

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **II.1. Sistem Informasi**

##### **II.1.1. Sistem**

Sistem adalah sekumpulan komponen yang saling berhubungan, bekerja sama untuk mencapai suatu tujuan dengan menerima input serta menghasilkan output dalam proses transformasi yang teratur .

Sistem adalah sebuah tatanan yang terdiri atas sejumlah komponen fungsional (dengan tugas/fungsi khusus) yang saling berhubungan dan secara bersama-sama bertujuan untuk memenuhi suatu proses/pekerjaan tertentu. Sebagai contoh, sistem kendaraan terdiri dari: komponen starter, komponen pengapian, komponen penggerak, komponen pengerem, komponen kelistrikan- speedometer, lampu dan lain-lain. Komponen-komponen tersebut diatas memiliki tujuan yang sama yaitu untuk membuat kendaraan tersebut bisa dikendarai dengan nyaman dan aman. Contoh lain yaitu sistem perguruan tinggi, yang terdiri dari dosen, mahasiswa, kurikulum, dan lain-lain. Sistem ini bertujuan untuk menghasilkan mahasiswa-mahasiswa yang memiliki kemampuan di bidang ilmunya. (Kusrini ; 2008: 4)

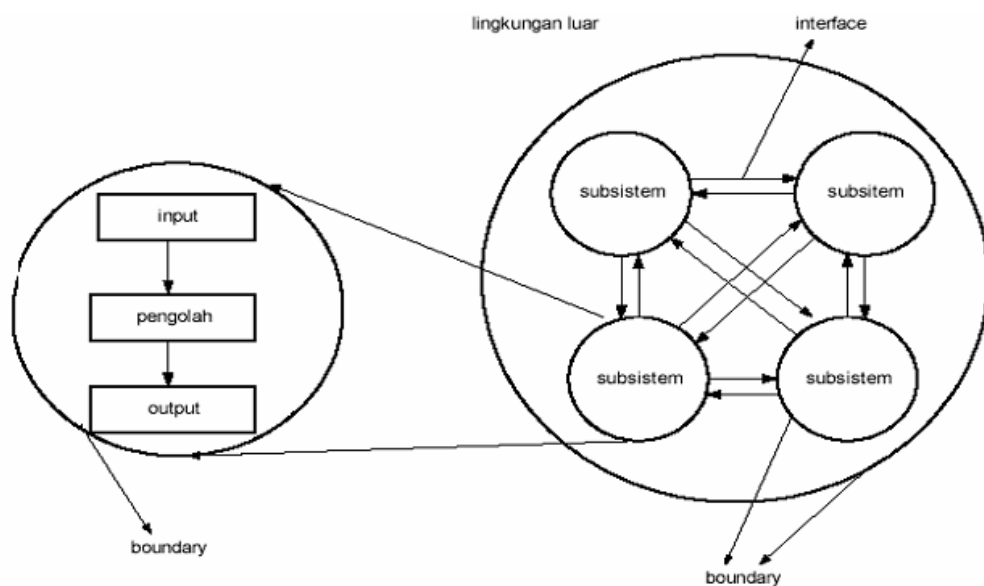
##### **II.1.2. Konsep Dasar Sistem**

Sistem adalah kumpulan dari elemen-elemen yang berinteraksi untuk mencapai suatu tujuan tertentu. Menurut Jerry FithGerald ; sistem adalah suatu

jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan suatu kegiatan atau menyelesaikan suatu sasaran tertentu. (Kusrini ; 2008: 4)

### II.1.3. Karakteristik Sistem

Karakteristik sistem dapatlah digambarkan sebagai berikut :



**Gambar II.1. Karakteristik Sistem**  
(Sumber : Kusrini; 2008:5)

#### 1. Komponen Sistem (*Components*)

Suatu sistem terdiri dari sejumlah komponen yang saling berinteraksi, yang artinya saling bekerja sama membentuk satu kesatuan. Komponen-komponen sistem atau elemen-elemen sistem dapat berupa suatu subsistem atau bagian-bagian dari sistem. Setiap sistem tidak peduli betapapun kecilnya, selalu mengandung komponen-komponen atau subsistem-subsistem. Setiap subsistem mempunyai sifat-sifat dari sistem untuk menjalankan suatu fungsi tertentu dan mempengaruhi proses sistem

secara keseluruhan. Jadi, dapat dibayangkan jika dalam suatu sistem ada subsistem yang tidak berjalan/berfungsi sebagaimana mestinya. Tentunya sistem tersebut tidak akan berjalan mulus atau mungkin juga sistem tersebut rusak sehingga dengan sendirinya tujuan sistem tersebut tidak tercapai.

## 2. Batas Sistem (*Boundary*)

Batas sistem (*boundary*) merupakan daerah yang membatasi antara suatu sistem dengan sistem yang lainnya atau dengan lingkungan luarnya. Batas sistem ini memungkinkan suatu sistem dipandang sebagai satu kesatuan. Batas suatu sistem menunjukkan ruang lingkup (*scope*) dari sistem tersebut.

## 3. Lingkungan Luar Sistem (*Environments*)

Lingkungan luar dari suatu sistem adalah apapun diluar batas dari sistem yang mempengaruhi operasi sistem. Lingkungan luar sistem dapat bersifat menguntungkan dan dapat juga bersifat merugikan sistem tersebut. Lingkungan luar yang menguntungkan merupakan energi dari sistem dan dengan demikian harus tetap dijaga dan dipelihara. Sedang lingkungan luar yang merugikan harus ditahan dan dikendalikan, kalau tidak maka akan mengganggu kelangsungan hidup dari sistem.

