

BAB II

LANDASAN TEORI

2.1 Klasifikasi dan Prediksi

Menurut Han (2007) klasifikasi dan prediksi adalah dua bentuk dari analisis data yang dapat digunakan untuk mengekstrak model yang mendeskripsikan kelas data yang penting atau untuk memprediksi tren data di masa depan. Analisis tersebut dapat membantu untuk menyediakan pemahaman yang lebih baik dari data yang besar. Jika klasifikasi memprediksi label yang bersifat kategori (diskrit dan tidak berurutan), prediksi memodelkan fungsi nilai yang kontinyu.

2.1.1 Klasifikasi

Klasifikasi adalah teknik dimana sebuah model atau *classifier* dikonstruksi untuk memprediksi label yang bersifat kategori, seperti “aman” atau “beresiko” untuk data aplikasi pinjaman, “ya” atau “tidak” untuk data marketing, atau “perawatan A”, “perawatan B”, atau perawatan “C” untuk data medis. Kategori ini dapat direpresentasikan oleh nilai diskrit, dimana susunan dari nilai tidak mempunyai arti. Sebagai contoh, nilai 1, 2, dan 3 mungkin digunakan untuk merepresentasikan perawatan A, B, dan C dimana tidak ada urutan yang dinyatakan dari grup perawatan tersebut.

2.1.2 Prediksi

Prediksi numerik adalah langkah yang dilakukan dalam memprediksi nilai yang kontinyu (atau berurutan) untuk nilai yang diberikan. Menurut Han (2007) sejauh ini pendekatan yang paling banyak digunakan untuk prediksi numerik (yang kemudian disebut dengan prediksi saja) adalah regresi, sebuah metodologi statistik yang dikembangkan oleh Sir Francis Galton, seorang ahli matematika yang juga merupakan sepupu dari Charles Darwin. Pada kenyataannya, banyak tulisan yang menggunakan istilah “regresi” dan “prediksi numerik” sebagai sinonim. Namun, seperti yang sudah diketahui bahwa beberapa teknik klasifikasi (seperti *backpropagation*, *support vector machines*, dan *k-nearest-neighbour classifier*) dapat diadaptasi untuk prediksi. Maka untuk lebih jelasnya, dalam bahasan kali ini digunakan teknik regresi sebagai sarana prediksi.

Analisis regresi dapat digunakan untuk memodelkan hubungan antara satu atau lebih variabel *independen* atau prediktor dan sebuah variabel *dependen* atau respon (yang bernilai kontinyu). Dalam konteks *data mining*, variabel prediktor adalah atribut yang berkepentingan mendeskripsikan baris data (yakni yang menyusun vektor atribut). Secara umum, nilai dari variabel prediktor adalah diketahui. Variabel respon adalah apa yang akan diprediksikan. Ketika diberikan sebuah baris data yang dideskripsikan oleh variabel prediktor, maka akan diprediksikan nilai yang bersesuaian dengan variabel respon.

Analisis regresi adalah pilihan yang bagus manakala semua nilai dari variabel prediktor bernilai kontinyu juga. Banyak permasalahan dapat dipecahkan oleh Regresi Linier, dan akan lebih banyak lagi dapat dikerjakan dengan

mengaplikasikan transformasi kepada variabel-variabel yang bersesuaian sehingga masalah non-linier dapat diubah menjadi masalah linier.

2.2 Statistika

Statistika adalah ilmu yang mempelajari bagaimana merencanakan, mengumpulkan, menganalisis, menginterpretasi, dan mempresentasikan data. Singkatnya, statistika adalah ilmu yang berkaitan dengan data. Statistika berbeda dengan statistik, dimana statistika adalah ilmu yang berkaitan dengan data, sedangkan statistik adalah data, informasi atau hasil penerapan algoritma statistika pada suatu data.

Statistika dibedakan menjadi 2 yaitu statistika deskriptif dan inferensial. Statistika deskriptif lebih berpusat pada penjelasan mengenai karakteristik data dan biasanya disajikan dalam ukuran-ukuran statistika seperti mean, median, simpangan baku, dan lainnya. Selain itu juga bisa disajikan melalui tabel, diagram pie, histogram, kurva, dan lainnya. Melalui statistika deskriptif, data lebih mudah dibaca dan dimengerti.

