terdiri dari nama perusahaanya, seperti terlihat pada gambari berikut ini :



Gambar 2. 1 Logo Perusahaan

2.1.3 Visi dan Misi

Adapun visi dan misi dari PT. Trimitra Tunas Sakti adalah :

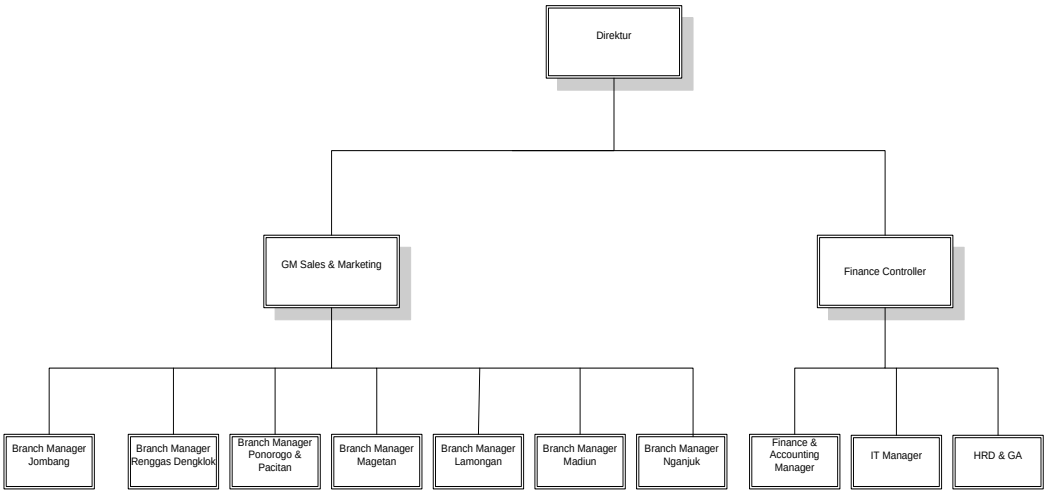
1. Visi

   Menjadi perusahaan yang handal dan terpercaya di industri telekomunikasi.

2. Misi
   - Memberikan pelayanan terbaik kepada pelanggan
   - Menghasilkan laba yang maksimal untuk perusahaan

2.1.4 Struktur Organisasi Perusahaan

Berikut ini merupakan struktur organisasi yang terdapat di PT. Trimitra Tunas Sakti :



Gambar 2. 2 Struktur organisasi perusahaan

Dan berikut ini merupakan tugas dan wewenang dari masing – masing bagian pada struktur organisasi perusahaan diatas :

a. Direktur

Tugas dan wewenang yang dimiliki oleh direktur diantaranya adalah :

1. Menentukan kebijakan tertinggi perusahaan.
2. Bertanggung jawab terhadap keuntungan dan kerugian perusahaan.
3. Mengangkat dan memberhentikan karyawan persusahaan.
4. Memelihara dan mengawasi kekayaan perseroan terbatas.
5. Bertanggung jawab dalam memimpin dan membina efektif dan efisien.
6. Mewakili perusahaan, mengadakan perjanjian – perjanjian, merencanakan dan mengawasi pelaksanaan tugas personalia yang bekerja pada perusahaan.
7. Menyusun dan melaksanakan kebijakan umum pabrik sesuai dengan kebijakan RUPS (Rapat Umum Pemegang Saham).
8. Menetapkan besarnya deviden perusahaan.

b. *General Manager* dan *Marketing*

Manager umum adalah manager yang memiliki tanggung jawab seluruh bagian atau fungsional pada suatu perusahaan atau organisasi. Manager umum memimpin beberapa unit bidang fungsi pekerjaan yang mengepalai beberapa atau seluruh manager fungsional. Pada perusahaan yang berskala kecil mungkin cukup diperlukan satu orang manager umum, sedangkan pada perusahaan atau organisasi yang berkaliber besar biasanya memiliki beberapa orang manager umum yang bertanggung-jawab pada area tugas yang berbeda-beda.

Tugas dan tanggung jawab dari seorang *Manager* Penjualan atau *Marketing Manager* adalah:

1. Bertanggung jawab terhadap Manager Umum
2. Menetapkan tujuan dan sasaran jalannya operasional perusahaan dan strategi penjualan kepada konsumen (strategi bisnis perusahaan).

3. Membuat analisa terhadap pangsa pasar dan menentukan strategi penjualan terhadap konsumen atau pelanggan.
4. Menganalisis laporan yang dibuat oleh bawahannya.
5. Mengoptimalkan kerja staf dan administrasi dibawah wewenangnya untuk mencapai tujuan perusahaan.
6. Memberikan pelayanan yang prima kepada setiap konsumen atau pelanggan.
7. Manajer pemasaran bertanggung-jawab terhadap perolehan hasil penjualan dan penggunaan dana promosi.
8. Manajer pemasaran membina bagian pemasaran dan membimbing seluruh karyawan dibagian pemasaran

c. Brach Manager

Tugas dan wewenang yang dimiliki oleh *branch manager* adalah sebagai berikut :

1. Merencanakan langkah strategis cabang, mengatur penjadwalan kunjungan dan target sales untuk pencapaian target penjualan secara maksimal
2. Memonitor dan mengevaluasi pencapaian target penjualan secara berkelanjutan
3. Memantau tugas penagihan kolektor dan tempo pembayaran customer
4. Mereview dan memastikan kesiapan sales order untuk proses pengiriman barang
5. Berkoordinasi dengan pusat dan cabang lain untuk penentuan wilayah penjualan dan koordinasi target penjualan
6. Memonitor dan mengevaluasi pasar dan kompetitor untuk melihat kedudukan cabang dengan pasar sejenis di area yang sama, menganalisa kebutuhan pasar untuk menyusun dan mengusulkan strategi penjualan
7. Menjalankan tugas-tugas terkait lainnya dalam upaya pencapaian target cabang

d. Finance controller

Tugas dan wewenang yang dimiliki oleh *finance controller* adalah sebagai berikut :

1. Mengarahkan, mengkoordinasikan dan mengukur kinerja keuangan perusahaan atas operasi harian seluruh bagian divisi dan departemen dalam perusahaan untuk tujuan perencanaan keuangan dan manajemen anggaran.
2. Memantau dan menganalisis transaksi keuangan harian, bulanan dan tahunan dan dibandingkan dengan anggaran.
3. Mengarahkan dan mengkoordinasikan pembiayaan utang dan pembayaran utang dengan lembaga eksternal.
4. Melakukan koordinasi dengan seluruh bagian dalam perusahaan untuk penyusunan laporan keuangan bulanan dan tahunan, perencanaan dan proyeksi keuangan sesuai rencana bisnis perusahaan.
5. Melakukan analisa keuangan untuk membuat usulan pertimbangan bagi atasan untuk proses negosiasi kontrak dan pengambilan keputusan atas produk investasi, rencana jangka pendek dan jangka panjang perusahaan, anggaran dan target keuangan setiap bagian dan perusahaan secara keseluruhan, pembuatan kebijakan dan perbaikan prosedur.
6. Menyajikan data keuangan yang diperlukan untuk komite perencanaan dan pembuatan kebijakan keuangan.
7. Melakukan proses audit keuangan langsung dan memberikan rekomendasi untuk perbaikan prosedural.
8. Memastikan dokumen yang berhubungan dengan kegiatan akuntansi tercatat dalam pembukuan yang akurat dan benar.
9. Memastikan dan menghitung kewajiban perpajakan yang timbul dari transaksi yang ada.
10. Memastikan konsistensi penerapan sistem akuntansi yang telah ditetapkan baik oleh perusahaan sesuai dengan standar akuntansi keuangan yang berlaku di Indonesia.

