

BAB II

LANDASAN TEORI

2.1 Data mining

Data Mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode *Data Mining* ini dapat digunakan untuk mengambil keputusan di masa depan. *Data Mining* ini juga dikenal dengan istilah *pattern recognition* (Fajar Astuti, 2013).

Data Mining, sering juga disebut *knowledge discovery in database* (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari *Data Mining* ini biasa dipakai untuk memperbaiki pengambilan keputusan di masa depan. Sehingga istilah *pattern recognition* sekarang jarang digunakan karena ia termasuk bagian dari *Data Mining* (Santosa, 2007).

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database. *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 (Turban, 2005).

Data mining merupakan bidang dari beberapa keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar (Larose, 2006).

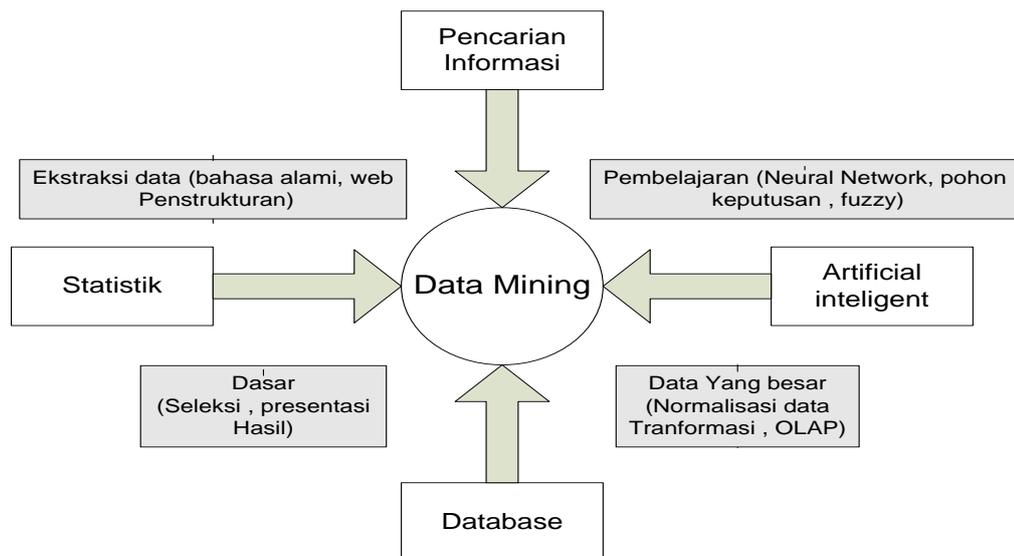
Kemajuan luar biasa yang terus berlanjut dalam bidang *data mining* didorong oleh beberapa faktor, antara lain :

1. Pertumbuhan yang cepat dalam kumpulan data.
2. Penyimpanan data dalam *data warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam database yang baik.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk *data mining* (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

Berdasarkan definisi-definisi yang telah disampaikan, hal penting yang terkait dengan *data mining* adalah :

1. *Data mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan *data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Hubungan yang dicari dalam *data mining* dapat berupa hubungan antara dua atau Pernyataan tersebut menegaskan bahwa dalam *data mining* otomatisasi tidak menggantikan campur tangan manusia. Manusia harus ikut aktif dalam setiap fase dalam proses *data mining*. Kehebatan kemampuan algoritma *datamining* yang terdapat dalam perangkat lunak analisis yang terdapat saat ini memungkinkan terjadinya kesalahan penggunaan yang berakibat fatal. Pengguna mungkin menerapkan analisis yang tidak tepat terhadap kumpulan data dengan menggunakan pendekatan yang berbeda. Oleh karenanya, dibutuhkan pemahaman tentang statistik dan struktur model matematika yang mendasari kerja perangkat lunak (Larose, 2006).



Gambar 2.1 Bidang ilmu *Data Mining*
 Sumber : Santoso, H. (2012).