## 4. Penghubung (*Interface*)

Penghubung merupakan media penghubung antara satu subsistem dengan subsistem lainnya. Melalui penghubung ini memungkinkan sumber-sumber daya mengalir dari satu subsistem ke yang lainnya.

Keluaran (output) dari satu subsistem akan menjadi masukan (input) untuk subsistem lainnya dengan melalui penghubung. Dengan penghubung satu subsistem dapat berintegrasi dengan subsistem yang lainnya membentuk satu kesatuan.

#### 5. Masukan (*Input*)

Masukan adalah energi yang dimasukkan ke dalam sistem. Masukan dapat berupa masukan perawatan (*maintenance input*) dan masukan sinyal (*signal input*). Maintenance input adalah energi yang dimasukkan supaya sistem tersebut dapat beroperasi. Signal input adalah energi yang diproses untuk didapatkan keluaran. Sebagai contoh didalam sistem komputer, program adalah *maintenance input* yang digunakan untuk mengoperasikan komputernya dan data adalah signal input untuk diolah menjadi informasi.

#### 6. Keluaran (*Output*)

Keluaran adalah hasil dari energi yang diolah dan diklasifikasikan menjadi keluaran yang berguna dan sisa pembuangan. Keluaran dapat merupakan masukan untuk subsistem yang lain atau kepada super sistem. Misalnya untuk sistem komputer, panas yang dihasilkan adalah keluaran yang tidak berguna dan merupakan hasil sisa pembuangan, sedang informasi adalah keluaran yang dibutuhkan.

#### 7. Pengolah (*Process*)

Suatu sistem dapat mempunyai suatu bagian pengolah yang akan merubah masukan menjadi keluaran. Suatu sistem produksi akan mengolah masukan berupa bahan baku dan bahan-bahan yang lain menjadi keluaran

berupa barang jadi. Sistem akuntansi akan mengolah data-data transaksi menjadi laporan-laporan keuangan dan laporan-laporan lain yang dibutuhkan oleh manajemen.

#### 8. Sasaran (*Objectives*) atau Tujuan (*Goal*)

Suatu sistem pasti mempunyai tujuan atau sasaran. Kalau suatu sistem tidak mempunyai sasaran, maka operasi sistem tidak akan ada gunanya. Sasaran dari sistem sangat menentukan sekali masukan yang dibutuhkan sistem dan keluaran yang akan dihasilkan sistem. Suatu sistem dikatakan berhasil bila mengenai sasaran atau tujuannya. Perbedaan suatu sasaran (*objectives*) dan suatu tujuan (*goal*) adalah, *goal* biasanya dihubungkan dengan ruang lingkup yang lebih luas dan sasaran dalam ruang lingkup yang lebih sempit. Bila merupakan suatu sistem utama, seperti misalnya sistem bisnis perusahaan, maka istilah *goal* lebih tepat diterapkan. Untuk sistem akuntansi atau sistem-sistem lainnya yang merupakan bagian atau subsistem dari sistem bisnis, maka istilah *objectives* yang lebih tepat. Jadi tergantung dari ruang lingkup mana memandang sistem tersebut. Seringkali tujuan (*goal*) dan sasaran (*objectives*) digunakan bergantian dan tidak dibedakan. (Kusrini ; 2008: 8)

#### **II.1.4. Informasi**

Informasi adalah data yang dapat dianalogikan dengan data – data , yang belum di kelolah dan harus diolah untuk menjadi informasi yang akurat. Agar informasi yang penulis sajikan lebih bermanfaat maka terlebih dahulu dibuat

aliran informasi yang lebih jelas dan lengkap. Berkaitannya dengan penyedia informasi bagi manajemen dalam mengambil suatu keputusan, yang diperoleh harus berkualitas, maka kualitas dari informasi tergantung :

1. Akurat

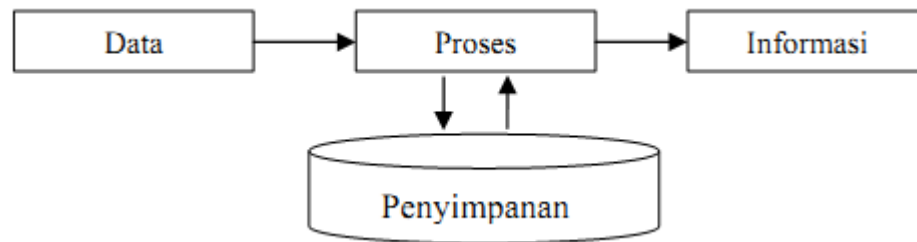
Akurat berarti bahwa informasi harus bebas dari kesalahan - kesalahan dan tidak biasa ( menyesatkan ) dan jelas mencerminkan maksudnya. Informasi harus akurat karena dari sumber informasi sampai ke penerimaan informasi kemungkinan banyak terjadi gangguan yang dapat merubah informasi atau merusak informasi tersebut.

2. Relevansi

Relevansi berarti bahwa informasi benar – benar berguna bagi suatu tindakan dan keputusan oleh seseorang

3. Tepat waktu

Tepat waktu berarti bahwa informasi yang datang pada penerimaan tidak boleh terlambat. Informasi yang sudah usang tidak akan mempunyai nilai lagi. Karena informasi merupakan landasan dalam pengambilan keputusan. Dewasa ini mahalnya nilai informasi disebabkan harus cepatnya informasi itu di dapat, Untuk lebih jelasnya informasi merupakan hasil atau output dari proses informasi data (Kusrini ; 2008: 15) . Hal ini dapat dilihat seperti gambar dibawah ini :



**Gambar II.2. Proses Data Menjadi Informasi**  
*(Sumber : Kusrini ; 2008:15)*

Sedangkan Sistem Informasi adalah kumpulan suatu sistem jaringan yang menghubungkan beberapa jaringan, yang mana telah dirangkai dengan sempurna dengan suatu struktur yang terbentuk untuk melaksanakan suatu kegiatan atau fungsi utama suatu perusahaan (Vervolt Keith ; 2003: 105).

