BAB 2
TINJAUAN PUSTAKA

2.1. Tinjauan Tempat Penelitian

Pada bagian tinjauan tempat penelitian ini akan dijelaskan beberapa hal mengenai tempat penelitian. Tinjauan penelitian meliputi profil perusahaan, logo perusahaan, visi dan misi, serta struktur organisasi perusahaan.

2.1.1. Profil Perusahaan

Mimosabi Handmadeshoes berdiri pada tanggal 5 April 2008 di Cimahi, didirikan oleh Venny Ahlani dan istrinya Lulu Sayyidatu dan memulai pemasarannya dengan metode *online marketing.* Saat ini Mimosabi telah memiliki 44 distributor resmi yang berada di beberapa kota-kota besar di Indonesia.

Pada saat itu Mimosabi memproduksi sepatunya berdasarkan permintaan. Pembuatan sepatu dilakukan di bengkel sepatu di Cibaduyut dibantu oleh rekan. Tanggapan pasar yang positif ditandai dengan datangnya pesanan dalam jumlah yang besar dari berbagai kota di Indonesia. Karena alasan tersebut Mimosabi menyiapkan tempat yang lebih dekat yaitu di daerah Cimahi.

Mimosabi tidak hanya mendapatkan tanggapan positif dari para penggemarnya, tetapi banyak dilirik juga oleh media cetak dan elektronik yang kemudian banyak meliput kegiatan usaha Mimosabi, adapun media – media yang pernah meliput Mimosabi antara lain TransTV, NET, MetroTV, STV, MQTV, BandungTV, Tabloid Peluang Usaha, Business Opportunity, KOMPAS, Pikiran Rakyat, SINDO, dan lain-lain.

Liputan tersebut menambah tanggapan pasar menjadi jauh lebih besar sehingga banyak orang yang tertarik untuk menjadi distributor di kotanya masing-masing dan pada tahun 2012 Mimosabi sudah memiliki distributor di 44 kota.

Pada saat ini Mimosabi mengadakan perubahan manajemen dalam rangka lebih memfokuskan potensi dalam pengadaan produk dalam jumlah besar demi memenuhi permintaan pasar yang masih sangat potensial perkembangannya.

2.1.2. Logo Perusahaan

Logo dari CV. Mimosabi dapat dilihat pada Gambar 2.1.



Gambar 2.1 Logo CV. Mimosabi

2.1.3. Visi dan Misi

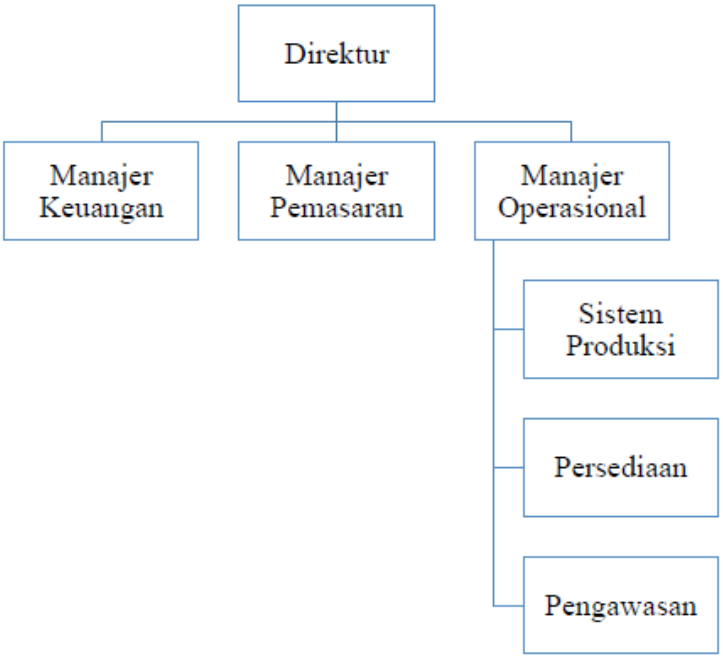
Berikut ini merupakan visi dan misi dari CV. Mimosabi :

Visi : Menjadi *brand* sepatu *handmade* terbesar di Indonesia.