Data mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefinisikan *data mining* adalah kenyataan bahwa *data mining* mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Gambar 2.1 menunjukkan bahwa *data mining* memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik, database, dan juga *information retrieval* (Pramudiono, 2007).

Istilah *data mining* dan *Knowledge Discovery in Database* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining*. Proses KDD secara garis besar dapat dijelaskan sebagai berikut :

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Pre-processing/Cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus KDD. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. *Transformation*

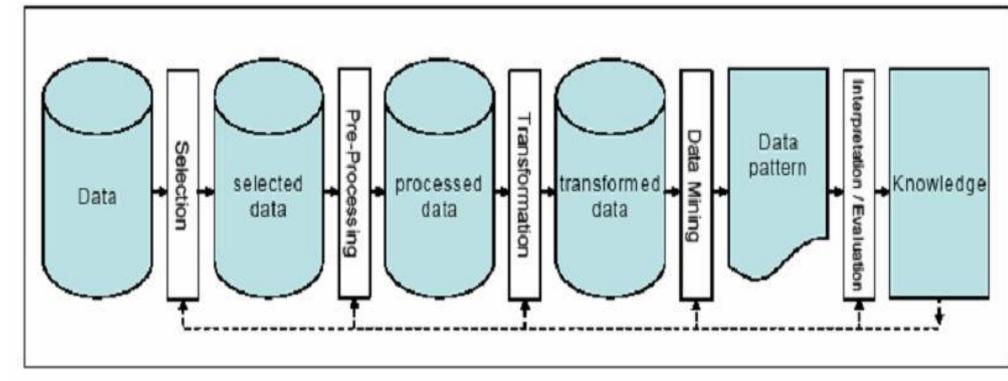
Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode dan algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation/Evaluation*

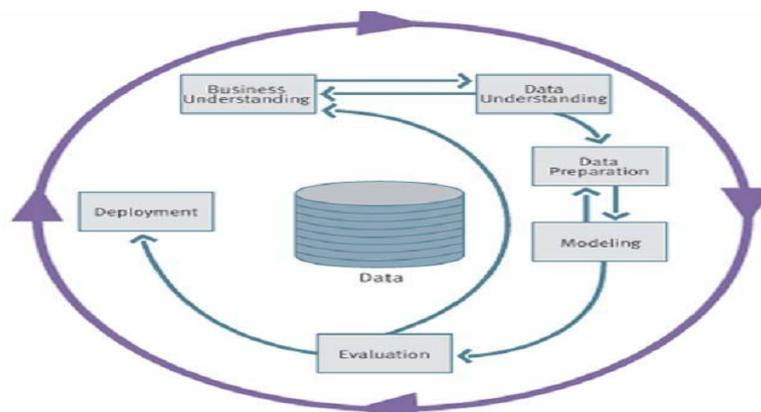
Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya. Penjelasan di atas dapat direpresentasikan pada Gambar 2.2



Gambar 2.2 Proses dari *Data Mining*
 Sumber : Santoso, H.(2012).

Cross-Industry Standart Process for Data Mining (CRISP-DM) yang dikembangkan tahun 1996 oleh analisis dari beberapa industri seperti Daimler Chrysler, SPSS dan NCR. CRISP-DM menyediakan standar proses *data mining* sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian.

Dalam CRISP-DM sebuah proyek *data mining* memiliki siklus hidup yang terbagi dalam enam fase Gambar 2.3. Keseluruhan fase berurutan yang ada tersebut bersifat adaptif. Fase berikutnya dalam urutan bergantung kepada keluaran dari fase sebelumnya. Hubungan penting antar fase digambarkan dengan panah. Sebagai contoh, jika proses berada pada fase *modeling*. Berdasar pada perilaku dan karakteristik model, proses mungkin kembali kepada fase *datapreparation* untuk perbaikan lebih lanjut terhadap data atau berpindah majukepada fase *evaluation*.



Gambar 2.3 Proses *Data Mining* Menurut CRISP-DM
 Sumber: Santoso, H. (2012).