Sistem informasi adalah sistem di dalam suatu organisasi yang mempertemukan kebutuhan pengolahan transaksi harian, membantu dan mendukung kegiatan operasi, bersifat manajerial dari suatu organisasi dan membantu mempermudah penyediaan laporan yang diperlukan. Sistem informasi dapat juga dianggap sebagai suatu departemen dan diperlukan pemisahan fungsi, dimana fungsi – fungsi tersebut dapat terdiri dari sistem analis, programmer, operator, yang dibawah oleh manajer sistem informasi (Zaki Baridwan ; 2009: 19).

Berdasarkan definisi di atas dapat disimpulkan bahwa sistem informasi adalah sekumpulan prosedur organisasi yang pada saat dilaksanakan akan memberikan informasi bagi pengambilan keputusan atau untuk pengendalian organisasi. Jadi sistem informasi himpunan dari berbagai sumber daya dalam organisasi yang

berinteraksi untuk menghasilkan informasi yang menunjang pengambilan keputusan operasi, manajerial, strategik.

## **II.2. Sistem Bisnis Cerdas**

### **II.2.1. Data Mining**

*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 et al ; 2005 ). Menurut Gartner Group *data mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose ; 2006).

Selain definisi di atas beberapa definisi juga diberikan seperti, “*data mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.” (Pramudiono; 2006). “*Data mining* adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya.”(Pramudiono; 2006).

“*Data mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data.” (Larose; 2006). “*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 : (Larose; 2006)

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 lebih dalam satu dimensi. Misalnya dalam dimensi produk, dapat di lihat keterkaitan pembelian suatu produk dengan produk yang lain. Selain itu, hubungan juga dapat dilihat antara dua atau lebih atribut dan dua atau lebih objek. (Ponniah; 2001).

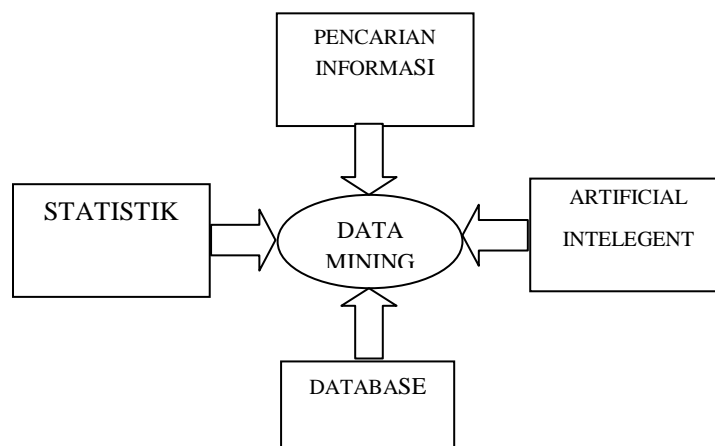
Sementara itu, penemuan pola merupakan keluaran lain dari *data mining*. Misalkan sebuah perusahaan yang akan meningkatkan fasilitas kartu kredit dari pelanggan, maka perusahaan akan mencari pola dari pelanggan-pelanggan yang ada untuk mengetahui pelanggan yang potensial dan pelanggan yang tidak potensial.

Beberapa definisi awal dari *data mining* meyertakan fokus pada proses otomatisasi. Berry dan Linoff, (2004) dalam buku *Data Mining Technique for Marketing, Sales, and Customer Support* mendefinisikan *data mining* sebagai

suatu proses eksplorasi dan analisis secara otomatis maupun semi otomatis terhadap data dalam jumlah besar dengan tujuan menemukan pola atau aturan yang berarti (Larose; 2006).

Tiga tahun kemudian, dalam buku *Mastering Data Mining* mereka memberikan definisi ulang terhadap pengertian *data mining* dan memberikan pernyataan bahwa “jika ada yang kami sesalkan adalah frasa secara otomatis maupun semi otomatis, karena kami merasa hal tersebut memberikan focus berlebih pada teknik otomatis dan kurang pada eksplorasi dan analisis”. Hal tersebut memberikan pemahaman yang salah bahwa *data mining* merupakan produk yang dapat dibeli dibandingkan keilmuan yang harus dikuasai (Larose; 2006).

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 data mining 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 II.3. Bidang Ilmu Data Mining**

*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 II.3 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; 2006).

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 (Fayyad; 1996).

### 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.

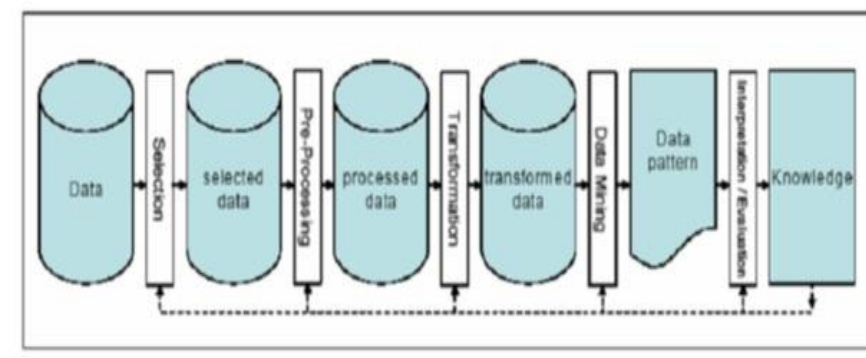
### 4. *Data mining*

*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 II.4 :