Sedangkan statistika inferensial terdiri dari 2 bagian umum yaitu pengujian hipotesis dan penaksiran parameter. Setelah data dikumpulkan maka dilakukan analisis data menggunakan berbagai macam metode statistika yang sesuai lalu menginterpretasikan hasil analisis tersebut. Statistika inferensial mengolah data dari *sample* dan melakukan pengujian atau melakukan pendekatan terhadap nilai atau karakter dari suatu populasi. Sebagai contoh pengujian korelasi, regresi, dan anova.

2.3 Regresi

Pengertian regresi secara umum adalah studi ketergantungan satu variabel tergantung (variabel dijelaskan atau variabel tak bebas) pada satu atau lebih variabel lain (variabel penjelas atau variabel bebas), dengan tujuan untuk menaksir atau memprediksi nilai rata-rata hitung atau rata-rata variabel tak bebas, berdasarkan nilai yang diketahui atau tetap (dalam pengambilan *sample* berulang-ulang) dari variabel penjelas. Analisis regresi digunakan untuk memodelkan hubungan antara satu atau lebih variabel *independen* atau prediktor dan sebuah variabel dependen atau respon (yang bernilai kontinyu).

2.3.1 Regresi Linear

Menurut Lungan (2006), regresi linear sederhana merupakan bagian regresi yang mencakup hubungan linear satu peubah tak bebas Y dengan satu peubah bebas X. Hubungan linear X dan Y dari satu populasi disebut garis regresi populasi yang dinyatakan persamaan sebagai berikut :

$$\mu_{Y,X} = E(Y/X) = \beta_0 + \beta_1 X$$

$\mu_{Y,X}$: rata-rata Y untuk nilai X tertentu

β_0 : jarak titik pangkal dengan titik potong garis regresi dengan sumbu Y (*intercept*)

β_1 : Kemiringan (*slope* atau *gradient*) garis regresi

Kalau ingin menduga rataan μ_{Y,X_i} maka nilai Y perlu ditentukan untuk suatu nilai X_i tertentu. Nilai Y tersebut untuk X_i tertentu dinyatakan dengan Y_i . Nilai Y_i dan μ_{Y,X_i} pada umumnya tidak sama. Perbedaan tersebut tergantung pada

ketepatan model untuk menggambarkan keadaan yang sebenarnya dan ketepatan pengukuran peubah Y dan X.

Perbedaan antara Y_i dan μ_{Y,X_i} disebut galat acak (*random error*) dan dinyatakan dengan simbol ε_i . Dengan demikian :

$$\varepsilon_i = Y_i - \mu_{Y,X_i} \text{ atau } Y_i = \mu_{Y,X_i} + \varepsilon_i$$

Dari persamaan ini diperoleh model regresi linier sederhana dari suatu populasi sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

2.3.2 Regresi Linear Berganda

Menurut Lungan (2006), regresi linear berganda merupakan perluasan dari regresi linear sederhana. Jika regresi linear sederhana mempersoalkan tentang hubungan peubah tak bebas atau peubah kriteria (*respons*) dengan suatu peubah bebas (*deterministik*), maka pada regresi linear berganda mempersoalkan hubungan linear antara satu peubah tak bebas dengan beberapa peubah bebas.

Peubah tak bebas dapat berupa ukuran atau kriteria keberhasilan, sedangkan peubah bebas dapat berupa faktor-faktor penentu keberhasilan tersebut.

Jika banyaknya peubah bebas adalah k , maka model regresi populasi dapat dinyatakan dengan:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i$$

$\beta_0, \beta_1, \beta_2, \dots, \beta_k$ merupakan parameter yang disebut koefisien regresi parsial. ε_i = galat.