Enam fase CRISP-DM (*Cross Industry Standard Process for DataMining*) (Larose, 2006).

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)
 - a. Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkupbisnis atau unit penelitian secara keseluruhan.
 - b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan *data mining*.
 - c. Menyiapkan strategi awal untuk mencapai tujuan.

2. Fase Pemahaman Data (*Data Understanding Phase*)
 - a. Mengumpulkan data.
 - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
 - c. Mengevaluasi kualitas data.
 - d. Jika diinginkan, pilih sebagian kecil kelompok data yang mungkin mengandung pola dari permasalahan.

3. Fase Pengolahan Data (*Data Preparation Phase*)
 - a. Siapkan dari data awal, kumpulan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
 - b. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.
 - c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
 - d. Siapkan data awal sehingga siap untuk perangkat pemodelan.

4. Fase Pemodelan (*Modeling Phase*)
 - a. Pilih dan aplikasikan teknik pemodelan yang sesuai.
 - b. Kalibrasi aturan model untuk mengoptimalkan hasil.
 - c. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan *data mining* yang sama.
 - d. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk

menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik *data mining* tertentu.

5. Fase Evaluasi (*Evaluation Phase*)

- a. Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektivitas sebelum disebarkan untuk digunakan.
- b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
- c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
- d. Mengambil keputusan berkaitan dengan penggunaan hasil dari *datamining*.

6. Fase Penyebaran (*Deployment Phase*)

- a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.
- b. Contoh sederhana penyebaran: Pembuatan laporan.
- c. Contoh kompleks Penyebaran: Penerapan proses *data mining* secara paralel pada departemen lain. Informasi lebih lanjut mengenai CRISP-DM dapat dilihat di www.crisp-dm.org

2.2 Pengelompokan *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose, 2006).

1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan

kecendrungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecendrungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai 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, 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 di masa mendatang.

Contoh prediksi dalam bisnis dan penelitian adalah:

- a. Prediksi harga beras dalam tiga bulan yang akan datang.
- b. Prediksi presentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan.

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 mahal, pendapatan sedang, dan pendapatan murah.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah:

- a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.

- b. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- c. Mendiagnosa penyakit seorang pasien untuk mendapatkan termasuk kategori apa.

5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *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 (homogen), yang mana 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 suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
- c. Melakukan pengklusteran terhadap ekspresi dari gen, dalam jumlah besar.

6. Asosiasi

Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam suatu 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. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli bersamaan.

2.3 Data

Dalam *Webster's New World's Dictionary* tertulis bahwa datum: *something known or assumed*. Artinya, datum (bentuk tunggal data) merupakan suatu yang diketahui / dianggap. Dengan demikian, data dapat memberi gambaran tentang suatu keadaan atau persoalan. Sedangkan, data menurut kamus Oxford Dictionary adalah *The Facts*. Jadi, dapat disimpulkan bahwa data adalah sesuatu yang nyata diketahui atau dianggap yang dipakai untuk keperluan suatu analisa, diskusi, presentasi ilmiah atau tes statistic. Data dapat dibagi menjadi 4 bagian, yaitu: (Ong, 2013).

2.3.1 Jenis – Jenis Data Berdasarkan Sifatnya

Jenis-jenis data dapat dibagi menurut sifatnya, menurut sumbernya, menurut cara memperolehnya dan menurut waktu pengumpulannya. Menurut sifatnya data dapat terbagi menjadi dua jenis, yaitu data kualitatif (non-metrik) dan data kuantitatif (metrik). Kemudian jenis data kualitatif terbagi lagi menjadi dua jenis, yaitu data nominal dan data ordinal. Begitu pula dengan jenis data kuantitatif terbagi menjadi dua jenis, yaitu data interval dan data rasio (Hidayat & Istiadah 2011).

a. Data Kualitatif

Data kualitatif secara sederhana dapat disebut data yang bukan berupa angka. Ciri utama data kualitatif didapat dengan cara menghitung, sehingga tidak memiliki nilai desimal. Selain itu data kualitatif memiliki ciri-ciri tidak bisa dilakukan operasi matematika, seperti penambahan, pengurangan, perkalian dan pembagian.