11. Memastikan bahwa setiap pencatatan dalam pembukuan yang material dapat diungkapkan.
12. Memastikan seluruh pencatatan akurat dan konsisten terhadap standar akuntansi keuangan yang berlaku.

2.2 Landasan Teori

Landasan teori ini akan menjelaskan teori – teori yang digunakan dalam pembangunan sistem ini.

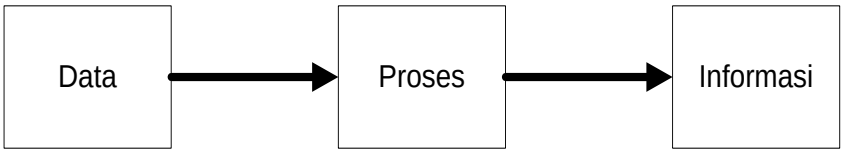
2.2.1 Konsep Data dan Informasi

2.2.1.1 Pegertian Data

Secara konspetual, data adalah deskripsi mengenai benda, kejadian, aktivitas dan transaksi, yang tidak memiliki makna atau tidak berpengaruh secara langsung kepada pengguna [7].

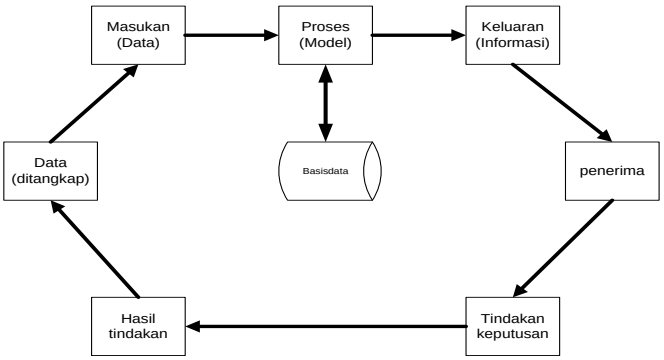
2.2.1.2 Pengertain Informasi

Informasi sebagai data yang telah di proses sedemikian rupa sehingga meningkatkan pengetahuan seseorang yang menggunakan data tersebut [7].



Gambar 2. 3 Transformasi data menjadi informasi

Gambar 2.4 memperlihatkan bagaimana siklus informasi yang menggambarkan pengolahan data menjadi informasi dan pemakaian informasi untuk mengambil keputusan, hingga akhirnaya dari tindakan hasil pengambilan keputusan tersebut dihasilkan data kembali [7].



Gambar 2. 4 Siklus informasi

2.2.2 Basis Data (*Database*)

Basis data (*Database*) adalah sekumpulan informasi bermanfaat yang diorganisasikan ke dalam tata cara yang khusus [8].

Database adalah kumpulan data, umumnya mendeskripsikan aktivitas suatu organisasi yang berhubungan atau. Misalnya, database suatu *minimarket* atau toko mungkin berisi informasi mengenai hal breikut :

- Entitas produk, entitas transaksi, entitas detail transaksi, entitas karyawan, entitas pemasok barang.
- Hubungan antar entitas, seperti produk yang dijual akan dicatat di detail transaksi, detail transaksi akan di catat di trasansaksi, setiap transaski akan disahkan oleh karyawan, dan pemasok akan memasukan produk – produk ke toko.

2.2.3 *Database Management System* (DBMS)

Management sistem basisi data adalah perangkat lunak yang di rancang untuk membantu dalam hal pemeliharaan dan utilitas kumpulan data dalam jumlah besar [8].

Database management sistem atau DBMS, adalah perangkat lunak yang didesain untuk membantu memelihara dan memanfaatkan kumpulan data yang besar. Kebutuhan terhadap sistem tersebut, termasuk juga penggunaanya, berkembang secara cepat. Alternatif penggunaan DBMS adalah untuk menyimpan data dalam file dan menulis kode aplikasi tertentu untuk penggunanya. Penggunaan DBMS memiliki beberapa manfaat penting diantaranya adalah [8]:

a. Kemandirian data

   Program aplikasi idealnya tidak diekspose pada detail representasi dan penyimpanan data. DBMS menyediakan suatu pandangan abstrak tentang data yang menyembunyikan detail tersebut.

b. Akses data efisien

   DBMS menamfaatkan berbagai teknik yang canggih untuk menyimpan dan mengambil data secara efisien. Fitur ini

terutama penting jika data disimpan pada alat penyimpanan eksternal.

c. Integritas dan keamanan data

Jika data selalu diakses melalui DBMS, maka DBMS dapat memanfaatkan batasan integritas. Misalnya, sebelum menyimpan informasi gaji untuk suatu karyawan, DBMS dapat memeriksa bahwa besarnya tidak melebihi anggaran departemen. DBMS juga dapat memanfaatkan kontrol akses yang menentukan data apa yang boleh dilihat oleh kelas pengguna yang berbeda.

d. Administrasi data

Ketika beberapa pengguna berbagi data, pemutusan administrasi data dapat memberikan perbaikan yang signifikan. Para profesional berpengalaman yang memahami sifat data yang akan diolah, dan memahami bagaimana kelompok pengguna yang berbeda menggunakan data tersebut, dapat menjadi tanggung jawab untuk mengatur representasi data untuk meminimalkan redudansi dan untuk memfinetune penympanan data guna melakukan pengambilan data yang efisien.

e. Akses konkuren dan *crash recovery*

DBMS menjadwalkan akses konkuren pada data dalam cara tertentu pada data alam cara tertentu sehingga pengguna data memandang data sebagi data yang sedang diakses oleh hanya satu pengguna pada satu waktu. Lebih lanjut, DBMS memproteksi pengguna dari efek kegagalan sistem.

f. Waktu pengembangan aplikasi terkurangi

DBMS mendukung fungsi penting yang merupkan hal yang bisa bagi banyak aplikasi untuk mengakses data dalam DBMS. Hal ini, dalam kaitanya dengan *interface* level tinggi pada data, memfasilitasi pengembangan aplikasi yang cepat. Aplikasi DBMS kemungkinan besar menjadi lebih akurat daripada

aplikasi yang berdiri sendiri karena banyak tugas penting ditangani oleh DBMS.

Manipulasi basis data meliputi pembuatan pernyataan (*query*) untuk mendapatkan informasi tertentu, melakukan pembaharuan atau penggantian (*update*) data, serta pembuatan *report* dari data. Tujuan utama DBMS adaah untuk menyediakan tinjauan abstrak dari data bagi pengguna. Jadi sistem menyembunyikan informasi mengenai bagaimana data disimpan dan dirawat, tetapi data tetap dapat diambil.

2.2.4 Pengertian Data Mining

Data mining atau sering disebut juga dengan *knowledge discovery in database* (KDD), adalah kegiatan yang meliputi penugmpulan, pemakaian data historis untuk menemukan keterurutan, pola atau hubungan dalam set data berukuran besar, keluaran dari *data mining* ini dapat dipakai untuk memperbaiki pengambilan keputusan dimasa depan [1].

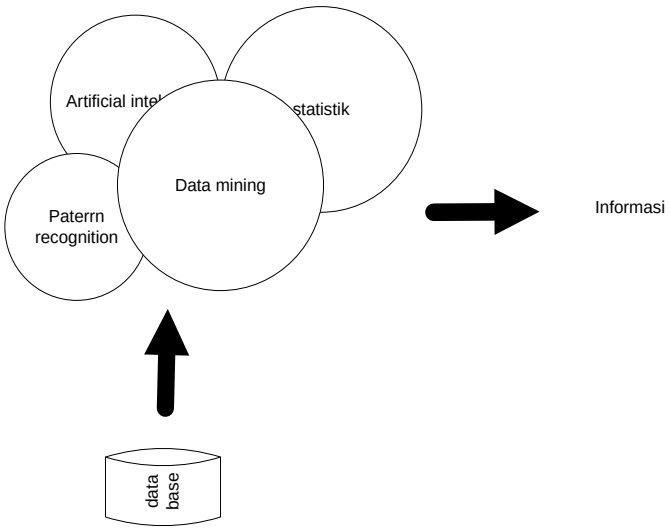
Data mining didefinisikan sebagai suatu set teknik yang digunakan secara otomatis untuk mengeksplorasi secara menyeluruh dan membawa ke permukaan raslasi – relasi yang kompleks pada set data yang sangat bersar [3].

