BAB 2
TINJAUAN PUSTAKA

2.1 Profil Institusi

Iteung Pusat Oleh-oleh Bandung, merupakan tempat penjualan cinderamata dan oleh-oleh khas Bandung yang telah ada sejak tahun 2013. Toko ini berada di Jalan Anggrek No. 55 Bandung dan telah memiliki beberapa cabang di antaranya di Carrefour-Transmart, Yogya Kepatihan, Cihampelas Batara Hotel, dan Natural Lembang. Iteung Pusat Oleh-oleh Bandung merupakan toko oleh-oleh dengan format swalayan sehingga konsumen dapat memilih dan mengambil sendiri produk yang diinginkan. Produk yang ditawarkan oleh toko Iteung yaitu berupa makanan atau camilan khas Bandung, pakaian, tas, hiasan atau pajangan, dan beragam kerajinan tangan yang hampir semuanya berasal dari daerah Bandung.

2.1.1 Logo

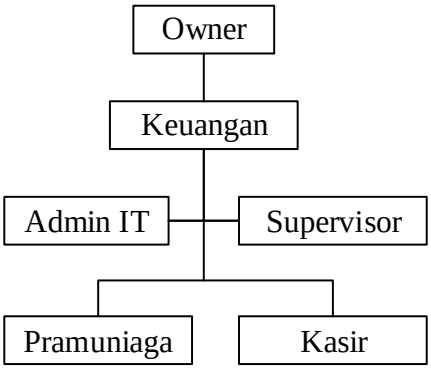
Berikut ini adalah logo Iteung Pusat Oleh-oleh Bandung



Gambar 2.1 Logo Iteung Pusat Oleh-oleh Bandung

2.1.2 Struktur Organisasi

Struktur organisasi perusahaan adalah gambar suatu bagan yang menerangkan posisi atau jabatan dan struktur kerja pegawai dalam suatu perusahaan. Berikut ini adalah struktur organisasi di Iteung Pusat Oleh-oleh Bandung.



Gambar 2.2 Struktur Organisasi di Iteung Pusat Oleh-oleh Bandung

2.1.3 Deskripsi Kerja

1. *Owner*

   Selain menjadi pemilik perusahaan, *owner* bertugas untuk mengawasi kerja para karyawan, pengambil keputusan dan penanggung jawab semua aktivitas yang berada di Iteung Pusat Oleh-oleh Bandung.

2. Keuangan

   Bagian keuangan bertugas merencanakan, menganggarkan, memeriksa, mengelola, dan menyimpan dana yang dimiliki oleh perusahaan.

3. Admin IT

   Admin IT bertugas mengelola semua data yang ada di Iteung Pusat Oleh-oleh Bandung seperti data penjualan, data gudang, data produk dan lain – lain.

4. *Supervisor*

   Supervisor bertugas untuk Mengawasi semua aktivitas karyawan dan membatu aktivitas di Iteung Pusat Oleh-oleh Bandung.

5. Pramuniaga

   Bertugas untuk melayani *customer* yang akan membeli produk di Toko, merapikan produk, juga menyimpan letak produk untuk dipajang.

6. Kasir

   Bertugas untuk melayani konsumen yang melakukan transaksi di toko Iteung.

2.2 Landasan Teori

Landasan teori membahas tentang materi atau teori apa saja yang digunakan sebagai bahan acuan dalam pembuatan tugas akhir ini. Landasan teori yang diuraikan merupakan hasil studi literatur baik dari buku atau situs internet.

2.2.1 Data

Data adalah segala fakta, angka atau teks yang dapat diproses oleh komputer. Data dapat digunakan sebagai *input* dan menghasilkan sebuah informasi. Data adalah sesuatu yang belum memiliki arti dan masih membutuhkan suatu pengolahan. Dalam data terdapat himpunan data (*data-set*) yang merupakan kumpulan dari objek dan atributnya. Atribut merupakan sifat atau karakteristik dari suatu objek yang biasanya dikenal sebagai *variabel*, *field*, karakteristik atau fitur.