Contoh data kualitatif adalah data gender, data golongan darah, data tempat tinggal atau data jenis pekerjaan. Agar dapat dilakukan proses pada data kualitatif atau nonmetric, data tersebut harus diubah ke

dalam bentuk angka, proses ini dinamakan kategorisasi. Data kualitatif dibedakan menjadi dua jenis, yaitu data nominal dan data ordinal (Santoso, 2010).

1. Data Nominal adalah jenis data kualitatif yang digunakan mengidentifikasi, mengklasifikasi, atau membedakan objek. Data nominal merupakan data yang paling rendah dalam level pengukuran data. Semua data memiliki posisi setara dalam arti tidak ada data yang memiliki tingkat yang lebih atau kurang dibandingkan dengan data yang lain. Jenis data nominal ini tidak memiliki jarak, urutan dan titik origin (Hidayat & Istiadah, 2011).
2. Data Ordinal adalah jenis data kualitatif namun memiliki level lebih tinggi dari data nominal. Data ordinal memiliki karakteristik nominal tapi terdapat perbedaan derajat, urutan, atau peringkat dalam objek tersebut (posisi data tidak setara) (Hidayat & Istiadah 2011).

b. Data Kuantitatif

Data kuantitatif dapat disebut sebagai data berupa angka dalam arti sebenarnya. Jadi, berbagai jenis operasi matematika dapat dilakukan pada data kuantitatif. Data kuantitatif merupakan data yang didapat dengan jalan mengukur sehingga bisa mempunyai nilai desimal. Contoh data kuantitatif adalah tinggi badan, usia, penjualan barang, dan sebagainya. Sebagai contoh, tinggi badan seseorang bisa bernilai 165 cm atau 165.5 cm. Seperti pada jenis data kualitatif, jenis data kuantitatif juga terbagi menjadi dua, yaitu data interval dan data rasio (Santoso, 2010).

2.3.2 Jenis – Jenis Data Menurut Sumbernya

Pembagian jenis data menurut sumbernya didasarkan pada sumber perolehan data tersebut, yaitu data internal dan data eksternal (Ong, 2013).

- a. Data internal adalah data yang dikumpulkan oleh suatu organisasi untuk menggambarkan keadaan atau kegiatan organisasi yang bersangkutan serta berguna untuk keperluan kegiatan harian dan pengawasan internal. Misalnya, data penjualan, data produksi suatu perusahaan, data keuangan, data kepegawaian, dan lain sebagainya.
- b. Data eksternal adalah data yang dikumpulkan untuk menggambarkan suatu keadaan atau kegiatan di luar organisasi tersebut. Contoh dari data eksternal seperti data jumlah penduduk dan data pendapatan nasional yang didapat dari kantor pusat statistik setempat. Suatu perusahaan memerlukan data eksternal seperti jumlah penduduk untuk memprediksi potensi permintaan, sedangkan data pendapatan nasional untuk menentukan tingkat daya beli masyarakat yang berguna untuk dasar kebijakan tingkat harga.

2.3.3 Jenis Data Menurut Cara Memperolehnya

Berdasarkan cara memperolehnya, data dapat dibedakan menjadi dua jenis, yaitu data primer dan data sekunder (Ong, 2013).

- a. Data primer adalah data yang dikumpulkan dan diolah sendiri oleh organisasi atau perorangan langsung dari objeknya. Misalnya suatu perusahaan ingin mengetahui konsumsi rata-rata suatu produk terhadap penduduk di suatu daerah dengan cara melakukan wawancara langsung kepada penduduk setempat.
- b. Data sekunder adalah data yang diperoleh dalam bentuk jadi dan telah diolah oleh pihak lainnya. Biasanya data sekunder ini dalam bentuk publikasi.

2.3.4 Jenis Data Menurut Waktu Pengumpulannya

Berdasarkan waktu pengumpulannya, data dapat dibedakan menjadi dua jenis, yaitu data cross section dan data berkala (time series) (Ong, 2013).