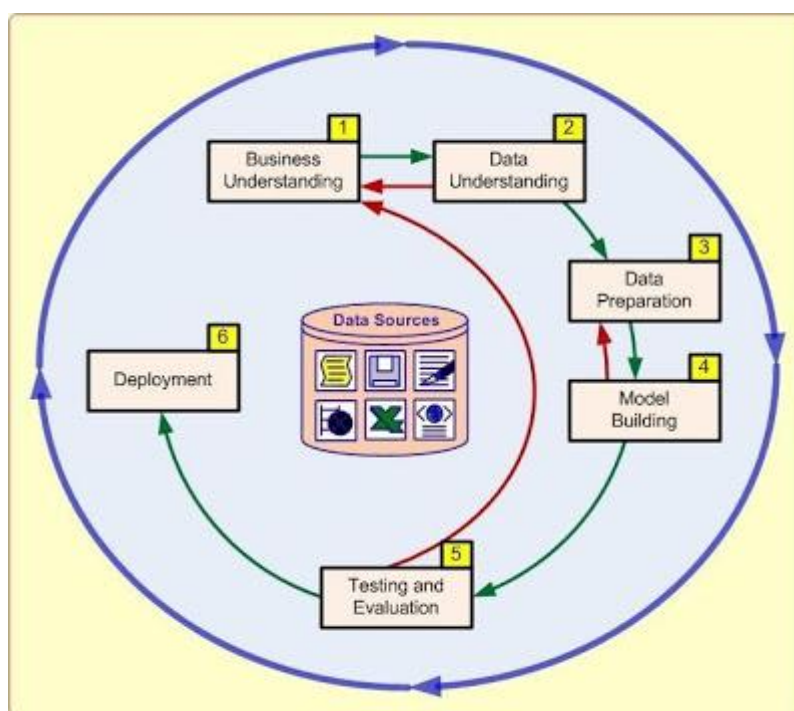


**Gambar II.4. Proses dari Data Mining**

*(Sumber: Statistical Package for the Social Science ; 2004)*

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 II.4. 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 data *preparation* untuk perbaikan lebih lanjut terhadap data atau berpindah maju kepada *fase evaluation*.



Source: Adapted from CRISP-DM.org.

**Gambar II.5. Proses Data Mining Menurut CRISP-DM**

*(Sumber: Cross Industry Standart Process ; 2005)*

Enam fase CRISP-DM ( *Cross Industry Standard Process for Data Mining*) (Larose; 2006) :

1. Fase Pemahaman Bisnis ( *Business Understanding Phase* )
  - a. Penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup 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 data mining.

#### 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](http://www.crisp-dm.org)

### **II.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 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 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 prediksi minat pembeli barang elektronik pada pelanggan berdasarkan umur pelanggan, jenis kelamin, dan kegunaan. Hubungan antara minat pembeli barang elektronik 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 :

- Prediksi harga barang elektronik dalam tiga bulan yang akan datang.
- Prediksi presentase kenaikan pembelian tahun depan jika penjualan barang dinaikan. Beberapa metode dan teknik yang digunakan

dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

- Melakukan pengklusteran terhadap transaksi, dalam jumlah besar.

#### 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 :

- Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
- Memperkirakan apakah suatu pengajuan barang elektronik oleh nasabah merupakan suatu kredit yang baik atau buruk.
- Mendiagnosis barang elektronik 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 :

- Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memilikidana pemasaran yang besar.
- Untuk tujuan audit akutansi, yaitu melakukan pemisahan terhadap prilaku finansial dalam baik dan mencurigakan.
- Melakukan pengklusteran terhadap barang elektronik, dalam jumlah besar.

## 6. Asosiasi

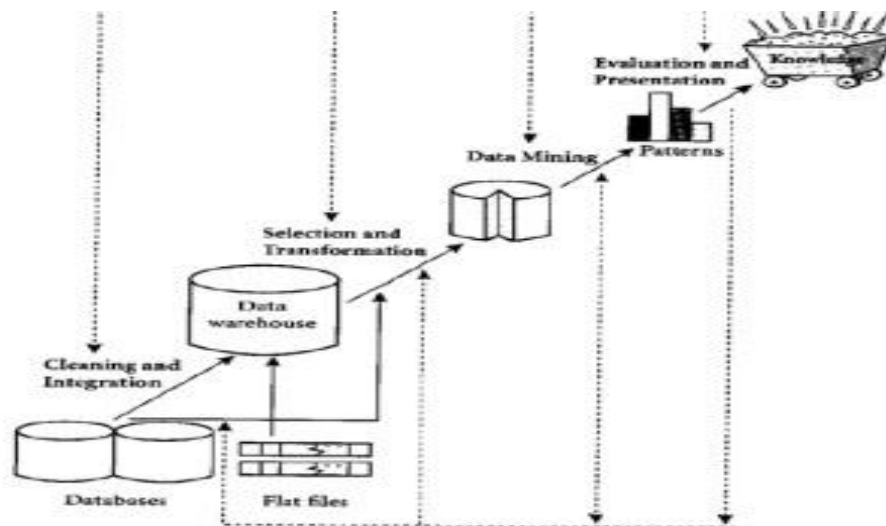
Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Contoh asosiasi dalam bisnis dan penelitian adalah :

- Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran upgrade layanan yang diberikan.

- Menemukan barang yang dibeli secara bersamaan dan barang yang tidak pernah dibeli bersamaan. Untuk mendukung penelitian ini penulis menggunakan Algoritma apriori.