Misi : Menebar manfaat dengan berbagi pekerjaan untuk kemajuan bersama.

2.1.4. Struktur Organisasi Perusahaan

Struktur organisasi CV. Mimosabi dapat dilihat pada Gambar 2.2.



Gambar 2.2 Struktur Organisasi CV. Mimosabi

Keterangan posisi :

1. Direktur

   Direktur mempunyai tugas untuk menawarkan visi dan misi, memimpin rapat umum, mengambil keputusan, menjalankan tanggung jawab, melakukan koordinasi peningkatan mutu pelayanan, penelitian, pengabdian dan kerja sama, bertanggung jawab atas semua proses bisnis perusahaan.

2. Manajer Keuangan

   Manajer keuangan mempunya tugas dalam membuat perencanaan keuangan, membuat rencana pemasukan dan pengeluaran serta kegiatan-kegiatan lainnya untuk periode tertentu, mencari dan mengolah sumber dana yang ada untuk operasional kegiatan perusahaan, mengumpulkan dana perusahaan serta menyimpan dan mengamankan dana tersebut.

3. Manajer Pemasaran

   Manajer pemasaran mempunyai tugas merancang dan melakukan kegiatan pemasaran, merumuskan strategi dan melakukan koordinasi kegiatan promosi.

4. Manajer Operasional

   Manajer operasional mempunyai tugas mengambil keputusan yang berkaitan dengan fungsi operasi dan sistem transformasi, perancangan dan desain.

5. Sistem produksi

   Sistem produksi mempunyai tugas pokok untuk menjalankan proses pembuatan dan pengiriman produk.

6. Persediaan

   Bagian persediaan bertanggung jawab dalam penyediaan pergudangan dan penyediaan bahan baku.

7. Pengawasan

   Bagian pengawasan mempunyai tugas pokok menjaga, merawat, dan memperbaiki peralatan mesin-mesin yang ada. Dan mengawasi sistem pengiriman produk.

2.2. Landasan Teori

Pada bagian landasan teori ini akan menjelaskan teori-teori yang akan digunakan sebagai acuan dalam penelitian yang akan dilakukan.

2.2.1. Data

Data adalah suatu nilai yang merepresentasikan fakta di dunia nyata dan mewakili suatu objek seperti manusia (pegawai, siswa, pembeli, pelanggan), barang, hewan, peristiwa, konsep, keadaan, dsb. Nilai tersebut lalu direkam dalam bentuk angka, huruf, simbol, teks, gambar, atau kombinasinya [3].

2.2.2. Basis Data

Basis data adalah sekumpulan data yang saling terhubung dan mempunyai skema tertentu.

Menurut Fathansyah [3], basis data sendiri dapat didefinisikan dalam sejumlah sudut pandang seperti :

1. Himpunan kelompok data (arsip) yang saling berhubungan yang diorganisasi sedemikian rupa agar kelak dapat dimanfaatkan kembali dengan cepat dan mudah.
2. Kumpulan data yang saling berhubungan yang disimpan secara bersama sedemikian rupa dan tanpa pengulangan (redudansi) yang tidak perlu, untuk memenuhi berbagai kebutuhan.
3. Kumpulan *file* atau tabel atau arsip yang saling berhubungan yang disimpan dalam media penyimpanan elektronis.

2.2.3. *Database Management System*

Database management system adalah sistem yang digunakan untuk membuat dan mengelola basis data. Tujuan dari *DBMS* adalah menyediakan lingkungan yang mudah dan aman untuk penggunaan dan perawatan *database*.