- a. Data cross section adalah data yang dikumpulkan dalam suatu periode tertentu, biasanya menggambarkan keadaan atau kegiatan dalam

periode tersebut. Misalnya, hasil sensus penduduk tahun 2014 menggambarkan keadaan Indonesia pada tahun 2014 menurut, umur, jenis kelamin, agama, tingkat pendidikan, dan sebagainya.

- b. Data berkala (time series) adalah data yang dikumpulkan dari waktu ke waktu. Tujuannya adalah untuk menggambarkan perkembangan suatu kegiatan dari waktu ke waktu. Misalnya, perkembangan produksi di suatu perusahaan selama lima tahun terakhir, perkembangan penjualan produk selama lima tahun terakhir, dan lain sebagainya. Jenis data ini juga sering disebut sebagai data historis.

2.4 Metode Sampling

Penarikan contoh (*Sampling*) merupakan teknik utama yang di gunakan untuk seleksi data. Proses ini di gunakan untuk persiapan penyelidikan dan analisis data akhir. Penarikan contoh di gunakan dalam data mining karena pengolahan himpunan data yang di inginkan secara mahal atau menghabiskan waktu.

Prinsip utama untuk penarikan contoh yang efektif adalah sebagai berikut :

- a. Menggunakan suatu contoh yang akan dapat bekerja hampir sebaik penggunaan data- set keseluruhan, jika contoh tersebut representatif (bersifat mewakili).
- b. Suatu contoh dikatakan representatif jika paling tidak mempunyai sifat yang sama (dengan yang diperhatikan) sebagai himpunan data asli.

Ada beberapa tipe penarikan contoh yaitu :

- a. Penarikan contoh acak sederhana. Kemungkinan ada suatu pemilihan item tertentu yang sama.
- b. Penarikan contoh tanpa penggantian (replacement). Setiap item yang dipilih, dikeluarkan dari populasinya.
- c. Penarikan contoh dengan penggantian (replacement). Objek –objek tidak dikeluarkan dari populasi saat objek tersebut dipilih untuk sample. Dalam penarikan contoh dengan penggantian, objek yang sama dapat diambil lebih dari satu kali.

- d. Penarikan contoh bertingkat (stratified sampling).membagi data dalam beberapa bagian (partisi), kemudian menarik contoh dari tiap partisi.

Ukuran dari sample sangat menentukan keakuratan dari pengambilan kesimpulan jika diterapkan sebuah algoritma data mining. Semakin besar ukuran sample maka hasilnya akan semakin akurat. Namun demikian perlu dipertimbangkan juga waktu komputasi. Untuk itu perlu ditentukan ukuran yang tepat dan tidak terlalu besar. Untuk sample 8000 titik ,objek terlihat sangat jelas. Jumlah sample 500, bentuk objek sama sekali tidak terlihat. Sedangkan untuk sample 2000, bentuk objek cukup terlihat. Sehingga, untuk mendapatkan hasil objek yang di inginkan dengan jumlah sample tidak terlalu besar [mengurangi waktu komputasi],jumlah sample 2000 bisa di pilih.

2.5 Laptop / Komputer

Menurut Hamacher, Vranesic, & Zaky (Wimatra,2008) mengatakan bahwa komputer didefinisikan sebagai sebuah mesin penghitung elektronik yang cepat dapat menerima informasi input digital, memprosesnya sesuai dengan suatu program yang tersimpan di memorinya (stored program) dan menghasilkan output informasi.