Data mining adalah proses yang menggunakan teknik statistik, perhitungan, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan terkait dari berbagai basis data besar [4].

Dari pengertian mengenai *data mining* di atas maka dapat disimpulkan bahwa data mining adalah suatu disiplin ilmu yang digunakan untuk mencari suatu pengetahuan dari tumupkan data tersebut dengan menggunakan sebuah metode *data mining*. *Data mining* membantu pengguna dengan menemukan hubungan dan menyajikan dengan cara yang dapat dipahami sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan.

Data mining adalah satu ilmu di bidang komputer yang cukup banyak penerapanya, ditunjang kekayaan dan keanekaragaman berbagai bidang ilmu lainya seperti *artificial intelligence*, *database*, statistik, pemodelan

matematika pengolahan citra, dan sebagainya, sehingga penerapan *data mining* menjadi kasus semakin luas.



Gambar 2. 5 Akar ilmu *data mining*

Munculnya data minng didasarkan pada jumlah data yang tersimpan dalam basis data semakin besar. Misalnya dalam sebuah toko, ada beberapa transaksi pelanggan yang terjadi dalam sehari dan beberapa juta data yang sudah tersimpan dalam sebulan. Dalam perusahaan dapat menghasilkan beberapa juta data produksi barang, ataupun apabila mendapatkan sebuah rekomendasi untuk membeli barang dari sebuah sistem belanja ketika membeli suatu produk.

2.2.5 Arsitektur Dara Mining

Pada umumnya sistem *data mining* terdiri dari komponen – komponen berikut [5]:

a. *Database, data warehouse*

   Media dalam ha ini bisa jadi berupa *database, data warehouse, spreadsheet,* atau jenis – jenis penampungan informasi lainnya. *Data cleaning* dan *data integration* dapat dilakukan pada data tersebut.

b. *Database,* atau *data warehouse server*

   Database atau *data warehouse server* bertanggungjawab untuk menyediakan data yang relevan berdasarkan permintaan dari *user* pengguna *data mining*.

c. Basis pengalaman (*knowladge base*)

Merupakan basis pengetahuan yang digunakan sebagai panduan dalam pencarian pola.

d. *Data mining engine*

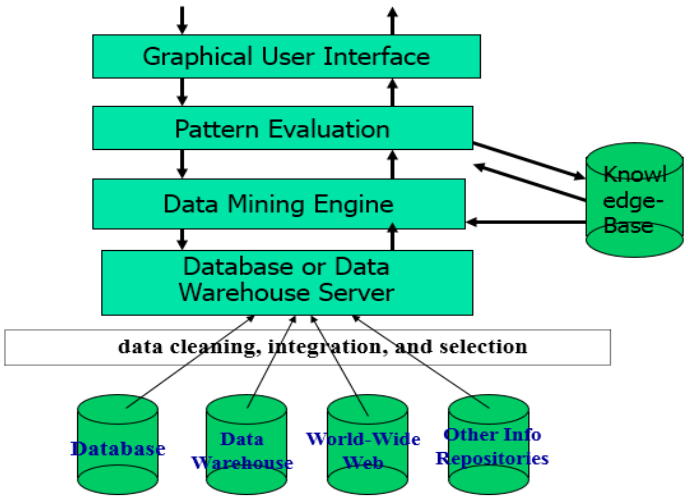
Bagian dari *software* yang menjalankan program berdasarkan algoritma yang ada.

e. *Pattern evaluation module*

Bagian dari *software* yang berfungsi untuk menemukan *paterrn* atau pola –pola yang yang terdapat di dalam *database* yang diolah sehingga nantinya proses *data mining* dapat menentukan *konwladge* yang sesuai.

f. *Graphical user interface*

Bagian ini merupakan sarana antara *user* dan sistem *data mining* untuk berkomunikasi, dimana *user* dapat berinteraksi dengan sistem melalui *data mining query*, untuk menyediakan informasi yang dapat membantu *user* untuk melakukan *browsing* pada *databse* dan *data warehouse*, mengevaluasi *pattern* yang telah dihasilkan, dan menampilkan *pattern* tersebut dengan tampilan yang berbeda – beda.

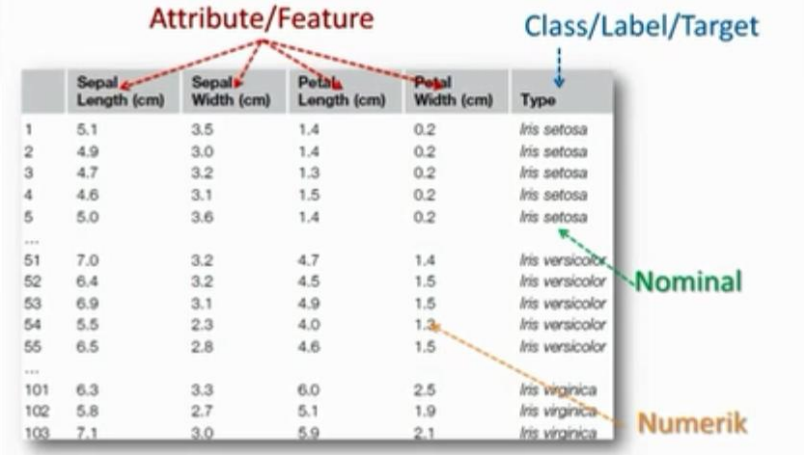


Gambar 2. 6 Arsitektur *data mining*

2.2.6 Himpunan Data Dalam Data Mining

Himpunan data yang akan dibahas dalam *data mining* ini adalah sebuah data yang terstruktur yang memiliki atribut, dan kelas atau label. Tipe data yang digunakan memiliki dua jenis yaitu :

1. Tipe data numerik adalah suatu tipe data yang dapat dikalkulasi (dapat dilakukan penjumlahan, pengurangan, perkalian, pembagian).
2. Tipe data nominal adalah suatu tipe data yang tidak dapat dikalkulasi, contoh datanya adalah data jenis kelamin yang isinya adalah laki – laki, atau perempuan.



	Sepal Length (cm)	Sepal Width (cm)	Petal Length (cm)	Petal Width (cm)	Type
1	5.1	3.5	1.4	0.2	*Iris setosa*
2	4.9	3.0	1.4	0.2	*Iris setosa*
3	4.7	3.2	1.3	0.2	*Iris setosa*
4	4.6	3.1	1.5	0.2	*Iris setosa*
5	5.0	3.6	1.4	0.2	*Iris setosa*
...					
51	7.0	3.2	4.7	1.4	*Iris versicolor*
52	6.4	3.2	4.5	1.5	*Iris versicolor*
53	6.9	3.1	4.9	1.5	*Iris versicolor*
54	5.5	2.3	4.0	1.3	*Iris versicolor*
55	6.5	2.8	4.6	1.5	*Iris versicolor*
...					
101	6.3	3.3	6.0	2.5	*Iris virginica*
102	5.8	2.7	5.1	1.9	*Iris virginica*
103	7.1	3.0	5.9	2.1	*Iris virginica*

Gambar 2. 7 Himpunan Data

2.2.7 Metode Data Mining

Pekerjaan yang berkaitan dengan *data mining* dapat dibagi menjadi empat kelompok [6] :

a. Model prediksi (*Prediction modelling*)

Pekerjaan ini berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variabel ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat.