### II.2.3. Konsep Data Mining

Sumber data untuk keperluan pendeteksian minat pembeli barang elektronik dapat berasal dari data pemasaran, pengawasan, survei pasar dan data lapangan. Selanjutnya di analisis untuk mengetahui perilaku data koleksi barang dengan mengungkapkan pola-pola tersembunyi untuk mengubah data menjadi informasi yang berguna, bisa berupa pengelompokan data, maupun ranking data minat pembeli sesuai keperluan user, maka harus ada input kebutuhan yang diinginkan user, dapat dilihat pada gambar berikut:



**Gambar II.6. Konsep Data Mining**  
(Sumber : Eko Budiyanto; 2009:8-9)

### II.3. Algoritma Apriori

Algoritma apriori termasuk jenis aturan asosiasi pada data mining. Selain apriori, yang termasuk pada golongan ini adalah metode Generalized Rule Induction dan Algoritma Hash Based. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut affinity analysis atau market basket analysis. Analisis asosiasi atau association rule adalah teknik data mining untuk menemukan aturan asosiatif antara suatu item. Algoritma Apriori yang bertujuan untuk menemukan frequent itemsets dijalankan pada sekumpulan data. Analisis Apriori didefinisikan suatu proses untuk menemukan semua aturan apriori yang memenuhi syarat minimum untuk support dan syarat minimum confidence. Support adalah nilai penunjang, atau persentase kombinasi sebuah item dalam database. Rumus support sebagai berikut:  $\text{Support (A)} = (\text{jumlah transaksi mengandung A} / \text{Total transaksi}) \times 100\%$ . Sedangkan confidence adalah nilai kepastian yaitu kuatnya hubungan antar item dalam sebuah apriori. Confidence bisa dicari setelah pola frekuensi munculnya sebuah item ditemukan. Rumus untuk menghitung confidence adalah sebagai berikut: Contoh misal ditemukan aturan  $A \Rightarrow B$  maka:  $\text{Confidence } P(B|A) = (\text{Total transaksi mengandung A dan B} / \text{transaksi mengandung A}) \times 100\%$ .

## II.4. Sekilas Tentang SQL Server 2008

### II.4.1. Pengenalan Database, DBMS dan RDBMS

Basis data (atau *database*) adalah kumpulan informasi yang disimpan di dalam komputer secara sistematis sehingga dapat diperiksa menggunakan suatu program komputer untuk memperoleh informasi dari basis data tersebut (Achmad Solihin ; 2010 : 6)

Database digunakan untuk menyimpan informasi atau data yang terintegrasi dengan baik di dalam komputer.

Untuk mengelola database diperlukan suatu perangkat lunak yang disebut DBMS (*Database Management System*). DBMS merupakan suatu sistem perangkat lunak yang memungkinkan user (pengguna) untuk membuat, memelihara, mengontrol, dan mengakses database secara praktis dan efisien. Dengan DBMS, user akan lebih mudah mengontrol dan memanipulasi data yang ada.

Sedangkan RDBMS atau *Relationship Database Management System* merupakan salah satu jenis DBMS yang mendukung adanya *relationship* atau hubungan antar tabel. Di samping RDBMS, terdapat jenis DBMS lain, misalnya *Hierarchy DBMS*, *Object Oriented DBMS*, dan sebagainya.(Achmad Solihin ; 2010 : 6)

1. Beberapa Istilah Database.

- a. *Table*

Sebuah tabel merupakan kumpulan data (nilai) yang diorganisasikan ke dalam baris (*record*) dan kolom (*field*). Masing-masing kolom memiliki nama yang spesifik dan unik.

*b. Field*

*Field* merupakan kolom dari sebuah table. *Field* memiliki ukuran type data tertentu yang menentukan bagaimana data nantinya tersimpan.

*c. Record*

*Record* merupakan sebuah kumpulan nilai yang saling terkait.

*d. Key*

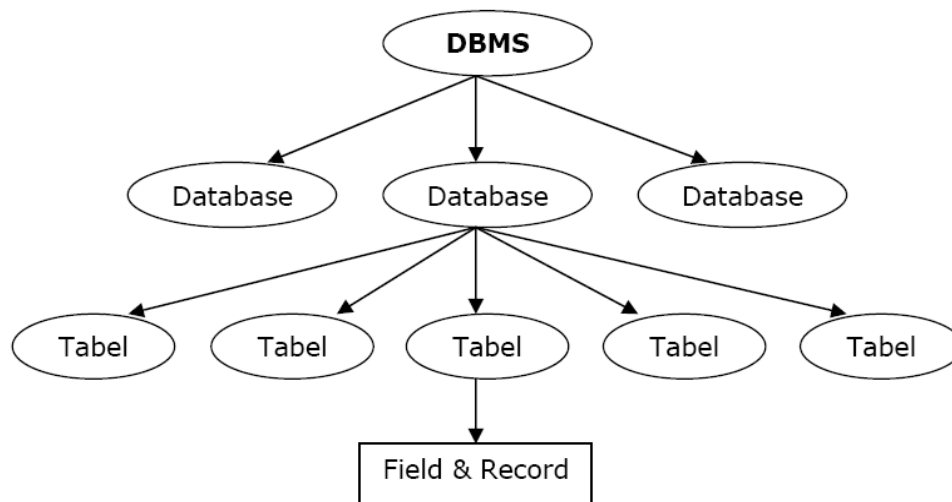
*Key* merupakan suatu field yang dapat dijadikan kunci dalam operasi tabel. Dalam konsep database, *key* memiliki banyak jenis diantaranya *Primary Key*, *Foreign Key*, *Composite Key*, dan lain-lain.