2.5.1 Elemen – Elemen yang Terdapat pada Laptop/ Komputer

Seperti yang sudah disebutkan sebelumnya, (Wimatra, (2008)) membuat sebuah rangkuman mengenai definisi dan juga elemen – elemen yang dapat mendukung bekerjanya sebuah sistem komputer. Ketiga elemen tersebut adalah :

A. Brainware

Brainware merupakan elemen pertama dari sebuah komputer. Brainware mengacu pada user alias manusia yang menggunakan dan juga mengoperasikan komputer tersebut. Brainware sendiri sering dikategorikan menjadi dua jenis, yaitu :

a. Laterat

Laterat mengacu pada user atau brainware yang paham mengenai seluk beluk komputer. Biasanya adalah seorang programmer, teknisi komputer, dan juga ahli – ahli komputer.

b. Kompeten

Kompeten merujuk pada brainware alias user komputer yang bisa melakukan pengopereasian dari sebuah komputer. Tidak perlu memahami hingga seluk beluk komputer, yang penting dapat mengoperasikannya.

B. Hardware

Hardware komputer atau perangkat keras komputer merupakan sekumpulan perangkat keras yang memiliki bentuk fisik dan secara nyata bisa dilihat dan dipegang. Lebih lengkapnya, anda bisa membaca mengenai perangkat keras komputer.

Berikut ini adalah beberapa contoh hardware :

a. Processor

b. RAM Komputer

c. VGA

d. Harddisk

e. Keyboard

f. Mouse

g. Monitor

C. Software

Software atau perangkat lunak mengacu pada sebuah sistem yang dapat membantu menjalankan sebuah fungsi dari masing – masing hardware komputer, serta membantu usernya dalam menjalankan sebuah komputer.

Software bisa terdiri dari :

a. System software, seperti Windows, Linux, Mac

b. Programming software, seperti software – software untuk membuat program

c. Application software seperti aplikasi word, games, dan banyak lagi.

2.5.2 Manfaat laptop/ Komputer

Laptop/ Komputer sendiri bagi kita sudah pasti memiliki banyak sekali manfaat.

Berikut ini adalah beberapa manfaat komputer :

- a. Membantu mempermudah pekerjaan, seperti melakukan pengetikan dokumen, pencatatan data, dan sebagainya.
- b. Dapat menjadi sumber informasi.
- c. Membantu menyediakan hiburan bagi usernya.
- d. Bisa saling terhubung dengan menggunakan konsep jaringan computer.
- e. Dapat melakukan koneksi terhadap jaringan internet.

2.6 Klasifikasi Laptop

Dalam memudahkan customer dalam mencari barang berupa laptop yang diinginkan maka di klasifikasikan sesuai dengan kategori masing - masing laptop secara otomatis. Jika laptop dapat melakukan klasifikasi laptop secara otomatis maka dapat menghemat waktu yang diperlukan mengingat jumlah laptop pada toko computer dapat bertambah banyak dalam waktu yang singkat sehingga pencarian laptop pada toko tersebut lebih mudah.

Selain dapat mengefisiensi penggunaan waktu, klasifikasi laptop juga dapat membuat system pada toko computer tersebut terlihat rapi karena laptop di klasifikasikan berdasarkan kategori yang sesuai dengan hardware dan software. Hal ini juga dapat membantucustomer yang ingin mencari barang atau produk yang diinginkannya secara mudah. Seiring dengan perkembangan teknologi saat ini, jumlah laptop di Indonesia mengalami peningkatan yang cukup pesat dari tahun ke tahun. Hasil survey menunjukkan jumlah pengguna laptop di Indonesia pada tahun 2011 mencapai sekitar 40 juta dan saat ini terdapat sekitar 63 juta pengguna.

Perkembangan teknologi dan internet memang sangat cepat dan memberi pengaruh signifikan dalam segala aspek kehidupan masyarakat. Laptop membantu masyarakat untuk mempermudah pekerjaan baik berupa tugas ataupun pekerjaan kantor, berkomunikasi, serta mendapatkan informasi mengenai apapun dengan mudah dan cepat. Seiring dengan berjalannya waktu serta diiringi perkembangan

teknologi, laptop kini dapat dimanfaatkan sebagai sarana untuk melakukan transaksi perdagangan yang dikenal dengan istilah jual beli online.