Model prediksi ini digunakan ketika tipe data yang dimiliki oleh himpunan data merupakan tipe data numerik, dan ciri khas dari model prediksi ini merupakan *dataset* dalam bentuk *time series.* Contoh algoritma yang dapat digunakan dalam model prediksi ini adalah algoritma *neural network* (ANN).

Contoh pekerjaan yang menggunakan jenis klasifikasi adalah untuk melakukan prediksi harga saham dengan *clas* atau label berupa tanggal penutupan saham.



Row No.	Close	Date	Open	High	Low	Volume
1	1286.570	Apr 11, 2006	1296.600	1300.710	1282.960	223288000(
2	1288.120	Apr 12, 2006	1286.570	1290.930	1286.450	193810000(
3	1289.120	Apr 13, 2006	1288.120	1292.090	1283.370	189194000(
4	1285.330	Apr 17, 2006	1289.120	1292.450	1280.740	179465000(
5	1307.280	Apr 18, 2006	1285.330	1309.020	1285.330	259544000(
6	1309.930	Apr 19, 2006	1307.650	1310.390	1302.790	244731000(
7	1311.460	Apr 20, 2006	1309.930	1318.160	1306.380	251292000(
8	1311.280	Apr 21, 2006	1311.460	1317.670	1306.590	239263000(
9	1308.110	Apr 24, 2006	1311.280	1311.280	1303.790	211733000(
10	1301.740	Apr 25, 2006	1308.110	1310.790	1299.170	236638000(
11	1305.410	Apr 26, 2006	1301.740	1310.970	1301.740	250269000(
12	1309.720	Apr 27, 2006	1305.410	1315	1295.570	277201000(
13	1310.610	Apr 28, 2006	1309.720	1316.040	1306.160	241992000(

Gambar 2. 8 Contoh data saham

Dari data saham tersebut dilakukan pembelajaran misalnya dengan *neural network* sehingga didapakan sebuah pengetahuan, pengetahuan yang didapatkan berupa pengahuan berapa harga sahama kedepan.

b. Analisis *Cluster*

Analisis *cluster* / kelompok melakukan pengelompokan ke dalam sejumlah kelompok berdasarkan kesamaan karakteristik masing – masing data pada kelompok – kelompok yang ada. Data – data yang masuk ke dalam batas kesamaan dengan kelompoknya akan bergabung ke dalam kelompok tersebut, dan akan terpisah dalam kelompok yang berbeda jika keluar dari batas kesamaan kelompok tersebut.

Metode *clustering* digunakan apabila dataset yang dimiliki tidak memiliki label, metode *clustering* ini tidak memerlukan guru dalam melakukan pembelajaranya atau biasa disebut dengan *unsupervised learning*, hal ini yang membedakan model *clustering* dengan model yang lainya.

pekerjaan yang berkaitan dengan *cluster* adalah bagaimana bisa mengetahui pola pembelian barang oleh konsumen pada waktu – waktu tertentu. Dengan mengetahui pola kelompok pembelian tersebut, maka perusahaan / *retailer* dapat menentukan jadwal promosi yang dapat diberikan sehingga dapat membantu meningkatkan omzet penjualan.

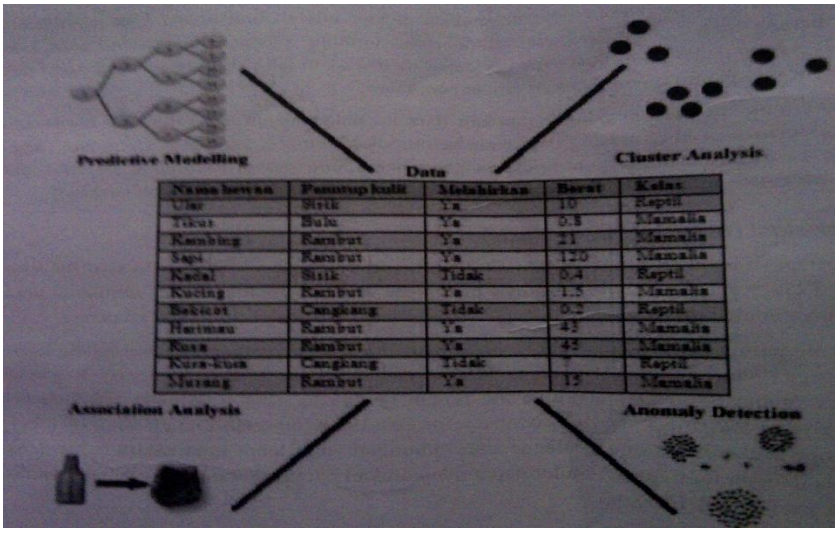
c. Analisis Asosiasi

Analsisi asosiasi digunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang ditemukan biasanya merepresentasikan bentuk aturan implikasi atau subset fitur. Tujuanya adalah untuk menemukan pola yang menarik dengan cara yang efisien.

Penerapan yang paling dekat dengan kehidupan sehari – hari adalah analisis data keranjang belanja. Jika Ibu rumah tangga akan membeli kebutuhan rumah tangga (misalnya beras) di sebuah *super market*, maka sangat besar kemungkinan ibu rumah tangga tersebut juga akan membeli kebutuhan rumah tangga yang lain, misalnya minyak atau telur, dan tidak munggkin (atau jarang) membeli barang lain seperti topi atau buku. Dengan mengetahui hubungan yang lebih kuat antara beras dengan telur daripada beras dengan topi, maka *retailer* dapat menentukan barang – barang yang sebaiknya disediakan dalam jumlah yang cukup banyak.

d. Deteksi Anomali

Pekerjaan deteksi anomali berkaitan dengan pengamatan sebuah data dari sejumlah data secara signifikan mempunyai karakteristik yang berbeda dari sisa data yang lain. Data – data yang karakteristiknya menyimpang (berbeda) dari data yang lain disebut dengan *outlier*. Deteksi anomali dapat diterapkan pada sistem jaringan untuk mengetahui pola data yang memasuki jaringan sehingga dapat diketahui adanya penyusupan jika pola kerja data yang datang berbeda [6].



Nama hewan	Penutup kulit	Melahirkan	Berat	Kelas
Ular	Sisik	Ya	10	Reptil
Tikus	Bulu	Ya	0.8	Mamalia
Kambing	Rambut	Ya	21	Mamalia
Sapi	Rambut	Ya	120	Mamalia
Kadal	Sisik	Tidak	0.4	Reptil
Kucing	Rambut	Ya	1.5	Mamalia
Bekicot	Cangkang	Tidak	0.2	Reptil
Harimau	Rambut	Ya	45	Mamalia
Rusa	Rambut	Ya	45	Mamalia
Kura-kura	Cangkang	Tidak	?	Reptil
Murang	Rambut	Ya	15	Mamalia

Gambar 2. 9 Pekerjaan utama *data mining*

2.2.8 Pengelompokan (*Clustering*)

Clustering merupakan pekerjaan yang memisahkan data / *vector* ke dalam sejumlah kelompok (*cluster*) menurut karakteristiknya masing – masing. Data – data yang memiliki kemiripan karakteristik yang berkumpul dalam *cluster* yang sama, dan data – data dengan karakteristik berbeda akan terpisah dalam *cluster* yang berbeda. Tidak diperlukan label baru bisa diberikan ketika *cluster* sudah terbentuk. Karena tidak adanya target label kelas untuk setiap data maka *clustering* sering disebut juga pembelajaran tidak terbimbing (*unsupervised learning*) [6].