Dalam *DBMS* pengguna difasilitasi untuk mendefinisikan skema basis data yang disebut sebagai *data definition language (DDL).* Selain itu pengguna juga

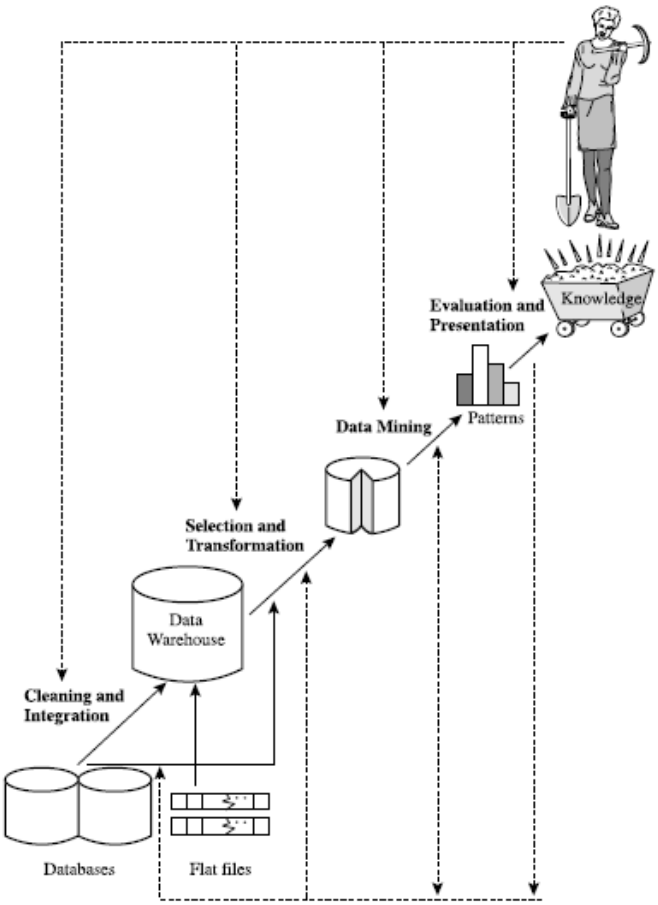
dapat membuat, mengubah, dan membuat *query* data yang biasanya disebut dengan *data manipulation language* (DML) [4].

2.2.4. Data Mining

Data *mining* adalah kegiatan mengekstraksi atau menambang pengetahuan dari data yang berukuran atau berjumlah besar. Nama lain yang sering digunakan dan memiliki makna yang sama dengan data *mining* adalah *knowledge mining from data*, *knowledge extraction*, *pattern analysis*, *data archaeology*, *data dredging*.

Data *mining* berusaha memecahkan masalah dengan menganalisis sekumpulan data yang sudah ada dalam basis data. Sedangkan hasil dari proses dari data mining adalah penemuan pengetahuan, informasi, pola, atau model.

Sebagian lain melihat data *mining* hanya sebagai satu langkah penting pada proses KDD (*Knowledge Discovery from Data*). Gambar 2.3 merupakan langkah-langkah dari proses KDD [5].



Gambar 2.3 Tahapan KDD [5]

a. Data *cleaning and Integration*

   Data *cleaning* berguna untuk menghilangkan *noise* data yang tidak konsisten serta mengintegrasikan data, di mana sumber data yang terpecah dapat disatukan.

b. Data *selection and transformation*

   Memilih data yang sesuai dengan analisis yang akan dilakukan dan mengubahnya agar sesuai untuk proses data *mining*..

c. Data *mining*

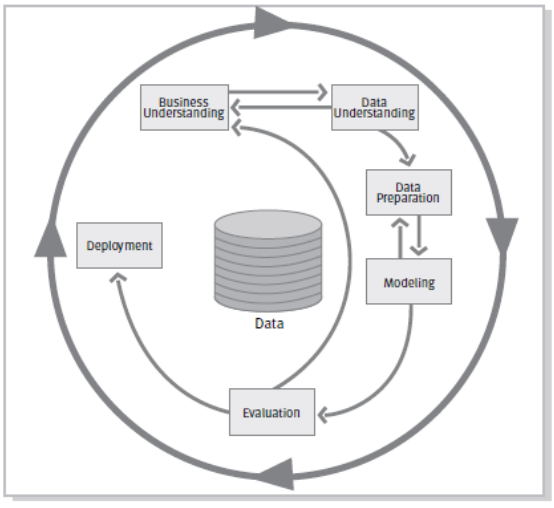
   Proses penerapan metode yang digunakan untuk menambang pola atau pengetahuan dari data.

d. *Evalution and Knowledge Presentation*

   Mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan serta memvisualisasikannya kepada pengguna.