*e. SQL*

*SQL* atau *Structured Query Language* merupakan suatu bahasa (*language*) yang digunakan untuk mengakses database. *SQL* sering disebut juga sebagai *query*. (Achmad Solihin ; 2010 : 7)

## 2. Hierki Database

Dalam konsep *database*, urutan atau hierarki database sangatlah penting. Urutan atau hierarki *database* digambarkan seperti pada gambar II.7. dibawah ini



**Gambar II.7. Hirarki Database**  
 (Sumber : Achmad Solihin ; 2010:7)

#### II.4.2. SQL Server 2008

SQL Server 2008 R2 (dahulu nama kode SQL Server “Kilimanjaro”) diumumkan di TechEd 2009, dan dirilis untuk manufaktur pada 21 April 2010. SQL Server 2008 R2 menambahkan fitur tertentu untuk SQL Server 2008 termasuk data master manajemen sistem dicap sebagai Master Data Services , manajemen pusat data entitas master dan hirarki. Juga *Multi Server Management*, terpusat konsol untuk mengelola beberapa SQL Server 2008 *instance* dan jasa, termasuk database relasional, *Reporting Services*, *Analysis Services & Integration Services*. Adapun tampilan utama SQL Server 2008 R2 :



**Gambar II.8. SQL Server 2008 R2**  
(Sumber : Achmad Solihin ; 2010:7)

SQL Server 2008 R2 meliputi sejumlah layanan baru, termasuk *PowerPivot* untuk *Excel* dan *SharePoint*, *Master Data Services*, *StreamInsight*, *Report Builder 3.0*, *Reporting Services Add-in* untuk *SharePoint*, sebuah lapis fungsi Data dalam Visual Studio yang memungkinkan kemas database berjenjang sebagai bagian dari aplikasi, dan *Utility SQL Server* yang bernama UC (*Utility Control Point*), bagian dari AMSM (Aplikasi dan *Multi-Server Management*) yang digunakan untuk mengelola beberapa SQL Server. *Microsoft SQL Server* tersedia dalam beberapa edisi, dengan set fitur yang berbeda dan menargetkan pengguna yang berbeda. Edisi ini adalah:

### 1. SQL Server Compact Edition (SQL CE)

Edisi kompak adalah mesin database tertanam. Berbeda dengan edisi lain dari SQL Server, SQL CE mesin didasarkan pada SQL Mobile (awalnya dirancang untuk digunakan dengan perangkat genggam) dan tidak berbagi binari yang sama. Karena ukurannya yang kecil (1 MB DLL jejak), ia memiliki fitur mencolok mengurangi diatur dibandingkan dengan edisi lainnya. Misalnya, mendukung subset dari tipe data standar, tidak mendukung

prosedur yang tersimpan atau Tampilan atau batch multi-pernyataan (antara keterbatasan lain). Hal ini terbatas pada ukuran basis data maksimal 4 GB dan tidak dapat dijalankan sebagai layanan *Windows, Compact Edition* harus di-host oleh aplikasi menggunakannya. Versi 3.5 mencakup pekerjaan yang cukup yang mendukung ADO.NET Layanan Sinkronisasi.

## **2. SQL Server Developer Edition**

*SQL Server Developer Edition* mencakup fitur yang sama seperti *SQL Server Enterprise Edition*, tetapi dibatasi oleh lisensi yang akan hanya digunakan sebagai pengembangan dan sistem pengujian, dan bukan sebagai server produksi. Edisi ini tersedia untuk *download* oleh siswa secara gratis sebagai bagian dari *Microsoft 's DreamSpark program*.

## **3. Embedded Edition SQL Server 2005 (SSEE)**

*SQL Server 2005 Edition Embedded* adalah contoh bernama dikonfigurasi secara khusus dari mesin database *SQL Server Express* yang hanya dapat diakses oleh Jasa Windows tertentu.

## **4. SQL Server Enterprise Edition**

*SQL Server Enterprise Edition* adalah edisi fitur lengkap dari SQL Server, termasuk mesin database inti dan add-on jasa, sedangkan termasuk berbagai alat untuk membuat dan mengelola sebuah *cluster SQL Server*.

## **5. SQL Server Edition Evaluasi**

*SQL Server Edition Evaluasi*, juga dikenal sebagai *Trial Edition*, memiliki semua fitur dari Edisi *Enterprise*, tetapi terbatas pada 180 hari, setelah itu alat akan terus berjalan, tetapi layanan server akan berhenti.

## **6. SQL Server Express Edition**

*SQL Server Express Edition* adalah *free edition* skala bawah, *SQL Server*, yang meliputi database engine inti. Meskipun tidak ada batasan pada jumlah database atau pengguna yang didukung, hal ini terbatas untuk menggunakan satu prosesor, memori 1 GB dan 4 file database GB (10 GB file database dari *SQL Server 2008 Express R2*). Seluruh database disimpan dalam satu .mdf file, dan dengan demikian sehingga cocok untuk deployment XCOPY. Hal ini dimaksudkan sebagai pengganti MSDE. Dua edisi tambahan menyediakan superset dari fitur tersebut tidak di Edisi Express asli. Yang pertama adalah *SQL Server Express dengan Tools*, yang meliputi *SQL Server Management Studio Basic Server*. *SQL Express* dengan Advanced Layanan menambahkan teks lengkap kemampuan pencarian dan layanan pelaporan.

## **7. SQL Server Fast Track**

*SQL Server Fast Track* ini khusus untuk skala data perusahaan pergudangan penyimpanan dan pengolahan intelijen bisnis, dan berjalan pada-arsitektur perangkat keras referensi yang dioptimalkan untuk *Fast Track*.