Teknik *clustering* banyak diterapkan dalam berbagai bidang, seperti kedoktrean, teknik *clustering* dapat digunakan untuk mengelompokan jenis – jenis penyakit berbahaya berdasarkan karakteristik atau sifat – sifat penyakit pasien. Dalam bidang keshatan dapat digunakan untuk mengelompokan jenis makanan berdasarkan kandungan kalori, vitamin, protein. Dalam bidang klimatologi dapat digunakan untuk mengetahui pola angin dan kondisi cuaca di udara sehingga bias diketahui wilayah – wilayah yang retang terhadap cuaca buruk, dan sebagainya.

Clustering dapat dibedakan menjadi dua tujuan, yaitu *clustering* untuk pemahaman dan *clustering* untuk penggunaan. Jika tujuanya untuk pemahaman maka *cluster* yang terbentuk harus menangkap struktr alami

data. Biasanya proses *clustering* dalam tujuan ini hanya sebagai proses awal untuk kemudian dilanjutkan dengan pekerjaan inti seperti *summarization* (rata – rata, standar deviasi), pelebelan kelas pada setiap kelompok untuk kemudian digunakan sebagai data latih klasifikasi, dan sebaginya. Sementara jika tujuanya untuk penggunaan, biasanya tujuan utama untuk mencari *prototype cluster* yang paling representative terhadap data dan memberikan abstraksi dari setiap objek data dalam *cluster* di mana sebuah data terletak di dalamnya.

Berdasrkan strukturnya, *clustering* terbagi menjadi dua, yaitu hirarki dan partisi. Dalam pengelompokan berbasis hirarki (*hierarchical clustering*), satu data tunggal bisa dianggap sebuah cluster, dua atau lebih *cluster* kecil dapat bergabung menjadi sebuah *cluster* bersar. Begitu seterusnya hingga semua data dapat bergabung menjadi sebuah *cluster*. Pengelompokan berbasis partisi membagi set data ke dalam sejumlah *cluster* yang tidak bertumpang – tindih antara satu *cluster* dengan *cluster* yang lain, artinya setiap data hanya menjadi anggota satu *cluster* saja.

Teknik *clustering* berdasarkan keanggotaanya dibedakan menjadi dua bagian, yaitu eklusif dan tumpang-tindih. Dalam kategori eklusif, sebuah data bias dipastikan hanya menjadi satu *cluster* dan tidak menjadi anggota di *cluster* yang lain. Metode clustering yang termasuk kategori ini adalah *K – Means, DBSCAN,* dan *SOM.* Sementara yang termasuk kategori tumpang-tindih adalah metode *clustering* yang membolehkan sebuah data menjadi anggota di lebih dari satu *cluster,* misalnya *Fuzy C-Means*.

Hal yang terpenting dalam *clustering* adalah matriks yang digunakan untuk mengukur ketidak miripan data yang dikelompokan. Penggunaan matriks yang berbeda dapat memberikan hasil yang berbeda, tergantung kasus yang diselesaikan. Matriks yang paling banyak digunakan adalah matriks *Euclidean.* Secara geometris matriks ini memberikan jarak terpendek antara dau data [6].

Menghitung jarak dengan menggunakan Euclidean :

$d = \sqrt{\sum_{i=1}^{r}(X\,i - Yi)^2}$ ……..……..(2.1)

Dimana :

d = Jarak

r = jumlah fitur dalam vektor

Xi = nilai fitur *X* ke i

Yi = nilai fitur *Y* ke i

2.2.8.1 Teknik *Clustering*

Tujuan utama dari metode *cluster* adalah pengelompokan sejumlah data atau obyek ke dalam *cluster* (grup) sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin. Dalam *clustering* akan berusaha untuk menenpatkan obyek yang mirip (jaraknya dekat) dalam satu *cluster* sangat mirip satu sama lain dan berbeda dengan obyek dalam klaster – klaster yang lain.

Ada dua pendekatan dalam teknik *clustering,* yaitu partisi dan hirarki. Dalam teknik partisi akan mengelompokan obyek X_1, X_2, X_3, …., X_n ke dalam *k cluster.* Ini bias dilakukan dengan menentukan pusat *cluster* awal, lalu dilakukan relokasi obyek berdasarkan kriteria tertentu samapi dicapai pengelompokan yang optimum. Dalam *cluster* hirarki dimulai dengan membuat *m cluster* dimana setiap *cluster* beranggotakan satu obyek dan berakhir dengan satu *cluster* diaman anggotanya adalah *m* obyek. Pada setiap tahap dalam prosedurnya, satu *cluster* digabung satu *cluster* yang lain. Penghentian atau *cut-off* bias dilakukan ketika jumlah *cluster* yang diinginkan sudah tercapai [6].

2.2.8.2 Analisis *Cluster* Berbasis Hirarki

Dalam statistik, pengelompokan berbasis hirarki adalah metode analisis *cluster* yang berusaha untuk membangun sebuah hirarki *cluster*. Strategi untuk mengelompokan berbasis hirarki pada umumnya terdapat dua jenis, yaitu *agglomerative* dan *divisive*.

Agglomerative merupakan sebuah metode pengelompokan berbais hirarki dengan pendekatan *bottom up,* yaitu proses pengelompokan dimulai dari masing – masing data sebagai satu buah *cluster,* kemudian secara

rekursif mencari *cluster* terdekat sebagai pasangan untuk bergabung sebagai *cluster* yang lebih besar. Proses terbut diulang terus sehingga tampak bergerak ke atas membentuk hirarki.

Divisive merupakan metode pengelompokan data berbasis hirarki dengan pendekatan *top down*, yaitu proses pengelompokan dimulai dari satu *cluster* yang berisi semua data, kemudian secara rekursif memecah *cluster* menjadi dua *cluster* sampai *cluster* hanya berisi data tunggal (data itu sendiri). Untuk cara ini, dibutuhkan adalah keputusan *cluster* yang manakah yang akan dipecah pada setiap langkah dan bagaimana cara memechakanya.

2.2.8.3 *Agglomerative Hierarchical Clustering* (AHC)

Agglomerative hierarchical clustering merupakan algoritma yang digunakan untuk mengukur kedekatan diantara dua *cluster*. Terdapat tiga cara yang digunakan AHC, diantaranya adalah, metode *single linkage, complete linkage, average linkage*.

Single linkage memberikan hasil apabila *cluster* – *cluster* digabugkan menurut jarak yang paling dekat diantara dua *cluster*. *Complete linkage* terjadi apabila kelompok – kelompok digabung menurut jarak antara yang paling jauh diantara dua *cluster. Average linkage* digabungkan menurut jarak rata – rata antara pasangan – pasangan anggota masing – masing himpunan diantara dua *cluster*.

Berikut ini merupakan tahapan dari algoritma *Agglomerative* :

1. Tentukan jumlah *cluster* yang akan dibentuk.
2. Hitung matriks kedekatan berdasarkan jarak yang digunakan (misalkan dengan menggunakan matriks *euclidean ditance* yang terdapat dalam persamaan 2.1).
3. Ulangi langkah empat dan lima, hingga , menghasilkan jumlah *cluster*.
4. Gabung dua *cluster* terdekat berdasarkan parameter kedekatan yang ditemntukan (misalkan dengan metode *single linkage* pada persamaan 2.2).
5. perbaharui matriks kedekatan untuk mereflesikan kedekatan diantara *cluster* baru dan *cluster* yang tersisi [6].

Pada metode *single linkage* (*Min*), kedekatan diantara dua *cluster* ditentukan dari jarak terdekat (terkecil) diantara pasangan diantara dua dari *cluster* berbeda (satu dari *cluster* pertama satu dari *cluster* lainnya) atau disebut juga dengan kemiripan yang paling maksimal. Dengan cara seperti ini maka akan dimulai dari masing – masing data sebagai *cluster*, kemudain mencari tetangga terdekat dan menggunakan *single linkage* untuk menggabungkan dua *cluster* berikutnya hingga semuanya bergabung menjadi satu *cluster*.