2.7 Harga (Price)

Harga merupakan salah satu bagian yang sangat penting dalam pemasaran suatu produk karena harga adalah satu dari empat bauran pemasaran / marketing mix (4P = product, price, place, promotion / produk, harga, distribusi, promosi). Harga adalah suatu nilai tukar dari produk barang maupun jasa yang dinyatakan dalam satuan moneter.

Harga merupakan salah satu penentu keberhasilan suatu perusahaan karena harga menentukan seberapa besar keuntungan yang akan diperoleh perusahaan dari penjualan produknya baik berupa barang maupun jasa. Menetapkan harga terlalu tinggi akan menyebabkan penjualan akan menurun, namun jika harga terlalu rendah akan mengurangi keuntungan yang dapat diperoleh organisasi perusahaan.

2.7.1 Metode Penetapan Harga Produk

Ada beberapa metode penerapan harga produk sebagai berikut:

1. Pendekatan Permintaan dan Penawaran (supply demand approach).

Dari tingkat permintaan dan penawaran yang ada ditentukan harga keseimbangan (equilibrium price) dengan cara mencari harga yang mampu dibayar konsumen dan harga yang diterima produsen sehingga terbentuk jumlah yang diminta sama dengan jumlah yang ditawarkan.

2. Pendekatan Biaya (cost oriented approach).

Menentukan harga dengan cara menghitung biaya yang dikeluarkan produsen dengan tingkat keuntungan yang diinginkan baik dengan markup pricing dan break even analysis.

3. Pendekatan Pasar (market approach).

Merumuskan harga untuk produk yang dipasarkan dengan cara menghitung variabel-variabel yang mempengaruhi pasar dan harga seperti situasi dan kondisi politik, persaingan, sosial budaya, dan lain-lain.

2.8 Naive Bayes Classifier

Naive Bayes Classifier adalah metode pengklasifikasian dengan model peluang, dimana diasumsikan bahwa setiap atribut contoh (data sampel) bersifat saling lepas satu sama lain berdasarkan atribut kelas. Metode klasifikasi ini diturunkan dari penerapan teorema Bayes dengan asumsi independence (saling bebas). Naive Bayes Classifier merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Ciri utama dari Naive Bayes Classifier ini adalah asumsi yang sangat kuat (naif) akan independensi dari masing - masing kondisi/kejadian. Sebelum menjelaskan Naive Bayes Classifier ini, akan dijelaskan terlebih dahulu Teorema Bayes yang menjadi dasar dari metode tersebut. Pada teorema Bayes, bila terdapat dua kejadian yang terpisah.

Menurut Olson Delen (2008) menjelaskan Naive Bayes untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vektor informasi obyek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari "master" tabel keputusan.

Naive Bayes Classifier bekerja sangat baik dibanding dengan model classifier lainnya. Hal ini dibuktikan oleh Xhemali, Hinde Stone dalam jurnalnya "Naive Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages" mengatakan bahwa "Naive Bayes Classifier memiliki tingkat akurasi yang lebih baik dibanding model classifier lainnya".

Algoritma Naive Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasi dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu

memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak adahubungannya dengan ciri dari kelas lainnya.

Persamaan dari teorema *Bayes* adalah :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Di mana :

X : Data dengan class yang belum diketahui.

H : Hipotesis data merupakan suatu class spesifik.

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posterior probabilitas).

$P(H)$: Probabilitas hipotesis H (prior probabilitas).

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis.

Keuntungan penggunaan Metode Naïve Bayes Classifier adalah bahwa metoda ini hanya membutuhkan jumlah data pelatihan (training data) yg kecil unt menentukan estimasi parameter yg diperlukan dalam proses pengklasifikasian. Karena yg diasumsikan sebagai variable independent, maka hanya varians dr suatu variable dalam sebuah kelas yg dibutuhkan unt menentukan klasifikasi, bukan keseluruhan dari matriks kovarians.