2.2.5. Metode CRISP - DM

CRISP – DM (*Cross – Industry Standard of Data Mining*). CRISP – DM merupakan metodologi data *mining* yang digambarkan dengan model proses dan di dalamnya terdapat tugas-tugas yang harus dilakukan untuk melakukan proyek data *mining* [2]. Tahapan dari CRISP – DM dapat dilihat pada



Gambar 2.4 Model Proses CRISP - DM

1. Business *Understanding*

   Tahap awal ini berfokus pada pemahaman tujuan dan persyaratan proyek dari perspektif bisnis, kemudian mengubah pengetahuan tersebut ke dalam

definisi masalah data *mining* dan rencana awal dirancang untuk mencapai tujuan tersebut. Tahap pemahaman bisnis dibagi menjadi beberapa langkah, antara lain :

a. Penentuan Tujuan Bisnis

   Pada tahap ini akan menghasilkan *output* berupa latar belakang dan tujuan bisnis.

b. Penilaian Situasi

   Pada tahap ini akan menghasilkan *output* berupa inventori sumber daya, kendala dan resiko, serta biaya dan manfaat.

c. Penentuan Tujuan Data Mining

   Pada tahap ini akan menghasilkan *output* berupa sasaran data *Mining.*

d. Membuat Rencana Proyek

   Pada tahap ini akan menghasilkan *output* berupa rencana proyek dan penilaian awal untuk penggunaan alat dan teknik.

2. Data *Understanding*

   Pada fase pemahaman data dimulai dengan pengumpulan data awal dan memahami data yang akan digunakan dalam penelitian. Tahap pemahaman data dibagi menjadi beberapa langkah, antara lain :

   a. Pengumpulan Data Awal

      Pada langkah ini akan dikumpulkan data-data yang akan digunakan dalam data *mining.*

   b. Penjelasan Data

      Pada langkah ini akan dijelaskan data yang telah dikumpulkan di tahap awal. Dalam penjelasan data akan dijelaskan format dari data, kuantitas data, jumlah *record* dan *field* dalam setiap tabel, dsb.

   c. Eksplorasi Data

      Pada tahap ini akan dilakukan analisis data yang telah di jelaskan pada tahap sebelumnya. Salah satu cara mengeksplorasi data adalah menggunakan analisis statistik deskriptif [6]. Analisis statistik deskriptif berkaitan dengan mencari nilai-nilai di bawah ini :

      1) Nilai rata-rata dengan menggunakan persamaan (2.1) berikut ini :

$$\bar{\chi} = \frac{\sum_{i=1}^{n} x_i}{n}$$

2) Nilai standar deviasi dengan menggunakan persamaan (2.2) berikut ini :

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

3) Nilai kuartil 1, 2, dan 3 dengan menggunakan persamaan (2.3) berikut ini :

$$Letak\ Q_i = \frac{i(n+1)}{4}$$

4) Nilai *interquartile range* (IQR) dengan menggunakan persamaan (2.4) berikut ini :

$$IQR = Q_3 - Q_1$$

d. Verifikasi Kualitas Data

Tahap verifikasi kualitas data ini akan memeriksa data dengan cara melihat apakah terdapat *missing value* ataupun *outlier*. Metode yang dapat digunakan untuk mendeteksi *outlier* salah satunya yaitu dengan menggunakan *interquartile range* (IQR) [7], yaitu dengan menggunakan persamaan (2.5) berikut ini :

$$Batas\ bawah = Q_3 + (3 \times IQR)$$
$$Batas\ atas = Q_3 + (6 \times IQR)$$

3. Data *Preparation*

Tahap persiapan data mencakup semua kegiatan yang diperlukan untuk membangun data yang akan digunakan ke dalam pemodelan. Tahap persiapan data dibagi menjadi beberapa langkah, antara lain :

a. Pemilihan Data

Tahap pemilihan data bertujuan menentukan data yang akan digunakan. Pemilihan data ini meliputi pemilihan atribut ataupun pemilihan *record*.

b. Pembersihan Data

Tahap pembersihan data bertujuan untuk menghilangkan data berdasarkan data yang telah diversifikasi pada tahap verifikasi kualitas data.

c. Penyiapan Data Awal

Pada tahap ini akan disiapkan data yang telah dipilih dan dibersihkan untuk digunakan pada tahap pemodelan.

d. Integrasi Data

Pada tahap ini data akan digabungkan sesuai dengan kebutuhan untuk pemodelan.

e. Format Data

Pada tahap ini struktur data akan diformat disesuaikan dengan data yang dibutuhkan untuk pemodelan.