Berikut ini merupakan persamaan *single linkage* :

$$d(U,V) = \min\{d(U,V)\}; d(U,V) \in D \quad \text{.....(2.2)}$$

D = Matriks kedekatan jarak antar *cluster*.

2.2.8.4 Contoh Penerapan *Clustering* Menggunakan Algortima AHC

Contoh diberikan 5 data dan pengukuran jarak menggunakan jarak *Euclidean*, dengan menggunakan metode *single linkage.*

Tabel 2. 1 Set data untuk dihitung

Data ke –i	Fitur x	Fitur y
1	1	1
2	4	1
3	6	1
4	1	2
5	2	3

Yang pertama dilakukan adalah menghitung jarak dengan menggunakan *euclidean* pada semua pasangan dua data. Hasil dari perhitungan dengan menggunakan *euclidean* adalah sebagai berikut :

$$D = d(U,V) = \begin{bmatrix} 0 & 3 & 1 & 5 & 7 \\ 3 & 0 & 4 & 4 & 4 \\ 1 & 4 & 0 & 4 & 6 \\ 5 & 4 & 4 & 0 & 2 \\ 7 & 4 & 6 & 2 & 0 \end{bmatrix}$$

Dengan memperlakukan setiap data sebagai *cluster* maka *cluster* yang terbentuk pada saat ini adalah sebanyak 5 *cluster* atau jumlah *cluster*

K = jumlah data *n*, tahapan selanjutnya adalah memilih jarak dua *cluster* yang paling kecil.

$$\min\{d(U,V)\} = d13 = 1$$

Dari proses perhitungan dengan menggunakan metode *single linkage* maka didapatkan bawha jarak yang paling minimum adalah *cluster* 1 dan *cluster* 3, maka kedua *cluster* ini akan digabung. Untuk melanjutkan tingkat pengelompokan berikutnya maka jarak – jarak antara *cluster* (13) dengan *cluster* yang tersisia 2,4, dan 5 dihitung kembali dengan menggunakan metode *single linkage*. Jarak – jarak yang didapatkan adalah:

$d_{(13)2} = \min\{d_{12}, d_{(32)}\} = \min\{3,4\} = 3$

$d_{(13)4} = \min\{d_{14}, d_{(34)}\} = \min\{5,4\} = 4$

$d_{(13)5} = \min\{d_{15}, d_{(35)}\} = \min\{7,6\} = 6$

Dengan menghapus baris – baris dan kolom – kolom matriks D yang bersesuain dengan *cluster* 1 dan 3 dan menambahkan baris dan kolom untuk *cluster* 1 dan 3, dan menambahkan baris dan kolom untuk *cluster* (13), maka didapatkan matriks yang baru :

$$D = d(U,V) = \begin{matrix} (13) \\ 2 \\ 4 \\ 5 \end{matrix} \begin{bmatrix} 0 & 3 & 4 & 6 \\ 3 & 0 & 4 & 4 \\ 4 & 4 & 0 & 2 \\ 6 & 4 & 2 & 0 \end{bmatrix}$$

Tahapan selanjutnya adalah memilih kembali jarak dua *cluster* yang paling kecil.

$$\min\{d(U,V)\} = d45 = 2$$

Maka *cluster* yang terpilih adalah *cluster* 4 dan 5, maka *cluster* 4 dan *cluster* 5 digabung. Kemudian untuk menghitung jarak – jarak antara *cluster* (45) dengan *cluster* lain yang tersisa yaitu (13) dan 2 dihitung kembali dengan menggunakan metode *single linkage*. Jarak – jarak yang didapatkan adalah :

$d_{(45)(13)} = \min\{d_{41}, d_{43}, d_{51}, d_{53}\} = \min\{5,4,7,6\} = 4$

$d_{(45)2} = \min\{d_{42}, d_{(52)}\} = \min\{4,4\} = 4$

dengan menhapus baris – baris dan kolom – kolom matriks D yang bersesuaian dengan *cluster* 4 dan *cluster* 5 dan menambahkan baris dan kolom untuk *cluster* (45), maka didapatkan matriks jarak yang baru :

$$D = d(U,V) = \begin{matrix}(45)\\(13)\\2\end{matrix}\begin{bmatrix}0 & 4 & 4\\4 & 0 & 3\\4 & 3 & 0\end{bmatrix}$$

Selanjutnya dipilih kembali jarak dua *cluster* yang paling kecil.

$$\min\{d(U,V)\} = d(13)2 = 3$$

Terpilih *cluster* (13) dan 2, maka *cluster* (13) dan 2 digabung. Untuk melanjutkan tingkat *clustering* berikutnya, maka jarak – jarak antara *cluster* (13) dan 2 dengan *cluster* yang lain yang tersisisa yaitu (45) dihitung kembali dengan menggunakan metode *single linkage*. Jarak – jarak yang didapatkan adalah :

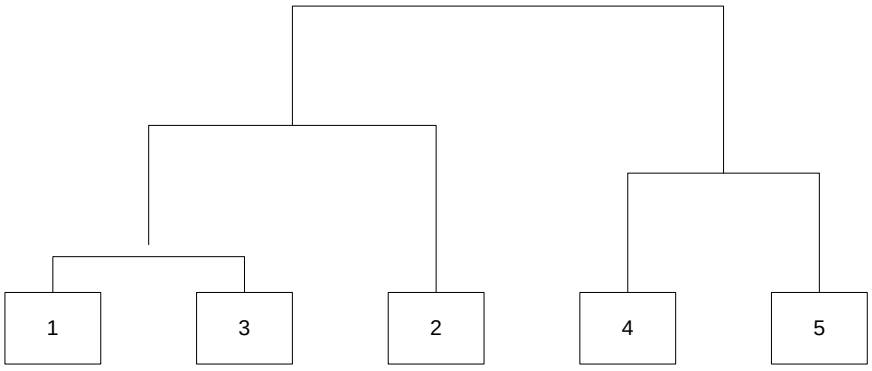
$d_{(123)(45)} = \min\{d_{14}, d_{43}, d_{24}, d_{25}, d_{34}, d_{35}\} = \min\{5,7,4,4,4,6\} = 4$

Dengan menghapus baris – baris dan kolom – kolom matriks D yang bersesuaian dengan *cluster* (13) dan 2, dan menambahkan baris dan kolom untuk *cluster* (123), maka matriks yang baru yang didapatkanya adalah :

$$D = d(U,V) = \begin{matrix}(123)\\(45)\end{matrix}\begin{bmatrix}0 & 4\\4 & 0\end{bmatrix}$$

Jadi *cluster* (123) dan (45) digabung membentuk *cluster* tunggal dari semua 5 data, (12345), ketika jarak terdekat mencapai 4.

Hasil *clustering* yang didapatkan dapat ditampilkan dalam bentuk dendogram, seperti terlihat dalam gambar 2.10 :



Gambar 2. 10 Dendrogram hasil *clustering* berbasis hirarki

2.2.9 Tahapan Data Mining

Dalam menyelesaikan penelitian data mining terdapat sebuah standar yang dapat digunakan untuk menyelesaikan penelitian data mining, standar tersebut dinamakan dengan *Cross – Industry Standard for Data Mining* (CRISP-DM). CRISP-DM merupakan suatu standar yang telah dikembangkan pada tahun 1996 yang ditunjukan untuk melakukan proses analisis dari satu industri sebagai strategi pemecahan masalah dari bisnis atau unit penelitian [2].