2.8 PHP

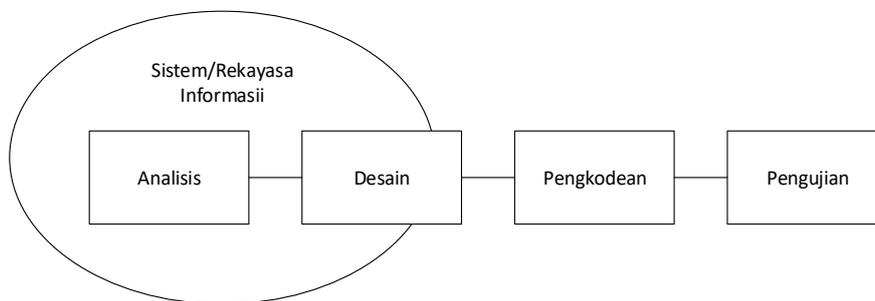
Madcoms (2008, p.195) PHP adalah salah satu bahasa pemrograman yang berjalan dalam sebuah *web server* dan berfungsi sebagai pengolahan data pada sebuah *server*. Untuk membuat *website* yang dinamis dan mudah untuk *diupdate* setiap saat dari *browser*, dibutuhkan sebuah program yang mampu mengolah data dari komputer *client* atau dari komputer *server* itu sendiri sehingga mudah dan nyaman untuk disajikan di *browser*.

2.9 MySQL

(Simarmata, J. 2006) MySQL adalah suatu *database* populer dengan pengembangan *Web*(*Web Developers*). Kecepatan dan ukuran yang kecil membuatnya ideal untuk *Web Site*. Ditambah lagi dengan fakta bahwa MySQL adalah *open source*, yang berarti gratis. (Simarmata, J. 2006) *software* MySQL terdiri dari MySQL Server, beberapa program utilitas membantu administrasi MySQL *Database*, dan beberapa *software* pendukung yang dibutuhkan MySQL Server

2.10 Metode Pengembangan Perangkat Lunak

Rosa dan Shalahuddin (2013), Model SDLC air terjun (*waterfall*) sering juga disebut model eksekusional linier (*sequential linear*) atau alur hidup klasik (*classiclifecycle*). Model air terjun menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau terurut dimulai dari analisis, desain, pengkodean, pengujian dan tahap pendukung (*support*). Berikut ini adalah gambar model air terjun :



Gambar 2.4 Metode *waterfall*
Sember : Rosa dan Shalahuddin (2013)

2.10.1 Analisis kebutuhan perangkat lunak

Proses pengumpulan kebutuhan dilakukan secara intensif untuk mespesifikasikan kebutuhan perangkat lunak agar dapat dipahami perangkat lunak seperti apa yang dibutuhkan oleh user. Spesifikasi kebutuhan perangkat lunak pada tahap ini perlu untuk didokumentasi.

2.10.2 Desain

Desain perangkat lunak adalah proses multi langkah yang *focus* pada desain pembuatan program perangkat lunak termasuk struktur data, arsitektur perangkat lunak, representasi antar muka, dan prosedur pengkodean.

2.10.3 Pembuatan kode program

Desain harus ditranslasikan ke dalam program perangkat lunak. Hasil dari tahap ini adalah program komputer sesuai dengan desain yang telah dibuat pada tahap desain.

2.10.4 Pengujian

Pengujian fokus pada perangkat lunak secara dari segi logik dan fungsional dan memastikan bahwa semua bagian sudah diuji. Hal ini dilakukan untuk meminimalisir kesalahan dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan.

2.10.5 Pendukung (*support*) atau pemeliharaan (*maintenance*)

Tidak menutup kemungkinan sebuah perangkat lunak mengalami perubahan ketika sudah dikirim ke *user*. Perubahan bisa terjadi karena adanya kesalahan yang muncul dan tidak terdeteksi saat pengujian atau perangkat lunak harus beradaptasi dengan lingkungan baru. Tahap pendukung dan pemeliharaan dapat mengulangi proses pengembangan mulai dari analisis spesifikasi untuk perubahan perangkat lunak yang sudah ada, tapi tidak untuk membuat perangkat lunak baru.