Salah satu himpunan data (*data-set*) adalah *record* data, yaitu data yang terdiri dari sekumpulan *record*, yang masing-masing terdiri dari satu set atribut yang tetap. Salah satu yang termasuk dalam tipe data *record* yaitu data transaksi. Data transaksi merupakan sebuah tipe khusus dari *record* data, di mana tiap *record* (transaksi) meliputi satu *set item* [7].

2.2.2 *Database*

Database adalah susunan *record* data operasional lengkap dari suatu organisasi atau perusahaan yang diorganisir dan disimpan secara terintegrasi dengan menggunakan metode tertentu dalam komputer sehingga mampu memenuhi informasi yang optimal yang dibutuhkan oleh para pengguna [8].

2.2.2.1 Data Pada *Database* dan Hubungannya

Ada 3 jenis data pada sistem *database*, yaitu [8]:

1. Data operasional dari suatu organisasi, berupa data yang tersimpan dalam basis data.
2. Data masukan (*input data*), data dari luar sistem yang dimasukkan melalui peralatan *input* (*keyboard*), yang dapat mengubah data operasional.

3. Data keluaran (*output data*), berupa laporan melalui peralatan *output* sebagai hasil dari dalam sistem yang mengakses data operasional.

2.2.2.2 Keuntungan dan kerugian Pemakai Sistem Database

Berikut adalah keuntungan pemakai sistem *database* [8]:

1. Terkontrolnya kerangkapan data dan inkonsistensi.
2. Terpeliharanya keselarasan data.
3. Data dapat dipakai secara bersama-sama.
4. Memudahkan penerapan standarisasi dan batas-batas pengamanan.
5. Terpeliharanya keseimbangan atas perbedaan kebutuhan data dari setiap aplikasi.
6. Program/data *independent*.

Di samping keuntungan, terdapat kerugian pemakai sistem *database*, yaitu:

1. Mahal dalam implementasinya.
2. Rumit/kompleks.
3. Penanganan proses *recovery backup* sulit.
4. Kerusakan pada sistem basis data dapat mempengaruhi.

2.2.3 Pengertian *Data Mining*

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan dalam *database*. *Data mining* merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar [9].

Menurut Pramudiono [2], "*Data mining* adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data."

Menurut Larose [9], 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 handal.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetensi bisnis untuk meningkatkan penguasaan pasar dalam *global*isasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk *data mining* (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

2.2.3.1 Konsep *Data Mining*

Data mining sangat diperlukan terutama dalam mengelola data yang sangat besar untuk memudahkan aktivitas *recording* suatu transaksi dan untuk proses *data warehousing* agar dapat memberikan informasi yang akurat bagi pengguna *data mining*. Alasan utama *data mining* sangat dibutuhkan dalam industri informasi karena tersedianya data dalam jumlah yang besar dan semakin besarnya kebutuhan untuk mengubah data tersebut menjadi informasi dan pengetahuan yang berguna karena sesuai fokus bidang ilmu ini yaitu melakukan kegiatan mengekstraksi atau menambang pengetahuan dari data yang berukuran atau berjumlah besar. Informasi inilah yang nantinya sangat berguna untuk pengembangan.

2.2.3.2 Metode *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu [9]:

1. Deskripsi

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

pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

   Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun dengan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

3. Prediksi

   Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang. Contoh prediksi dalam bisnis dan penelitian adalah:

   a. Prediksi harga beras dalam tiga bulan yang akan datang.
   b. Prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah dinaikkan.

4. Klasifikasi

   Dalam klasifikasi, terdapat target variabel kategori. sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori yaitu: pendapatan tinggi, pendapatan sedang, dan pendapatan rendah. Contoh lain klasifikasi dalam bisnis dan penelitian adalah:

   a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
   b. Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit 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 satu dengan yang lainnya dan tidak memiliki kemiripan dengan *record-record* dalam kluster lain. Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam suatu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal. Contoh pengklusteran dalam bisnis dan penelitian adalah:

a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk sebuah perusahaan yang tidak memiliki dana pemasaran yang besar.
b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik maupun mencurigakan.