4. *Modelling*

Pada tahap ini akan dilakukan pemilihan dan penerapan model yang sesuai berdasarkan tujuan yang akan dicapai. Tahap pemodelan dibagi menjadi beberapa langkah, antara lain :

a. Pemilihan Teknik Pemodelan

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

b. Pembuatan Model

Pada tahap ini akan dijelaskan teknik pemodelan yang telah dipilih pada tahap pemilihan teknik pemodelan.

c. Pengujian Model

Pada tahap ini model yang telah dipilih akan diterapkan terhadap kasus uji.

5. *Evaluation*

Pada tahap ini akan dilakukan evaluasi terhadap model yang telah dibuat dalam tahap pemodelan, untuk memastikan model benar-benar telah mencapai tujuan bisnis. Pada tahap evaluasi terdapat beberapa langkah, antara lain :

a. Evaluasi Hasil

Pada tahap ini akan dievaluasi hasil dari pemodelan apakah sudah sesuai dengan tujuan bisnis atau tidak.

b. Peninjauan Proses

Pada tahap ini akan dilakukan peninjauan menyeluruh proses data *mining* yang telah dilakukan untuk memastikan adanya faktor penting atau tugas yang terabaikan.

c. Menentukan Langkah Selanjutnya

Pada tahap ini akan diputuskan langkah selanjutnya, apakah hasil pemodelan akan dilanjutkan ke tahap *deployment* atau tidak.

6. *Deployment*

Pembuatan model umumnya bukan akhir dari sebuah proyek, bahkan jika tujuan dari model ini adalah untuk meningkatkan pengetahuan dari data, maka harus disajikan dalam bentuk yang dapat digunakan oleh pengguna. Pada tahap *deployment* terdapat beberapa tahap yaitu :

a. Perencanaan *Deployement*

Pada tahap ini akan menghasilkan *output* berupa perencanaan yang akan dilakukan pada tahap *deployment*.

b. Pemantauan Rencana dan Pemeliharaan

Pada tahap ini rencana yang telah dibuat akan dipantau dan dipelihara.

c. Membuat Laporan Akhir

Pada tahap ini akan menghasilkan *output* berupa laporan akhir dan presentasi akhir.

d. *Review* Proyek

Pada tahap ini akan menghasilkan dokumentasi *review* dari proyek yang telah dilakukan.

2.2.6. Data *Preprocessing*

Dalam data *mining*, kualitas data yang akan digunakan perlu diperhatikan. Ada beberapa faktor yang mempengaruhi kualitas data, antara lain keakuratan, keutuhan, konsistensi, aktualitas, dan penafsiran. Data *preprocessing* dapat memperbaiki kualitas data, sehingga dapat meningkatkan keakuratan dan efisiensi hasil dari data *mining* [8]. Beberapa kegiatan data *preprocessing* yang akan dilakukan dalam penelitian ini adalah sebagai berikut :

a. Penanganan *Missing value*

Missing value pada dasarnya tidak bermasalah bagi keseluruhan data, apalagi jika jumlahnya hanya sedikit. Namun jika persentase data yang hilang tersebut

cukup besar, maka perlu dilakukan pengujian apakah data yang mengandung banyak *missing value* masih layak diproses lebih lanjut ataukah tidak. Salah satu cara menangani *missing value* adalah dengan menghitung nilai rata-rata dari objek yang nilainya hilang, nilai rata-rata tersebut akan dijadikan nilai pengganti dari nilai yang hilang. Selain itu *missing value* juga bisa ditangani dengan memberikan nilai *default* pada setiap nilai yang hilang.