Jika ε_i diasumsikan = 0, maka diperoleh persamaan regresi linier berganda dari suatu populasi sebagai berikut :

$$\mu Y \cdot X_1, X_2, \dots, X_p = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

2.4 Pendidikan

Pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak yang mulia, serta keterampilan yang diperlukan dirinya dan masyarakat. Pendidikan biasanya berawal saat seorang bayi dilahirkan dan berlangsung seumur hidup. Pendidikan bisa saja berawal dari sebelum bayi terlahir seperti yang dilakukan oleh banyak orang dengan memainkan musik dan membaca kepada bayi dalam kandungan.

Menurut Horton dan Hunt, Lembaga pendidikan berkaitan dengan fungsi yang nyata sebagai berikut :

1. Mempersiapkan anggota masyarakat untuk mencari nafkah.
2. Mengembangkan bakat perseorangan demi kepuasan pribadi dan bagi kepentingan masyarakat.
3. Melestarikan kebudayaan.
4. Menanamkan keterampilan yang diperlukan bagi partisipasi dalam demokrasi.

2.5 Sekolah Menengah Atas (SMA)

Sekolah menengah atas adalah jenjang pendidikan menengah pada pendidikan formal di Indonesia setelah lulus dari sekolah menengah pertama atau sederajat. Sekolah menengah atas ditempuh dalam waktu 3 tahun. Pada tahun kedua, siswa SMA dapat memilih salah satu dari 3 jurusan yang ada yaitu IPA, IPS, dan Bahasa.

SMA diselenggarakan oleh pemerintah maupun swasta. Sejak diberlakukannya otonomi daerah pada tahun 2001, pengelolaan SMA negeri di Indonesia yang sebelumnya berada dibawah Departemen Pendidikan Nasional, kini menjadi tanggung jawab pemerintah daerah kabupaten/ kota. Sedangkan Departemen Pendidikan Nasional hanya berperan sebagai regulator dalam bidang standar nasional pendidikan. Secara struktural, SMA negeri merupakan unit pelaksana teknis dinas pendidikan kabupaten/ kota.

2.6 Perguruan Tinggi

Di Indonesia, Perguruan Tinggi dapat berbentuk akademi, institut, politeknik, sekolah tinggi, dan universitas. Perguruan tinggi dapat menyelenggarakan pendidikan akademik, profesi, dan vokasi dengan program diploma (D1, D2, D3, dan D4), sarjana (S1), magister (S2), doktor (S3), dan spesialis.

Pengelolaan dan regulasi perguruan tinggi di Indonesia dilakukan oleh Kementerian Pendidikan Nasional. Rektor Perguruan Tinggi merupakan pejabat eselon di bawah Menteri Pendidikan Nasional. Selain itu terdapat juga Perguruan Tinggi yang dikelola oleh kementerian atau lembaga pemerintah non kementerian

yang pada umumnya merupakan Perguruan Tinggi Kedinasan. Perguruan Tinggi Swasta di Indonesia, dikelola oleh masyarakat sesuai dengan perundang-undangan yang berlaku. Bimbingan dan pengawasan atas penyelenggaraan Perguruan Tinggi Swasta pada mulanya dilakukan oleh Lembaga Perguruan Tinggi Swasta (LPTS) yang dibentuk oleh pemerintah. LPTS ini kemudian menjadi Koordinasi Perguruan Tinggi Swasta yang disebut dengan KOPERTIS.

2.7 Konsep Dasar Sistem

Menurut Neuschel (1976), terdapat dua kelompok pendekatan di dalam mendefinisikan sistem, yaitu yang menekankan pada prosedurnya dan yang menekankan pada komponen atau elemennya. Pendekatan sistem yang lebih menekankan pada prosedur sistem adalah sebagai berikut:

“Sistem adalah suatu jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan suatu kegiatan atau untuk menyelesaikan suatu sasaran tertentu.”

Pendekatan sistem yang merupakan jaringan kerja dari prosedur lebih menekankan urutan-urutan operasi di dalam sistem. Prosedur didefinisikan oleh Neuschel (1976) sebagai berikut:

“Prosedur adalah suatu urutan-urutan operasi klerikal (tulis-menulis), biasanya melibatkan beberapa orang di dalam satu atau lebih departemen, yang diterapkan untuk menjamin penanganan yang seragam dari transaksi-transaksi bisnis yang terjadi.”