## **II.5. UML (*Unified Modelling Language*)**

UML singkatan dari *Unified Modeling Language* yang berarti bahasa pemodelan standar. (Chonoles ; 2003: bab 1) mengatakan sebagai bahasa, berarti UML memiliki sintaks dan semantik. Ketika kita membuat model menggunakan konsep UML ada aturan-aturan yang harus diikuti. Bagaimana elemen pada model – model yang kita buat berhubungan satu dengan yang lainnya harus mengikuti standar yang ada. UML bukan hanya sekedar diagram, tetapi juga menceritakan konteksnya. Ketika pelanggan memesan sesuatu dari sistem, bagaimana transaksinya ? Bagaimana sistem mengatasi error yang terjadi ? Bagaimana keamanan terhadap sistem yang kita buat ? Dan sebagainya dapat dijawab dengan UML. (Prabowo Pudjo Widodo & Herlawati ; 2011: 6)

UML dapat diaplikasikan untuk maksud tertentu, biasanya antara lain untuk :

1. Merancang perangkat lunak
2. Sarana komunikasi antara perangkat lunak dengan proses bisnis.
3. Menjabarkan sistem secara rinci untuk analisa dan mencari apa yang diperlukan sistem.
4. Mendokumentasi sistem yang ada, proses-proses dan organisasinya.

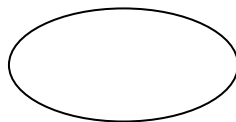
### **II.5.1. Diagram – Diagram UML**

Beberapa literatur menyebutkan bahwa UML menyediakan sembilan jenis diagram, yang lain menyebutkan delapan karena ada beberapa diagram yang digabung misalnya diagram komunikasi, diagram urutan dan diagram pewaktuan

digabung menjadi diagram interaksi. Namun demikian model-model itu dapat dikelompokkan berdasarkan sifatnya yaitu statis atau dinamis.(Prabowo Pudjo Widodo & Herlawati ; 2011: 10)

### *1. Use Case Diagram*

Menurut ( Pilone, 2005 ; bab 7.1 ) use case menggambarkan fungsi tertentu dalam suatu sistem berupa komponen, kejadian atau kelas. Sedangkan ( Whitten, 2004 ; 258 ) mengartikan use case sebagai urutan langkah-langkah yang secara tindakan saling terkait ( skenario ), baik terotomatisasi maupun secara manual, untuk tujuan melengkapi satu tugas bisnis tunggal. Use case digambarkan dalam bentuk ellips/oval. (Prabowo Pudjo Widodo & Herlawati ; 2011: 21)



**Gambar II.9. Simbol Use Case**  
(Sumber : Prabowo Pudjo Widodo & Herlawati ; 2011: 22)

Elemen use case terdiri dari :

- a. Diagram use case, disertai dengan narasi dan skenario.
- b. Aktor ( *actor* ), mendefinisikan entitas diluar sistem yang memakai sistem.
- c. Asosiasi ( *assosiations* ), mengindikasikan aktor mana yang berinteraksi dengan use case dalam suatu sistem.
- d. << include >> dan << extend >>. Merupakan indikator yang menggambarkan jenis relasi dan interaksi antar use-case.

- e. Generalisasi ( *generalization* ), menggambarkan hubungan turunan antar use case atau antar aktor. (Prabowo Pudjo Widodo & Herlawati ; 2011: 35)

## 2. Diagram Kelas ( *Class Diagram* )

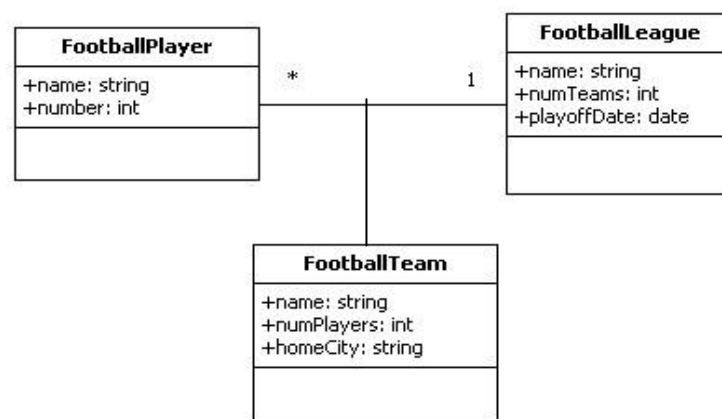
Diagram kelas adalah inti dari proses pemodelan objek. Baik forward engineering maupun reverse engineering memanfaatkan diagram ini. *Forward engineering* adalah proses perubahan model menjadi kode program sedangkan reverse engineering sebaliknya merubah kode program menjadi model. (Prabowo Pudjo Widodo & Herlawati ; 2011: 37).

Diagram kelas merupakan kumpulan kelas-kelas objek. Oleh karena itu pengertian kelas sangat penting sebelum merancang diagram kelas. (Whitten; 2004: 410 ) mengartikan kelas sebagai satu set objek yang memiliki atribut kelas objek ( objek class ). Secara alami, objek yang berupa buku analisis desain dan buku pemrograman terstruktur kita kelompokkan dalam satu kelas, yaitu kelas buku. Kedua objek memiliki atribut dan perilaku yang serupa. Contohnya, kedua objek mungkin memiliki atribut yang serupa seperti nomor ISBN, Judul, tanggal penerbitan, edisi, dan sebagainya. Demikian juga, kedua objek memiliki perilaku yang sama misalnya membuka dan menutup. (Prabowo Pudjo Widodo & Herlawati ; 2011: 39).