b. Penanganan *Outlier*

Outlier adalah data yang berbeda secara signifikan dari data yang lainnya. Salah satu cara yang dapat dilakukan untuk membersihkan data *outlier* adalah dengan melakukan *smoothing*. Salah satu teknik *smoothing* yang dapat dilakukan adalah metode *binning*. Metode *binning* membagi kumpulan data ke dalam beberapa partisi atau bin. Dimulai dengan mengurutkan setiap nilai pada sebuah atribut. Kemudian data yang sudah diurutkan dibagi ke dalam beberapa partisi atau bin yang memiliki frekuensi yang sama (*equal-frequency partitioning*). Ada dua cara *smoothing* dalam metode *binning*, yaitu *smoothing by bin means* dan *smoothing by bin boundaries*.

Dalam *smoothing by bin means* dilakukan dengan mengubah setiap nilai dalam bin dengan nilai rata-rata dari bin tersebut. Sedangkan dalam *smoothing by bin boundaries* setiap nilai dalam bin diubah menjadi batas bawah (minimum) dan batas bawah (maksimum) pada setiap bin.

2.2.7. *Clustering*

Clustering adalah proses mengelompokkan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas atau *cluster* [9]. Tujuannya menemukan *cluster* yang berkualitas dalam waktu yang layak. *Clustering* dalam data *mining* berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisa data. Kesamaan objek biasanya diperoleh dari kedekatan nilai-nilai atribut yang menjelaskan objek-objek data, sedangkan objek-objek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi.

Dengan menggunakan *clustering*, dapat diidentifikasi daerah yang padat, pola-pola distribusi secara keseluruhan dan keterkaitan yang menarik antara

atribut-atribut data. Dalam data *mining* usaha difokuskan pada metode-metode penemuan untuk *cluster* pada basis data berukuran besar secara efektif dan efisien.

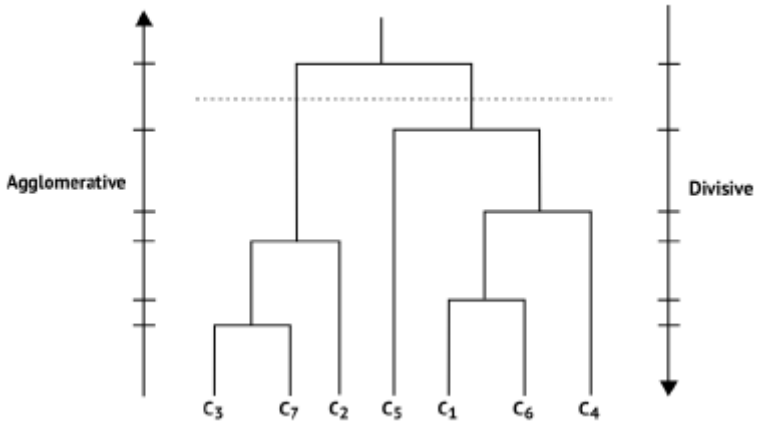
Secara garis besar ada beberapa kategori algoritma *clustering* yang dikenal yaitu:

a. Metode Partisi (*Partition-based Clustering*)

Dalam metode ini pengguna harus menentukan jumlah k partisi yang diinginkan lalu setiap data dites untuk dimasukkan pada salah satu partisi sehingga tidak ada data yang *overlap* dan satu data hanya memiliki satu *cluster*. Contohnya: algoritma K-Means, *mixture modeling*.

b. Metode Hierarki (*Hierarchical Clustering*)

Membuat hierarki berupa *dendogram* di mana data yang mirip akan disimpan di hierarki yang berdekatan. Metode ini terbagi menjadi dua yaitu *buttom-up* yang menggabungkan *cluster* kecil menjadi *cluster* lebih besar dan *top-down* yang memecah *cluster* besar menjadi *cluster* yang lebih kecil.
Contohnya: *Agglomerative Hierarchical Clustering* (AHC), *Divisive Hierarchical Clustering* (DHC).