6. Asosiasi

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

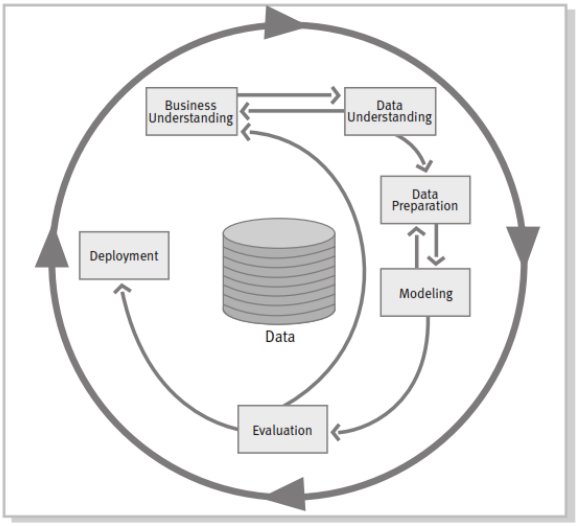
a. Meneliti jumlah konsumen dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran *upgrade* layanan yang diberikan.
b. Menemukan produk dalam supermarket yang dibeli secara bersamaan dan produk yang tidak pernah dibeli secara bersamaan.

2.2.3.2.1 *Cross-Industry Standard Process for Data Mining* (CRISP-DM)

Cross-Indutry Standard Prosess for Data Mining (CRISP-DM) yang dikembangkan tahun 1996 oleh analis 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. Keseluruhan fase berurutan yang ada tersebut bersifat adaptif. Fase berikutnya dalam urutan bergantung kepada keluaran dari fase sebelumnya. Hubungan antar fase digambarkan dengan panah. Sebagai contoh, jika proses berada pada fase *modelling*. Berdasar pada perilaku dan karakteristik model, proses mungkin harus kembali kepada fase *data preparation* untuk perbaikan lebih lanjut terhadap data atau berpindah maju kepada fase *evaluation*.

Enam fase yang ada dalam metode CRISP-DM tersebut dapat dilihat pada Gambar berikut ini:



Gambar 2.3 Fase metode CRISP-DM

Fase-fase dari CRISP-DM [9]:

1. *Business Understanding*
   a. Penentuan tujuan bisnis yang dilakukan oleh perusahaan dan kriteria sukses bisnis yang menjadi acuan berhasilnya perusahaan dalam menjalankan bisnis.
   b. Menjelaskan tujuan digunakannya metode *data mining* dalam bsnis yang dijalankan dan kriteria suksesnya metode *data mining*.
   c. Menilai situasi dari sumber daya yang dimiliki oleh perusahaan yang dapat mendukung proses *data mining*.

2. *Data Understanding*
    a. Mengumpulkan data.
    b. Mendeskripsikan data yang telah diperoleh.
    c. Mengevaluasi kualitas data.
    d. Jika diinginkan, pilih sebagian kecil grup data yang mungkin mengandung pola dari permasalahan.
3. *Data Preparation*
    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 dengan analisis yang akan dilakukan.
    c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
    d. Siapkan data awal sehingga siap untuk perangkat pemodelan.
4. *Modeling*
    a. Pilihan 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. *Evaluation*
    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. *Deployment*
   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.

2.2.3.3 *Association Rule*

Association rule adalah salah satu teknik utama atau prosedur dalam *market basket analysis* untuk mencari hubungan antar-item dalam suatu *data set* dan menampilkan bentuk *associaiton rule* [10]. *Association rule* (aturan asosiasi) akan menemukan pola tertentu untuk mengasosiasikan data yang satu dengan data yang lain.