Berikut ini merupakan tahapan – tahapan dalam CISP-DM [2] :

1. Pemahaman bisnis

   Tahapan pemahaman bisnis ini merupakan tahapan awal pada pemahaman tujuan dan kebutuhan proyek dari perspektif bisnis, kemudian mengubah pengetahuan tersebut menjadi sebuah masalah data mining dan rencana awal untuk mencapai tujuan. Dalam tahapan pemahaman bisnis ini dibagi menjadi beberapa bagian :

   a. Identifikasi tujuan bisnis

      Dalam tahapan ini bertujuan untuk memahami proses bisnis yang ingin dicapai.

   b. Pemahaman situasi

      Dalam tahapan pemahaman situasi ini untuk mencari fakta yang ada mengenai sumber daya yang ada, asumsi dan faktor yang harus dipertimbangkan dalam proses data mining.

c. Penentuan sasaran data mining

Dalam proses penentuan sasaran data mining ini adalah untuk menentukan kriteria sukses dari data mining.

2. Pemahaman Data

Tahap pemahaman data ini merupakan tahapan untuk memahami data yang berkaitan dengan penelitian yang akan dilakukan, dalam tahapan pemahaman data ini terdapat beberapa tahapan diantaranya adalah :

a. Pengumpulan data awal

Dalam pengumpulan data awal ini digunakan untuk mengumpulkan data yang akan digunakan.

b. Penjelasan data.

Data yang telah diperoleh dari tahapan pengumpulan data kemudian dijelaskan dalam tahapan penjelasan data ini.

c. Eksplorasi data

Tahapan eksplorasi data ini betujuan untuk mejelaskan data melalui statistika, ataupun dengan menggunakan visualisasi data.

1. Analisis statistik deskriptif

Dengan menggunakan analisis statistik deskriptif ini adalah untuk membantu terciptanya tujan dari data mining .Analisis statistik deskriptif digunakan adalah dengan mencari nilai – nilai dibawah ini :

a. Nilai minimal
b. Nilai maksimal
c. Nilai rata – rata dengan menggunakan rumus [12] :

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad \text{.....(2.3)}$$

d. Nilai standar deviasi dengan menggunakan rumus [12] :

$$s^2 = \frac{\sum(xi - x)^2}{n-1} \quad \text{.....(2.4)}$$

2. Visualisasi data

Dalam tahapan visualisasi data ini akan memeriksa data dengan cara menggambarkan bagaimana data yang ada, dengan cara melihat apakah terdapat *mising value*, ataupun *outlier.*

1. *Mising value.*
2. *Outlier*

   Outlier adalah data yang secara nyata berbeda dengan data yang lain [13]. Metode yang dapat digunakan untuk mendeteksi *outllier* berdasarkan teknis statistik. Metode ini menggunakan *threshold* untuk dinyatakan sebagai *outlier.* Perhitungan *thershold* menggunakan persamaan (2.5)

   Batas atas = mean + 2 * standar deviasi…….(2.5)

   Batas bawah = mean – 2 * standar deviasi…….(2.5)

d. Evaluasi data

   Dalam tahapan evaluasi ini bertujuan untuk mengevaluasi data yang telah dilakukan dalam tahapan eksplorasi data.

3. Persiapan Data

   Tahapan persiapan data ini merupakan tahapan yang mencakup semua kegiatan yang diperlukan untuk membangun *dataset* akhir (data yang akan digunakan dalam *modeling tools*) dari data mentah awal, dalam tahapan persiapan data ini terdapat beberapa tahapan diantaranya adalah :

   a. Pemilihan data

      Tahapan pemilihan data ini merupakan tahapan yang digunakan untuk memilih data yang akan digunakan, pemilihan tersebut meliputi pemilihan atribut ataupun pemilihan baris.

   b. Pembersihan data

      Tahapan pembersihan data merupakan tahapan untuk menghilangakan atau membersihkan data yang dihasilkan dalam tahapan evaluasi data.

   c. Penyiapan data awal

      Pada tahapan penyiapan data ini digunakan untuk menyiapkan data awal yang akan digunakan untuk tahapan pemodelan.

4. Pemodelan

   Tahapan pemodelan merupakan tahapan pemilihan model dan mengaplikasikan model yang sesuai. Dalam tahapan pemodelan ini terdapat beberapa tahapan diantaranya adalah :

   a. Memilih teknik pemodelan

      Dalam tahapan ini digunakan untuk memilih teknik pemodelan yang sesuai dengan permasalahan dan tujuan yang ingin dicapai.

   b. Pembuatan model

      Dalam tahapan ini dijelasakan mengenai teknik pemodelan yang telah dipilih

   c. Analisis pengujian model

      Dalam tahapan ini model yang telah dipilih dijalankan dengan menggunakan kasus uji.

5. Evaluasi

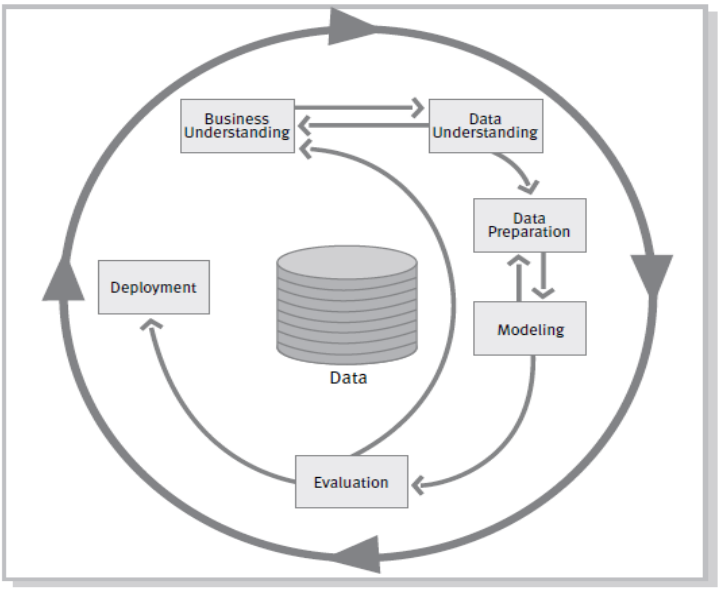
   Dalam tahapan ini akan dilakukan evaluasi terhadap model yang telah digunakan apakah model yang telah digunakan tersebut dapat mencapai tujuan yang ditetapkan pada fase pemahaman bisnis. Dalam tahapan evaluasi ini terdapat beberapa tahapan yang digunakan daintaranya adalah :

   a. Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan.
   b. Menetapkan apakah model yang digunakan sudah sesuai dengan tujuan pada fase awal.
   c. Menentukan apakah terdapat permasalahan dari tahapan pemahaman bisnis yang tidak tertangani.
   d. Mengambil keputusan yang berkaitan dengan penggunaan hasil dari *data mining.*

6. *Deployment*

   Tahap pembangunan ini merupakan tahapan implementasi untuk pembangunan aplikasi berupa representasi pengetahuan

yang telah diperoleh sihingga dapat digunakan oleh pengguna



Gambar 2. 11 Phase of the CRISP-DM References Model

2.2.10 Validitas *Cluster*

Validitas *cluster* ini digunakan untuk mengetahui berapa jumlah *cluster* yang tepat, salah satu matriks yang dapat digunakan untuk memvaliditas *cluster* adalah matriks *Silhouette Index* (SI), yaitu suatu matriks yang digunakan untuk mengevaluasi *cluster* dengan cara mengukur seberapa tepat sebuah data dalam suatu *cluster*. Untuk menghitung nilai Si dari sebuah data ke-*I*, terdapat dua kompnen aitu a_i dan b_i, a_i adalah rata – rata jarak ke-*i* terhadap semua data lainya dalam suatu *cluster*, sedangkan b_i didapatkan dengan menghitung rata – rata jarak data ke-*i*, kemudian diambil nilai terkecil [6].