Pendekatan sistem yang lebih menekankan pada elemen atau komponennya dalam mendefinisikan sistem, menurut Neuschel adalah “Sistem

adalah kumpulan dari elemen-elemen yang berinteraksi untuk mencapai suatu tujuan tertentu.”

2.8 Sistem Informasi

Sistem adalah suatu jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan suatu kegiatan atau untuk menyelesaikan suatu sasaran tertentu (Hartono,1999:1). Informasi adalah data yang diolah menjadi bentuk yang lebih berguna dan lebih berarti bagi yang menerimanya. Data merupakan bentuk yang masih mentah yang belum dapat bercerita banyak, sehingga perlu diolah lanjut (Hartono, 1999:8).

Informasi dapat menggambarkan kejadian nyata yang digunakan untuk pengambilan keputusan. Sumber dari informasi adalah data yang berbentuk huruf, simbol, alfabet, dan sebagainya. Sistem informasi mempunyai elemen utama, yaitu data yang menyediakan informasi, prosedur yang memberitahu pengguna bagaimana mengoperasikan sistem informasi, menyelesaikan masalah, membuat keputusan, dan menggunakan sistem informasi tersebut. Orang-orang dalam sistem informasi membuat prosedur untuk mengolah dan memanipulasi data sehingga menghasilkan informasi dan menyebarkan informasi tersebut ke lingkungannya.

Model dasar sistem adalah masukan, pengolahan, dan pengeluaran. Fungsi pengolahan informasi sering membutuhkan data yang telah dikumpulkan dan diolah dalam waktu periode sebelumnya. Oleh karena itu, dalam model sistem informasi ditambahkan pula media penyimpanan data. Maka fungsi pengolahan

informasi bukan lagi mengubah data menjadi informasi, tetapi juga menyimpan data untuk penggunaan lanjutan.

Sistem Informasi adalah data yang diolah menjadi bentuk yang lebih berguna dan lebih berarti bagi yang menerimanya. Data merupakan bentuk yang masih mentah yang belum dapat bercerita banyak, sehingga perlu untuk diolah lebih lanjut. Karena pada saat ini, para pembuat keputusan memahami bahwa informasi tidak hanya sekedar produk sampingan bisnis yang sedang dijalankan, namun juga sebagai bahan pengisi bisnis dan menjadi faktor kritis dalam menentukan kesuksesan atau kegagalan suatu usaha.

Untuk menghasilkan informasi yang berkualitas maka dibuatlah sistem informasi (Hartono, 1999 : 11). Sistem informasi adalah suatu sistem di dalam suatu organisasi yang mempertemukan kebutuhan pengolahan transaksi harian, mendukung operasi, bersifat manajerial, dan kegiatan strategi dari suatu organisasi serta menyediakan pihak luar tertentu dengan laporan-laporan yang diperlukan.

Informasi merupakan data yang telah diolah menjadi suatu bentuk yang berarti bagi penerimanya dan bermanfaat dalam pengambilan keputusan saat ini atau di masa yang akan datang. Sumber dari informasi adalah data, merupakan bentuk yang masih mentah dan belum dapat bercerita banyak, sehingga membutuhkan pengolahan lebih lanjut. Kualitas dari sistem informasi bergantung pada dua hal, yaitu:

1. Informasi harus akurat, dimana informasi tersebut harus bebas dari kesalahan.
2. Informasi tersebut harus relevan, supaya informasi tersebut dapat memberikan masukan bagi penerimanya.

Ada beragam definisi sistem informasi, sebagaimana tercantum pada Tabel 2.1 Dari berbagai definisi tersebut, dapat disimpulkan bahwa sistem informasi mencakup sejumlah komponen (manusia, komputer, teknologi informasi, dan prosedur kerja), ada sesuatu yang diproses (data menjadi informasi), dan dimaksudkan untuk mencapai suatu sasaran atau tujuan.