Untuk mencari *association rule* dari suatu kumpulan data, tahap pertama yang harus dilakukan adalah mencari *frequent itemset* terlebih dahulu. *Frequent itemset* adalah sekumpulan item yang sering muncul secara bersamaan. Setelah semua pola *frequent item* ditemukan, barulah mencari aturan *assosiatif* atau aturan keterkaitan yang memenuhi syarat yang telah ditentukan.

Jika diasumsikan bahwa produk yang dijual di swalayan adalah semesta, maka setiap produk akan memiliki *boolean* variabel yang akan menunjukkan keberadaannya atau tidak produk tersebut dalam suatu transaksi atau satu keranjang belanja. Pola *boolean* yang didapat digunakan untuk menganalisa produk yang dibeli secara bersamaan. Pola tersebut dirumuskan dalam sebuah *association rule*. Sebagai contoh konsumen biasanya akan membeli kopi dan susu yang ditunjukkan sebagai berikut:

Kopi → susu [*support* = 2%, *confidence* = 60%]

Association rule diperlukan suatu variabel ukuran yang ditentukan sendiri oleh *user* untuk menentukan batasan sejauh mana atau sebanyak apa *output* yang diinginkan *user*. *Support* dan *confidence* adalah sebuah ukuran kepercayaan dan

kegunaan suatu pola yang telah ditemukan, Nilai *support* 2% menunjukkan bahwa keseluruhan dari total transaksi konsumen membeli kopi dan susu secara bersamaan yaitu sebanyak 2%. Sedangkan *confidence* 60%, yaitu menunjukkan bila konsumen membeli kopi dan pasti membeli susu sebesar 60%.

Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter yaitu *support* dan *confidence*. *Support* (nilai penunjang) adalah persentase kombinasi item tersebut dalam *database*, sedangkan *confidence* (nilai kepastian adalah kuatnya hubungan antar–item dalam aturan asosiasi.

Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapat dari hasil pengolahan data dengan perhitungan tertentu. Umumnya ada dua ukuran, yaitu:

1. *Support*, suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item/itemset* dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu *item/itemset* layak untuk dicari *confidence*-nya (misal, dari keseluruhan transaksi yang ada, seberapa besar tingkat dominasi suatu item yang menunjukkan bahwa item A dan item B dibeli bersamaan).
2. *Confidence*, suatu ukuran yang menunjukkan hubungan antara 2 item secara *conditional* (misal, menghitung kemungkinan seberapa sering item B dibeli oleh konsumen jika konsumen tersebut membeli sebuah item A).

Kedua ukuran ini nantinya berguna dalam menentukan kekuatan suatu pola dengan membandingkan pola tersebut dengan nilai minimum kedua parameter tersebut yang ditentukan oleh pengguna. Bila suatu pola memenuhi kedua nilai *minimum parameter* yang sudah ditentukan sebelumnya, maka pola tersebut dapat disebut sebagai *interesting rule* atau *strong rule*.

2.2.3.3.1 Metodologi Dasar Analisis Asosiasi

Metodologi dasar *Association Rule* terbagi menjadi dua tahap, yaitu [11]:

1. Analisis pola frekuensi tinggi

   Tahap Tahap ini mencari pola *item* yang memenuhi syarat minimum dari nilai *support* dalam database. Menurut Larose, kita bebas menentukan nilai *minimum support* (*minsup*) dan *minimum confidence* (*mincof*) sesuai

kebutuhan [9]. Sebagai contoh, bila ingin menemukan data-data yang memiliki hubungan asosiasi yang kuat, *minsup* dan mincof-nya bisa diberi nilai yang tinggi. Sebaliknya, bila ingin melihat banyaknya variasi data tanpa terlalu mempedulikan kuat atau tidaknya hubungan asosiasi antara *item*-nya, nilai *minsup* dan mincofnya dapat diisi rendah [12]. Nilai *support* sebuah *item* diperoleh dengan rumus berikut.

$$\boldsymbol{Support}(\boldsymbol{A}) = \frac{\boldsymbol{Jumlah\ transaksi\ yang\ mengandung\ A}}{\boldsymbol{Jumlah\ total\ transaksi}} \times \mathbf{100}\% \textbf{ ... Persamaan 2.1}$$

Persamaan 2-1 menjelaskan bahwa nilai *support* didapat dengan cara membagi jumlah transaksi yang mengandung *item* A (satu *item*) dengan jumlah total seluruh transaksi.