Persamaan yang digunakan untuk mencari nilai a_i^j

$$a_i^j = \frac{1}{m1-1}\sum_{r=1}^{m1} d(x_i^j, x_r^j) \quad \text{.....(2.6)}$$

Persamaan yang digunakan untuk mencari nilai b_i^j

$$b_i^j = \min_{\substack{n=1,\dots,k \\ n\neq j}} \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ r\neq i}}^{m_n} d(x_i^j, x_r^n) \right\}, \quad i = 1,2,\dots,m_n$$

.....(2.7)

Untuk mendapatkan *Silhouette Index* (SI) data ke-*I* menggunakan persamaan berikut :

$$SI_i^j = \frac{b_i^j - a_i^j}{\max\left\{a_i^j, b_i^j\right\}}$$

……..……(2.8)

Nilai a_i untuk mengukur seberapa mirip sebuah data dengan *cluster* yang diikutinya, nilai yang semakin kecil menandakan semakin tepatna data tersebut dalam *cluster* tersebut. Nilai b_i yang besar menandakan seberapa jeleknya data terhadap *cluster* yang lain. Nilai SI yang terdapat dalam rentang [-1..+1]. Nilai SI yang mendekati 1 menandakan bahwa data tersebut semakin tepat dalam *cluster* tersebut. Nilai SI negatif menandakan bahwa data tersebut tidak tepat berada dalam *cluster* tersebut. SI bernilai 0 berarti data tersebut posisinya berada di perbatasan antar *cluster* [6].

Untuk nilai SI dari sebuah *cluster* didapatkan dengan menggunakan rata – rata nilai SI semua data bergabung dalam *cluster* teresebut, seperti pada persamaan berikut :

$SI_j = \frac{1}{mj} \sum_{i=1}^{mj} SI_i^j$……..(2.9)

Dan untuk nilai SI global didapatkan dengan menggunakan persamaan sebagai berikut :

$SI_j = \frac{1}{k} \sum_{j=1}^{k} SI_j$……..(2.10)

2.2.11 Alat – alat Pemodelan Sistem

Pemodelan sistem merupakan hal yang penting bagi kelangsungan sistem itu sendiri. Pemodelan sistem adalah suatu upaya untuk menjaga efektivitas sistem dalam memenuhi kebutuhan pengguna sistem. Pemodelan sistem dapat bererti menyusun sistem yang baru untuk menggantikan sistem yang lama secara keseluruhan atau memperbaiki sistem yang sudah ada [9].

2.2.11.1 *Entity Relationship Diagram* (ERD)

ERD adalah gambar atau diagram yang menunjukan informasi dibuat, disimpan, dan digunakan dalam sistem bisnis. Entitas biasanya menggambarkan jenis informasi yang sama. Dalam entitas digunakan untuk menghubungkan antar entitas yang sekaligus menunjukan hubungan antara data. Pada akhirnya ERD bisa juga digunakan untuk menunjukan aturan – aturan bisnis yang ada pada sistem informasi yang akan dibangun [8].

Entity relationship diagram ini memiliki beberapa elemen sebagai berikut [8]:

1. *Entity*

   Pada ERD, entitas digambarkan dengan menggunakan sebuah bentuk persegi panjang. Entitas adalah objek nyata yang dapat dibedakan dengan objek lain, misalnya mahasiswa, buku, karyawan dan lain – lain. Entitas diberi nama dengan kata benda. Entitas ini memiliki atribut, contohnya entitas karyawan memiliki atribut : NIK, Nama, Jabatan.

2. Relasi

   Dalam ERD, relasi dapat digambarkan dalam sebuah bentuk belah ketupat. Relasi yang digunakan untuk menyatakan hubungan antara satu entitas dengan entitas yang lainya. Untuk penamaan relasi biasanya menggunakan kata kerja dasar. Misalnya entitas mahasiswa berleasi dengan entitas buku makapenamaan dari relasinya adalah meminjam, jadi ketika diagram tersebut dibaca dapat lebih mudah.

3. Atribut

   Atribut adalah sesuatu yang menjelaskan mengenai maksud entitas maupun relasi, sehingga dikatakan atribut adalah elemen dari setiap entitas dan relasi.

4. Kardinalitas

   Kardinalitas relasi digunakan untuk menunjukan maksimum tupel yang dapat berelasi dengan entitas pada entitas yang lain. Terdapat tiga macam kardinalitas relasi, yaitu :

   a. *one to one*

      Tingkat hubungan satu ke satu, dinyatakan dengan uatau kejadian pada entitas pertama, hanya mempunyai suatu hubungan dengan entitas yang kedua, dan sebaliknya.

b. *One to Many* atau *Many to One*

Tingkat hubungan satu ke banyak adalah hubungan yang menyatakan dari satu entitas tersebut memiliki banyak kejadian dengan entitas yang berelasi, dan sebaliknya.

c. *Many to many*

Tingkat hubungan banyak ke banyak terjadi apabila kejadian pada sebuah entitas akan mempunyai banyak hubungan dengan kejadian entitas lainya, baik dilihat dari sisi entitas yang pertama maupaun dilihat dari sisi yang kedua. Ketika terdapat dua entitas yang memiliki kardinalitas *many to many* maka akan menghasilkan sebuah enititas yang baru.

2.2.11.2 *Data Flow Diagram* (DFD)

Data Flow Diagram (DFD) menggambarkan aliran data atau informasi di mana di dalamnya terlihat keterkaitan di antara data yang ada. Terdapat banyak simbol – simbol yang digunakan dalam pembuatan DFD. DFD merupakan salah satu alat analisis dan teknik pemodelan terbaik untuk menggambarkan proses dan kebutuhan fungsional dari suatu sistem [10].

DFD merupakan serangkain diagram yang menggambarkan kegiatan – kegiatan yang ada dalam satu sistem. Teknik pembuatan DFD dimulai dengan menggambarkan sistem secara global dan dilanjutkan dengan analisis masing – masing bagian.

2.2.11.3 Kamus Data

Kamus data adalah suaut daftar data elemen yang terorganisir dengan definisi yang tetap dan sesuai dengan sistem, sehingga user dan analis sistem mempunyai pengertian yang samam mengenai *input*, *output* dan komponen *data store*.

Pembuatan kamus data didasarkan pada alur data yang terdapat dalam DFD (*Data Flow Diagram*). Alur data pada DFD bersifat global (hanya menunjukan nama alur datanya tanpa menunjukan sturkutr dari alur data). Untuk menunjukan struktur dari alur data secara rinci maka dibentuklah kamus data [10].

2.2.12 PHP

PHP (*Hypertext Preprocessor*), bahasa pemrograman PHP merupakan bahasa berbentuk skrip yang ditempatkan dalam server dan diproses didalam server, hasilnya dikirim ke *client* tempat *user* menggunakan browser [11].

Secara khusus, PHP dirancang untuk membentuk aplikasi web dinamis (membentuk suatu tampilan berdasarkan permintaan terkini). Pada prinsipnya PHP mempunyai fungsi yang sama dengan skrip – skrip seperti ASP (*Active Server Page*), Could Fusion, ataupun Perl. PHP adalah sebuah bahasa pemrograman yang dapat digunakan secara *command line*, atau dapat dikatakan bahwa skrip PHP dapat dijalankan dengan melibatkan web browser maupun *browser*.

Pada saat ini PHP cukup popular sebagai piranti pemrograman web, tertama dilingkungan Linux. Walaupun demikian PHP sebenarnya juga dapat digunakan pada *server – server* yang berbasis UNIX, Windowa, Macintosh.

2.2.13 MySQL

MySQL adalah suatu perangkat lunak *database* relasi (*Relational Database Management Sistem* RDBMS), seperti halnya ORACLE, Postgresql,MSSQL dan sebagainya. MySQL dengan SQL merupakan dua hal yang berbeda, SQL (*Structured Query Language*) adalah sintak perintah – perintah atau bahawa pemrograman yang digunakan untuk pengolahan data base, sedangkan MySQL itu sendiri merupakan *software* yang digunakan untuk mengolah database [11].