Tabel 2.1 Definisi Sistem Informasi

Sumber	Definisi
Alter (1992)	Sistem informasi adalah kombinasi antara prosedur kerja, informasi, orang dan teknologi informasi yang diorganisasikan untuk mencapai tujuan dalam sebuah organisasi.
Hall (2001)	Sistem informasi adalah sebuah rangkaian prosedur formal dimana data dikelompokkan, diproses menjadi informasi, dan didistribusikan kepada pemakai.
Tuban, McLean, dan Wetherbe (1999)	Sebuah sistem informasi mengumpulkan, memproses, menyimpan, menganalisis, dan menyebarkan informasi untuk tujuan yang spesifik.

Istilah sistem informasi juga sering dikacaukan dengan sistem informasi manajemen (SIM). Kedua hal ini sebenarnya tidak sama. Sistem informasi manajemen merupakan salah satu jenis sistem informasi, yang secara khusus ditujukan untuk menghasilkan informasi bagi pihak manajemen dan untuk pengambil keputusan.

2.9 Analisis dan Perancangan Sistem

Menurut Kendall (2003), analisis sistem adalah penguraian dari suatu sistem informasi yang utuh kedalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan-permasalahan, kesempatan-kesempatan, hambatan-hambatan yang terjadi, dan kebutuhan-kebutuhan yang diharapkan sehingga dapat diusulkan perbaikan-perbaikannya.

Tahap analisis sistem terdapat langkah-langkah dasar yang harus dilakukan oleh analis sistem sebagai berikut:

1. *Identify*, yaitu mengidentifikasi masalah.
2. *Understand*, yaitu memahami kerja dari sistem yang ada.
3. *Analyze*, yaitu menganalisis sistem.
4. *Report*, yaitu membuat laporan hasil analisis.

Setelah tahap analisis sistem selesai dilakukan, maka analis sistem telah mendapatkan gambaran dengan jelas apa yang harus dikerjakan. Tiba waktunya sekarang bagi analis sistem untuk memikirkan bagaimana membentuk sistem tersebut. Tahap ini disebut dengan desain sistem.

Analisis dan perancangan sistem dipergunakan untuk menganalisis, merancang, dan mengimplementasikan peningkatan-peningkatan fungsi bisnis yang dapat dicapai melalui penggunaan sistem informasi terkomputerisasi.

2.10 Sistem Flow

System flow adalah bagan yang menunjukkan arus pekerjaan secara menyeluruh dari suatu sistem dimana bagan ini menjelaskan urutan prosedur-prosedur yang ada dalam sistem dan biasanya dalam membuat *system flow* sebaiknya ditentukan pula fungsi-fungsi yang melaksanakan atau bertanggung jawab terhadap sub-sistem yang ada (Hartono, 1999).

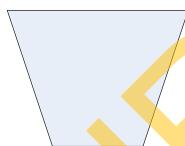
Terdapat berbagai macam bentuk simbol yang digunakan untuk merancang sebuah desain dari sistem, diantaranya adalah *terminator*, *manual operation*, *document*, *process*, *database*, *manual input*, *decision*, *off-line storage*, *on-page reference*, dan *off-page reference*.

Terminator merupakan bentuk simbol yang digunakan sebagai tanda dimulainya jalan proses sistem ataupun tanda akhir dari sebuah penggerjaan suatu sistem. Simbol dari *terminator* dapat dilihat pada Gambar 2.1.



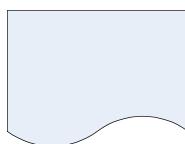
Gambar 2.1 Terminator

Manual operation digunakan untuk menggambarkan sebuah proses kerja yang dilakukan tanpa menggunakan komputer sebagai medianya (menggunakan proses manual). Simbol dari *operation* dapat dilihat pada Gambar 2.2.



Gambar 2.2 Manual Operation

Document merupakan simbol dari dokumen yang berupa kertas laporan, surat-surat, memo, maupun arsip-arsip secara fisik. Simbol dari *document* dapat dilihat pada Gambar 2.3.