Gambar 2.5 Dendrogram Hasil Pengelompokan Metode Hierarki

Dalam metode *clustering* hal yang penting adalah mengukur kemiripan atau ketidakmiripan data yang akan dikelompokkan. *Similaritas* (kemiripan) antara dua objek merupakan ukuran numerik dari seberapa mirip dua buah objek. *Similaritas* mempunyai nilai yang lebih tinggi jika objek-objek tersebut sama. *Range* nilai *similaritas* adalah [0,1]. *Dissimilaritas* (ketidakmiripan) antara dua objek merupakan ukuran numerik dari seberapa berbeda dua buah objek. *Dissimilaritas* bernilai lebih rendah jika objek-objek tersebut mirip. Minimum

dissimilaritas adalah 0 dan batas atasnya bervariasi. *Dissimiliratas* sering disebut juga dengan *distance* (jarak) [10]. Salah satu cara pengukuran jarak adalah menggunakan jarak *Euclidian* (*Euclidian Distance*), yang dinyatakan dengan persamaan (2.6) berikut :

$$dist = \sqrt{\sum_{k=1}^{n} (p_k - q_k)^2}$$

Di mana :

n adalah jumlah dimensi (*attribute*)

p_k dan q_k adalah atribut (komponen) ke-k atau objek data p dan q

Jika skala nilai dari objek-objek yang diukur berbeda maka diperlukan standarisasi.

2.2.8. *Agglomerative Hierarchical Clustering* (AHC)

Pengelompokan data menggunakan *agglomerative hierarchical clustering* (AHC) diawali dengan merepresentasikan setiap objek pada data sebagai satu kelompok, kemudian dilakukan perhitungan jarak (*distance measure*) antar kelompok tersebut. Setelah itu dua kelompok yang memiliki jarak terdekat digabungkan menjadi sebuah kelompok baru. Kemudian jarak antara kelompok yang baru dengan kelompok lain dihitung menggunakan salah satu metode perhitungan jarak antar kelompok (*single linkage*, *complete linkage*, *average linkage*, dll.) [9]. Proses perhitungan jarak dan penggabungan dua kelompok dilakukan secara iteratif. Berikut ini merupakan beberapa metode yang dapat digunakan untuk menghitung jarak antar kelompok :

a. *Single Linkage* (Jarak Terdekat)

   Single linkage disebut juga minimum *link*, di mana *similaritas* dari dua *cluster* didasarkan pada dua titik paling dekat dari dua *cluster* yang berbeda. Untuk mencari *single linkage* dapat menggunakan persamaan (2.7) berikut ini :

$$d_{uv} = \min\{d_{uv}\}, d_{uv} \in D$$

b. *Complete Linkage* (Jarak Terjauh)

   Complete linkage disebut juga *maximum link*, di mana jarak antara dua buah *cluster* dihitung dari jarak terjauh antara anggota *cluster* satu dengan

cluster kedua. Untuk mencari *complete linkage* dapat menggunakan persamaan (2.8) berikut ini :

$$d_{uv} = \max\{d_{uv}\}, d_{uv} \in D$$

c. *Average Linkage* (Jarak Rata-rata)

Average linkage disebut juga *group average*, di mana jarak antara dua buah *cluster* dihitung dari rata-rata jarak antara anggota *cluster* satu dengan *cluster* yang kedua. Untuk mencari *average linkage* dapat menggunakan persamaan (2.9) berikut ini :

$$d_{uv} = \text{average}\{d_{uv}\}, d_{uv} \in D$$

Berikut ini merupakan langkah-langkah dalam algoritma *Agglomerative Hierarchical Clustering* (AHC) [11]:

1. Setiap objek data dianggap sebagai *cluster*. Jika n = jumlah data dan k = jumlah *cluster*, maka k = n.
2. Menghitung matriks jarak antar *cluster* (menggunakan *euclidean distance*)
3. Cari dua *cluster* yang mempunyai jarak paling dekat (*single linkage*) lalu gabungkan dua *cluster* tersebut, sehingga k = k-1.
4. Jika k > 1, ulangi langkah 2 dan 3
5. Jika k = 1, perulangan selesai.
6. *Cluster* dihasilkan dari pemotongan dendrogram pada level yang tepat.

   Pemotongan dendrogram dapat dilakukan dengan berdasarkan pada agglomeration schedule. Jumlah cluster bisa didapatkan dengan mencari perbedaan koofisien jarak dari setiap iterasinya.