Kelas menggambarkan suatu group yang memiliki kesamaan dan perilaku. Kelas merupakan cetak biru suatu objek dalam suatu sistem orientasi objek. Dapat dikatakan kelas adalah sejenis alat pengklasifikasi. Sebagai contoh volkswag, toyota dan ford merupakan kumpulan mobil

sehingga kita dapat mengelompokkannya dalam menyatakan konsep yang dapat dilihat maupun abstrak.

Kelas dinyatakan dalam kotak yang terbagi menjadi beberapa kompartemen. Kompartemen adalah area dalam kelas yang berisi informasi. Kompartemen pertama berisikan nama kelas, berikutnya atribut dan yang terakhir operasinya. Kompartemen-kompartemen itu bisa disembunyikan untuk memperjelas diagram yang kita buat. (Prabowo Pudjo Widodo & Herlawati ; 2011: 40)



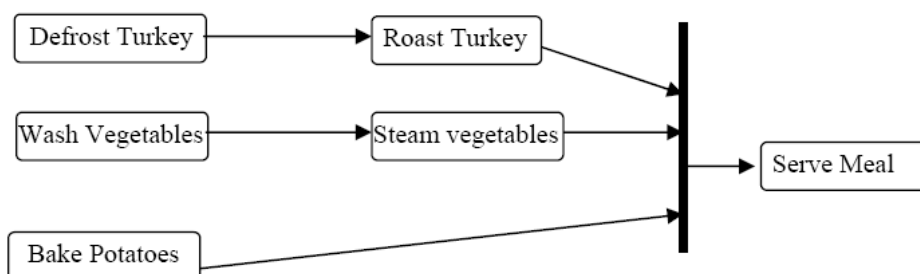
**Gambar II.10. Contoh Class Diagram**  
(Sumber : Prabowo Pudjo Widodo & Herlawati ; 2011: 69)

### 3. Activity diagram

Diagram aktifitas lebih memfokuskan diri pada eksekusi dan alur sistem dari pada bagaimana sistem itu dirakit. Diagram ini tidak hanya memodelkan software melainkan memodelkan model bisnis juga. Diagram aktifitas menunjukkan aktifitas sistem dalam bentuk aksi-aksi. Ketika digunakan dalam pemodelan software, diagram aktifitas merepresentasikan

pemanggilan suatu fungsi tertentu misalnya call. Sedangkan bila dimodelkan dalam pemodelan bisnis, diagram ini menggambarkan aktifitas yang dipicu oleh kejadian-kejadian di luar seperti pemesanan atau kejadian-kejadian internal. (Prabowo Pudjo Widodo & Herlawati ; 2011: 143)

### Prepare Meal



**Gambar II.11. Contoh Activity Diagram Sederhana**  
(Sumber : Prabowo Pudjo Widodo & Herlawati ; 2011: 163)

#### 4. Sequence Diagram

*Sequence Diagram* digunakan untuk menggambarkan perilaku pada sebuah skenario. Diagram ini menunjukkan sebuah contoh objek dan pesan yang diletakkan diantara objek-objek ini didalam *use case*. Diagram *sequence* dimaksudkan untuk mengembangkan komunikasi antara objek, bukan memanipulasi data saat berkomunikasi. (Prabowo Pudjo Widodo & Herlawati ; 2011: 174)

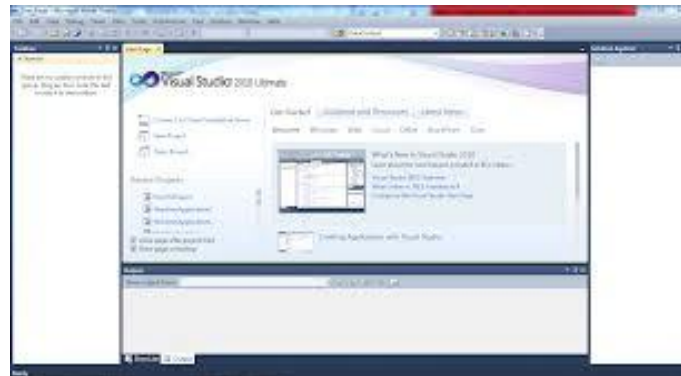
## **II.6. Microsoft Visual Studio 2010**

*Microsoft Visual Studio 2010* merupakan sebuah IDE (*Integrated Development Environment*) yang dikembangkan oleh Microsoft. IDE ini mencakup semua bahasa pemrograman berbasis .NET framework yang dikembangkan oleh Microsoft. Keunggulan *Microsoft Visual Studio 2010* ini antara lain adalah support *Windows 7* editor baru dengan WPF (*Windows Presentation Foundation*), dan banyak peningkatan mencakup semua bahasa dan fitur lainnya. (Patricle Pelland, Pascal Pare, dan Ken Haines ; 2011)

*Visual Studio .NET* adalah sebuah tools pengembangan perangkat lunak untuk membangun aplikasi ASP Web, layanan XML Web, aplikasi desktop, dan aplikasi mobile. Visual Basic .NET, Visual C++ .NET, Visual C# .NET, dan Visual J# .NET; semuanya menggunakan Integrated Development Environment (IDE) atau lingkungan pengembangan terintegrasi yang sama; yang membolehkan mereka untuk saling berbagi tools dan fasilitas dalam pembuatan solusi yang memadukan beberapa bahasa (*mixed-language solutions*). Selain itu, bahasa-bahasa ini mempengaruhi fungsionalitas dari NET Framework, dan menyediakan pengaksesan ke kunci teknologi yang menyederhanakan proses pengembangan dari aplikasi ASP Web dan layanan XML Web (Atmusudirjo, SP ; 2001; 42).

Gambar berikut menampilkan tampilan utama Microsoft Visual Studio

2010 :



**Gambar II.12. Tampilan Microsoft Visual Studio 2010**