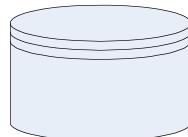
Gambar 2.3 Document

Process adalah sebuah bentuk kerja sistem yang dilakukan secara terkomputerisasi. Simbol dari *process* dapat dilihat pada Gambar 2.4.



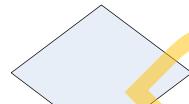
Gambar 2.4 Process

Database digunakan sebagai media penyimpanan data yang bersifat terkomputerisasi. Simbol dari *database* dapat dilihat pada Gambar 2.5.



Gambar 2.5 Database

Decision merupakan operator logika yang digunakan sebagai penentu keputusan dari suatu permintaan atau proses dengan dua nilai, benar dan salah. Simbol dari *decision* dapat dilihat pada Gambar 2.6.



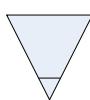
Gambar 2.6 Decision

Manual input digunakan untuk melakukan proses *input* ke dalam *database* melalui *keyboard*. Simbol dari *manual input* dapat dilihat pada Gambar 2.7.



Gambar 2.7 Manual Input

Off-line storage merupakan bentuk media penyimpanan yang berbeda dengan *database*, dimana media penyimpanan ini menyimpan dokumen secara manual atau lebih dikenal dengan nama arsip. Simbol dari *off-line storage* dapat dilihat pada Gambar 2.8.



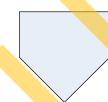
Gambar 2.8 Off-Line Storage

On-page reference digunakan sebagai simbol untuk menghubungkan bagan desain sebuah sistem apabila hubungan arus data yang ada terlalu jauh dalam permasalahan letaknya. Simbol dari *On-page reference* dapat dilihat pada Gambar 2.9.



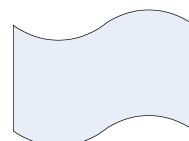
Gambar 2.9 *On-Page Reference*

Off-page reference memiliki sifat yang sedikit berbeda dengan *On-page reference*, karena simbol ini hanya digunakan apabila arus data yang ada dilanjutkan ke halaman yang berbeda. Simbol dari *off-page reference* dapat dilihat pada Gambar 2.10.



Gambar 2.10 *Off-Page Reference*

Paper tape merupakan sebuah simbol yang umumnya menggantikan bentuk penggambaran jenis pembayaran yang digunakan (misal: uang) dalam transaksi yang ada pada sistem yang dirancang. Simbol dari *paper tape* dapat dilihat pada Gambar 2.11.



Gambar 2.11 *Paper Tape*

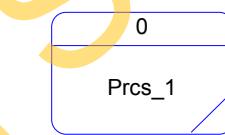
2.11 Data Flow Diagram (DFD)

Menurut Andri Kristanto (2008), Data Flow Diagram (DFD) adalah suatu model logika data atau proses yang dibuat untuk menggambarkan darimana asal

data dan kemana tujuan data yang keluar dari sistem, dimana data tersebut disimpan, proses apa yang menghasilkan data tersebut dan interaksi antara data yang tersimpan, dan proses yang dikenakan pada data tersebut.

Data Flow Diagram merupakan suatu metode pengembangan sistem yang terstruktur (*structure analysis and design*). Penggunaan notasi dalam data flow diagram sangat membantu untuk memahami suatu sistem pada semua tingkat kompleksitas. Pada tahap analisis, penggunaan notasi ini dapat membantu dalam berkomunikasi dengan pemakai sistem untuk memahami sistem secara logika.

Di dalam data flow diagram, terdapat empat simbol yang digunakan yaitu *process*, *external entity*, *data store*, dan *data flow*. Simbol *process* digunakan untuk melakukan suatu perubahan berdasarkan data yang diinputkan dan menghasilkan data dari perubahan tersebut. Simbol *process* dapat dilihat pada Gambar 2.12.



Gambar 2.12 Process

Pada bentuk gambar *process*, bagian atas berisi nomor untuk identitas proses. Suatu proses dengan nomor 0 (nol atau kosong) menandakan bahwa proses tersebut adalah sebuah *context diagram*. Diagram ini merupakan level tertinggi dari DFD yang menggambarkan hubungan sistem dengan lingkungan luarnya. Pembuatan *context diagram* dapat dilakukan dengan terlebih dahulu menentukan nama sistemnya, menentukan batasan dari sistem, dan menentukan *terminator* yang diterima atau diberikan daripada sistem untuk kemudian dilakukan penggambaran.