Sementara itu, nilai *support* dari 2 *item* diperoleh dari rumus berikut

$Support(A, B) = P(A \cap B)$

$$\boldsymbol{Support}(\boldsymbol{A}, \boldsymbol{B}) = \frac{\boldsymbol{Jumlah\ transaksi\ yang\ mengandung\ A\ dan\ B}}{\boldsymbol{Jumlah\ total\ transaksi}} \times \mathbf{100}\% \textbf{ ... Persamaan 2.2}$$

Persamaan 2-2 menjelaskan bahwa nilai *support item set* didapat dengan cara membagi jumlah transaksi yang mengandung *item* A dan *item* B (*item* pertama bersamaan dengan *item* yang lain) dengan jumlah total seluruh transaksi.

2. Pembentukan aturan asosiasi

   Setelah semua pola frekuensi tinggi ditemukan, kemudian mencari aturan asosiasi yang cukup kuat ketergantungan antar item.

   Dalam *antecedent* (pendahulu) dan *consequent* (pengikut) serta memenuhi syarat minimum untuk *confidence* aturan asosiatif A→B.

   Misalkan D adalah himpunan transaksi, di mana setiap transaksi T dalam D merepresentasikan himpunan item yang berada dalam I. I adalah himpunan item yang dijual. Misalkan kita memilih himpunan item A dan himpunan item lain B, kemudian aturan asosiasi akan berbentuk:

   Jika A, maka B (A→B)

   Dimana *antecedent* A dan *consequent* B merupakan *subset* dari I, dan A dan B di mana aturan:

   Jika A, maka B

   Tidak berarti

Jika B, maka A

Sebuah *item set* adalah himpunan item-item yang ada dalam I, dan i *item set*. Frekuensi *item set* merupakan *itemset* yang memiliki frekuensi kemunculan lebih dari nilai minimum yang telah ditentukan.

Nilai *confidence* dari aturan A→B diperoleh dari rumus berikut.

$$\boldsymbol{Confidence} = \frac{\boldsymbol{Jumlah\ transaksi\ yang\ mengandung\ A\ dan\ B}}{\boldsymbol{Jumlah\ transaksi\ yang\ mengandung\ A}}$$ **... Persamaan 2.3**

2.2.3.3.2 Algoritma CT-PRO

Algoritma ini merupakan salah satu algoritma pengembangan dari *FP-Growth*. Perbedaannya adalah di mana *FP-Growth* membuat *FP-Tree* sedangkan *CT-PRO* membuat *Compressed FP-Tree* (*CFP-Tree*) [13]. Pada tahap *mining* algoritma *CT-PRO* juga menggunakan pendekatan *bottom-up* dimana *item* pada *item* tabel dan *CFP-Tree* dilakukan *scan* dari jumlah terkecil hingga terbesar.

Algoritma *CT-PRO* memiliki tiga tahap yaitu:

1. Menemukan *item-item* yang *frequent.*
2. Membuat struktur data *CFP-Tree.*
3. Melakukan *mining frequent patterns.*

Langkah-langkah kerja algoritma *CT-PRO* [13] [14]:

1. Mencari *Frequent Item*, pada tahap ini terjadi proses-proses sebagai berikut:
   a. Dari *dataset* yang ada, dilakukan seleksi berdasarkan *minimum support* yang ditentukan sehingga menghasilkan *frequent item.*
   b. Dari *frequent item* yang telah terbentuk, dihitung frekuensi kemunculan setiap *item* sehingga menghasilkan *Global Item* tabel.
2. Membangun *CFP-Tree*, pada tahap ini terjadi proses-proses sebagai berikut:
   a. *Frequent item* yang telah didapatkan, diurutkan berdasarkan *Global item table* yang ada secara menurun (diurutkan mulai dari *item* berfrekuensi terbesar hingga terkecil).
   b. Dengan *frequent item* yang telah terurut ini dibentuk *Global CFP-Tree*, aturan pembentukan *Global CFP-Tree* sebagai berikut:

1) *CFP-Tree* terdiri dari *tree* yang memiliki *root* yang mewakili indeks dari *item* dengan tingkat kemunculan tertinggi dan kumpulan *subtree* sebagai anak dari *root*.
2) Jika I = {i1,i2, …, ik} adalah kumpulan dari *frequent item* dalam transaksi, *item* dalam transaksi akan dimasukkan ke dalam *CFP-Tree* dimulai dari *root subtree* yang merupakan i1 dalam *item* tabel.
3) *Root* dari *CFP-Tree* merupakan *level*-0 dari *tree*.
4) Setiap *node* dalam *CFP-Tree* memiliki empat *field* utama yakni *item-id*, *parent-id*, *Count* yang merupakan jumlah *item* pada *node* tersebut, dan *level* yang menunjukkan struktur data *tree* pada *node* tersebut dimulai dari *item* yang terdapat pada *item* tabel dengan *level* yang terdapat pada *CFP-Tree*.

3. *Mining*, pada tahap ini terjadi proses-proses sebagai berikut [13] [14]:
   a. Pada tahap *mining* ini, algoritma *CT-PRO* bekerja dengan melakukan *bottom-up mining* sehingga *Global item table* diurutkan mulai dari *item* berfrekuensi terkecil hingga terbesar.
   b. Untuk setiap *item* yang terdaftar pada *Global item table* yang telah diurutkan, dilakukan pencarian *node* yang berkaitan dengan *item* tersebut pada *Global CFP-Tree*. Dari semua *node* yang d*item*ukan untuk setiap *item* inilah yang disebut dengan *Local frequent item* dan digunakan untuk membuat *Local item table*.
   c. Pada pembuatan *Local item table* ini juga dilakukan berdasarkan jumlah *minimum support* yang telah ditentukan.
   d. Setelah itu, dibuat *Local CFP-Tree* berdasarkan *Local item table* yang terbentuk. Aturan pembentukan *Local CFP-Tree* sama dengan pembentukan *Global CFP-Tree,* yang membedakan adalah pada *Global CFP-Tree* yang digunakan dalam pembentukan *tree*-nya adalah *Global item table* yang terbentuk dari *Global item table* data sedangkan pada *Local CFP-Tree* yang digunakan dalam pembentukan *tree*-nya adalah *Local item table* yang terbentuk dari *Local frequent item.*

e. Dari *Local CFP-Tree* dibentuk *frequent pattern* sesuai dengan *item* yang di-*mining*.
f. Dari *frequent pattern* dihitung masing-masing *item* yang memenuhi dihitung *confidence*nya. Apabila memenuhi minimum *confidence* maka masing-masing *item* yang bersangkutan dijadikan sebagai *knowledge*.

2.2.4 Alat-Alat Pemodelan Sistem

Alat-alat pemodelan sistem adalah alat dan metode yang akan digunakan untuk memodelkan perancangan perangkat lunak dalam skripsi ini.

2.2.4.1 *Entity Relationship Diagram* (ERD)

ERD Menurut Nugroho (2002) adalah diagram yang memperlihatkan entitas-entitas yang terlibat dalam suatu sistem serta hubungan-hubungan (relasi) antar entitas tersebut.

2.2.4.2 Diagram Konteks

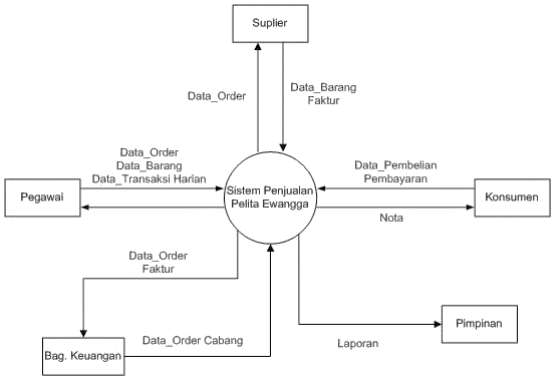
Diagram konteks adalah diagram aliran data pada tingkat paling atas yang merupakan penggambaran yang berfungsi untuk memperlihatkan interaksi atas hubungan langsung antara sistem informasi dengan lingkungannya. Diagram konteks menggambarkan sebuah sistem berupa sebuah proses yang berhubungan dengan satu atau beberapa entitas.

2.2.4.3 *Data Flow Diagram* (DFD)

Data flow diagram (DFD) adalah suatu alat pemodelan yang digunakan untuk memodelkan fungsi dari sistem, menggambarkan secara rinci mengenai sistem sebagai jaringan kerja antar fungsi yang berhubungan satu sama lain dengan menunjukkan dari dan kemana data mengalir serta menyimpannya. Pada umumnya DFD dimulai dari 0, 1, 2, dst. Level ke-0 disebut dengan diagram konteks yang menggambarkan sistem secara *global*.
Adapun level dalam DFD adalah:

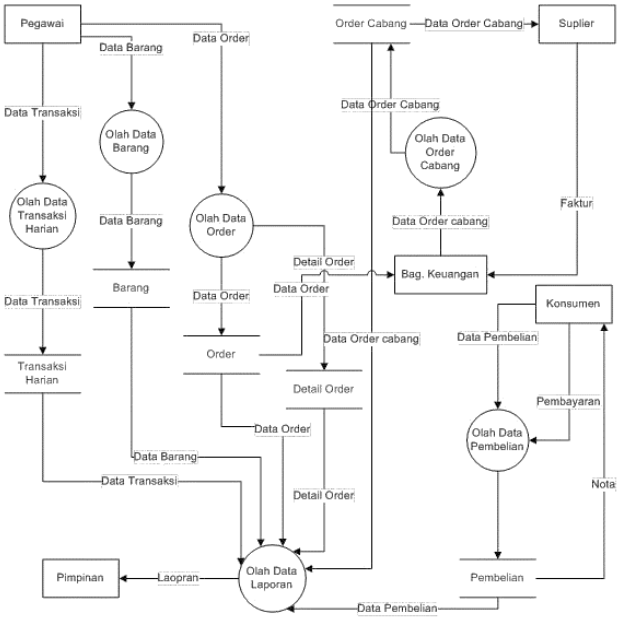
a. Diagram Konteks, merupakan satu lingkaran besar yang dapat mewakili seluruh proses yang terdapat di dalam suatu sistem.
b. DFD Level 1 merupakan satu lingkaran besar yang mewakili lingkaran-lingkaran kecil yang ada di dalamnya, dan merupakan pemecahan dari diagram konteks.
c. Diagram Rinci, merupakan diagram yang menguraikan proses apa yang ada dalam diagram DFD Level 1.

Contoh kasus dari penerapan Diagram Konteks:



Gambar 2.4 Contoh Kasus Penerapan Diagram Konteks

Contoh DFD dari diagram konteks di atas:



Gambar 2.5 Contoh Penerapan DFD

2.2.4.4 Kamus Data

Kamus data adalah suatu penjelasan tertulis mengenai data yang berada dalam *database*. Kamus data ikut berperan untuk menjelaskan arti aliran data dan menyimpan dalam menggambarkan *data flow diagram*, menjelaskan spesifikasi nilai dan satuan yang relevan terhadap data yang mengalir dalam sistem tersebut, mendeskripsikan komposisi paket data yang bergerak melalui aliran.