Nomor 1, 2, 3, dan seterusnya menandakan bahwa proses tersebut diartikan sebagai proses level-0 (nol) yang merupakan hasil turunan atau *decompose* dari proses *context diagram*. Proses level-0 membahas sistem secara lebih mendetail, baik dipandang dari segi kegiatan dari sebuah bagian, alur data yang ada, maupun *database* yang digunakan didalamnya. Pembuatannya dapat dilakukan dengan cara menentukan proses utama yang ada dalam sistem, menentukan alur data yang diterima dan diberikan masing-masing proses daripada sistem sambil memperhatikan konsep keseimbangan (alur data yang masuk atau keluar dari suatu level harus sama dengan alur data yang masuk dan keluar pada level berikutnya), memunculkan *data store* sebagai sumber maupun tujuan data (*optional*), menggambarkan diagram level-0, menghindari perpotongan arus data, dan melakukan pemberian nomor pada proses utama (nomor tidak menunjukkan urutan proses).

Nomor 1.1, 1.2, 2.1, 2.2, dan seterusnya merupakan sebuah proses turunan atau *decompose* dari proses level-0 yang disebut sebagai proses level-1 (satu). Proses level-1 menggambarkan detil kerja dari sebuah bagian dalam sebuah sistem. Penggambarannya dilakukan dengan cara menentukan proses yang lebih kecil (*sub-proses*) dari proses utama yang ada di level-0, menentukan apa yang diterima atau diberikan masing-masing sub-proses daripada sistem dan tetap memperhatikan konsep keseimbangan, memunculkan *data store* sebagai sumber maupun tujuan alur data (*optional*), menggambar DFD level-1, dan berusaha untuk menghindari perpotongan arus data. Hasil turunan akhir disebut sebagai *the lowest level*, dimana hasil akhir ini tergantung dari kompleksitas sistem yang ada.

External entity disimbolkan dengan bentuk persegi yang digunakan untuk menggambarkan pelaku-pelaku sistem yang terkait, dapat berupa orang-orang, organisasi maupun instansi. *External entity* dapat memberikan masukan kepada *process* dan mendapatkan keluaran dari *process*. Simbol *external entity* dapat dilihat pada Gambar 2.13.

Entt_2
Gambar 2.13 External Entity

Data store digunakan sebagai media penyimpanan suatu data yang dapat berupa *file* atau *database*, arsip atau catatan manual, lemari *file*, dan tabel-tabel dalam *database*. Penamaan *data store* harus sesuai dengan bentuk data yang tersimpan pada *data store* tersebut, misalnya tabel pelanggan, tabel detail penjualan, tabel detail pembelian, dan lain-lain. Simbol *data store* dapat dilihat pada Gambar 2.14.

1 | Stor_3
Gambar 2.14 Data Store

Data flow merupakan penghubung antara *external entity* dengan *process* dan *process* dengan *data store*. *Data flow* menunjukkan aliran data dari satu titik ke titik lainnya dengan tanda anak panah mengarah ke tujuan data. Penamaan *data flow* harus menggunakan kata benda, karena di dalam *data flow* mengandung sekumpulan data. Simbol *data flow* dapat dilihat pada Gambar 2.15.

Flow_6
Gambar 2.15 Data Flow

2.12 Entity Relationship Diagram (ERD)

ERD merupakan notasi grafis dalam pemodelan data konseptual yang mendeskripsikan hubungan antar penyimpanan. ERD digunakan untuk memodelkan struktur data dan hubungan antar data, karena hal ini relatif kompleks. Dengan ERD kita dapat menguji model dan mengabaikan proses apa yang harus dilakukan. ERD dapat dikategorikan menjadi tiga bagian, yaitu :

1. *One to one relationship*

Jenis hubungan antar tabel yang menggunakan bersama sebuah kolom *primary key*. Jenis hubungan ini tergolong jarang digunakan, kecuali untuk alasan keamanan atau kecepatan akses data. Misalnya satu departemen hanya mengerjakan satu jenis pekerjaan saja dan satu pekerjaan hanya dikerjakan oleh satu departemen saja.