Fungsi kamus data [9]:

1. Menjelaskan arti aliran data dan penyimpanan data dalam DFD.
2. Mendeskripsikan komposisi paket data (elemen data) yang bergerak melalui aliran data.
3. Mendeskirpsikan komposisi penyimpanan data.
4. Mendeskripsikan nilai dan satuan (struktur data) yang relevan bagi penyimpanan dan aliran.
5. Mendeskripsikan hubungan detail antar penyimpanan.

Tabel 2.1 Struktur Kamus Data

Nama Data	Keterangan
Deskripsi	Memilih *file* yang akan diproses
Bentuk Data	Format *file*
Struktur Data	Atribut yang digunakan

2.2.4.5 Spesifikasi Proses (*Process Specification*)

Digunakan untuk mendeskripsikan proses yang terjadi pada level paling rendah dari DFD. Struktur spesifikasi proses dapat dilihat di tabel lampiran.

Hubungan dengan *Prosess Specification* [11]:

1. Semua proses dalam DFD yang tidak dapat dipecah lagi dan harus didefinisikan dalam *process specification*.
2. Aliran data masuk (*input*) dan keluar (*output*) dalam DFD dan hubungan ke *data store* harus sesuai dan relevan dalam *process specification*.
3. Setiap elemen data dalam *process specification* harus:
    a. Nama dari aliran data atau *data store*.
    b. Atau komponen dalam kamus data atau *data store* yang berhubungan dalam DFD.

Tabel 2.2 Struktur Spesifikasi Proses

No. Urut	Proses	Keterangan
	No. Proses	
	Nama Proses	
	Source	
	Input	
	Output	
	Destination (tujuan)	
	Logika Proses	

2.2.5 MySQL

SQL (*Structured Query Language)* adalah sebuah bahasa yang dipergunakan untuk mengakses data dalam basis data relasional. SQL juga dapat diartikan sebagai antar muka standar untuk sistem manajemen basis data relasional, termasuk sistem yang beroperasi pada komputer pribadi. SQL memungkinkan seorang pengguna untuk mengakses informasi tanpa mengetahui di mana lokasinya atau bagaimana informasi tersebut disusun [15].

MySQL adalah sebuah perangkat lunak sistem manajemen basis data SQL (bahasa Inggris: *database management system*) atau DBMS yang *multi thread*, dan *multi-user*. MySQL dimiliki dan di sponsori oleh sebuah perusahaan komersial Swedia MySQL AB, di mana memegang hak cipta hampir atas semua kode sumbernya. Kedua orang Swedia dan satu orang Finlandia yang mendirikan MySQL AB adalah: David Axmark, Allan Larsson, dan Michael "Monty" Widenius.

2.2.6 PHP

PHP adalah bahasa pemrograman *script server-side* yang didesain untuk pengembangan web. Selain itu, PHP juga bisa digunakan sebagai bahasa pemrograman umum. PHP di kembangkan pada tahun 1995 oleh Rasmus Lerdorf, dan sekarang dikelola oleh The PHP Group.

PHP disebut bahasa pemrograman *server side* karena PHP diproses pada komputer *server*. Hal ini berbeda dibandingkan dengan bahasa pemrograman *client-side* seperti JavaScript yang diproses pada *web browser (client).*

Pada awalnya PHP merupakan singkatan dari *Personal Home Page*. Sesuai dengan namanya, PHP digunakan untuk membuat *website* pribadi. Dalam beberapa

tahun perkembangannya, PHP menjelma menjadi bahasa pemrograman web yang powerful dan tidak hanya digunakan untuk membuat halaman web sederhana, tetapi juga *website* populer yang digunakan oleh jutaan orang seperti wikipedia, wordpress, joomla, dll. Saat ini PHP adalah singkatan dari *PHP: Hypertext Preprocessor*.