2. *One to many relationship*

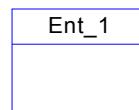
Jenis hubungan antar tabel dimana satu *record* pada satu tabel terhubung dengan beberapa *record* pada tabel lain. Jenis hubungan ini merupakan yang paling sering digunakan. Misalnya satu pekerjaan hanya dikerjakan oleh satu departemen saja, namun satu departemen dapat mengerjakan beberapa macam pekerjaan sekaligus.

3. *Many to many relationship*

Jenis hubungan ini merupakan hubungan antar tabel dimana beberapa *record* pada satu tabel terhubung dengan beberapa *record* pada tabel lain. Misalnya satu departemen mampu mengerjakan banyak pekerjaan, juga satu pekerjaan dapat ditangani oleh banyak departemen.

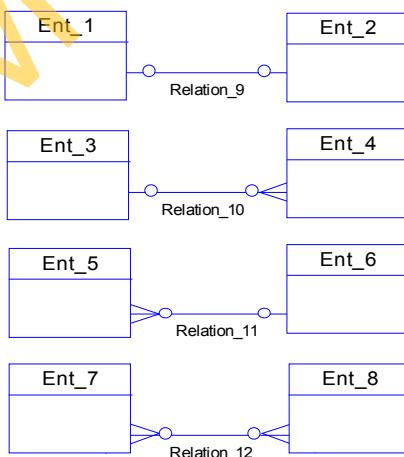
Entity Relationship Diagram, atau yang lebih dikenal dengan nama ERD, digunakan untuk mengimplementasikan, menentukan, dan mendokumentasikan kebutuhan-kebutuhan untuk sistem pemrosesan *database*. ERD menyediakan bentuk untuk menunjukkan struktur keseluruhan kebutuhan data dari pemakai. Adapun elemen-elemen yang terdapat pada ERD, adalah sebagai berikut:

1. *Entity* atau entitas, digambarkan dalam bentuk persegi seperti pada Gambar 2.16.



Gambar 2.16 Entity atau Entitas

2. *Relation* atau relasi merupakan penghubung antara entitas dengan entitas. Terdapat beberapa jenis relasi yang dapat digunakan, seperti *one-to-one*, *one-to-many*, *many-to-one*, dan *many-to-many*. Bentuk alur relasi secara detil dapat dilihat pada Gambar 2.17.



Gambar 2.17 Relation of Entity

2.13 Sistem Basis Data

Menurut Marlinda (2004:1), sistem basis data adalah suatu sistem menyusun dan mengelola *record* menggunakan komputer untuk menyimpan atau merekam serta memelihara dan operasional lengkap sebuah organisasi/perusahaan sehingga mampu menyediakan informasi optimal yang diperlukan pemakai untuk proses mengambil keputusan.

Pada sebuah sistem basis data terdapat komponen-komponen utama yaitu Perangkat Keras (*Hardware*), Sistem Operasi (*Operating System*), Basis Data (*Database*), Sistem (Aplikasi atau Perangkat lunak) pengelola Basis Data (DBMS), Pemakai (*User*), dan aplikasi (Perangkat lunak) lain (bersifat opsional).

2.14 SQL

SQL (*Structured Query Language*) adalah sebuah bahasa yang dipergunakan untuk mengakses data dalam basis data relasional. Bahasa ini secara *defacto* merupakan bahasa standar yang digunakan dalam manajemen basis data relasional. Secara umum, SQL terdiri dari 2 bahasa, yaitu *Data Definition Language* (DDL) dan *Data Manipulation Language* (DML). DDL digunakan untuk mendefinisikan, mengubah, serta menghapus basis data, dan objek-objek yang diperlukan dalam basis data, misalnya *table*, *view*, *user*, dan sebagainya. DML digunakan untuk memanipulasi data yang ada dalam suatu *table*.