2.2.9. *Agglomeration Schedule*

Agglomeration schedule adalah ringkasan numerik dari solusi *cluster* yang dihasilkan dari proses *clustering* menggunakan metode hierarki [12]. Berikut ini merupakan contoh dari *agglomeration schedule* :

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	8	11	1.260	0	0	7
2	6	7	1.579	0	0	4
3	2	9	1.625	0	0	7
4	1	6	2.318	0	2	6
5	3	5	2.619	0	0	8
6	1	10	3.670	4	0	10
7	2	8	4.420	3	1	8
8	2	3	4.505	7	5	9
9	2	4	4.774	8	0	10
10	1	2	5.718	6	9	0

Gambar 2.6 Contoh Agglomeration Schedule

Penjelasan tabel *agglomeration schedule* :

1. *Stage* merupakan tahapan atau iterasi yang telah dilakukan pada proses *clustering* menggunakan metode hierarki.
2. *Cluster combined* merupakan *cluster* yang digabungkan pada setiap *stage*nya.
3. *Cooficients* merupakan jarak dari dua *cluster* yang digabungkan pada setiap *stagenya* yang dicari dengan menggunakan rumus jarak.
4. *Stage cluster first appear* merupakan *stage* apabila *cluster* yang digabungkan telah digabungkan dengan *cluster* lain pada *stage* sebelumnya. Jika *cluster* yang digabungkan belum pernah digabungkan pada *stage* sebelumnya, maka akan diisi dengan nilai 0.
5. *Next stage* merupakan *stage* selanjutnya saat *cluster* akan digabungkan dengan *cluster* lain.

Berikut ini contoh penjelasan tahapan pada *agglomeration schedule* :

Pada tahap pertama, *cluster* yang digabungkan adalah *cluster* 8 dan *cluster* 11, karena memiliki jarak yang paling kecil.

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	8	11	1.260	0	0	7
2	6	7	1.579	0	0	4
3	2	9	1.625	0	0	7
4	1	6	2.318	0	2	6
5	3	5	2.619	0	0	8
6	1	10	3.670	4	0	10
7	2	8	4.420	3	1	8
8	2	3	4.505	7	5	9
9	2	4	4.774	8	0	10
10	1	2	5.718	6	9	0

Gambar 2.7 Agglomeration Schedule Tahap Pertama

Cluster yang telah digabungkan pada *stage* pertama akan digabungkan kembali pada *stage* ke-7.

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	8	11	1.260	0	0	7
2	6	7	1.579	0	0	4
3	2	9	1.625	0	0	7
4	1	6	2.318	0	2	6
5	3	5	2.619	0	0	8
6	1	10	3.670	4	0	10
7	2	8	4.420	3	1	8
8	2	3	4.505	7	5	9
9	2	4	4.774	8	0	10
10	1	2	5.718	6	9	0

Gambar 2.8 Agglomeration Schedule Next Stage

Pada *stage* 7, *cluster* yang telah digabungkan pada *stage* 1 dan *stage* 3 akan digabungkan kembali. *Cluster* yang akan dihasilkan berikutnya akan muncul pada *stage* 8.

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	8	11	1.260	0	0	7
2	6	7	1.579	0	0	4
3	2	9	1.625	0	0	7
4	1	6	2.318	0	2	6
5	3	5	2.619	0	0	8
6	1	10	3.670	4	0	10
7	2	8	4.420	3	1	8
8	2	3	4.505	7	5	9
9	2	4	4.774	8	0	10
10	1	2	5.718	6	9	0

Gambar 2.9 Agglomeration Schedule Stage 7

Ketika *cluster* yang akan digabungkan berjumlah sangat banyak, maka tabel *agglomeration schedule* yang dihasilkan akan bertambah panjang, tetapi akan mempermudah untuk melihat jarak terbesar dari seluruh *cluster* dibandingkan dengan melihat dari dendrogram. Jumlah *cluster* terbaik adalah selisih jarak terbesar.

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	8	11	1.260	0	0	7
2	6	7	1.579	0	0	4
3	2	9	1.625	0	0	7
4	1	6	2.318	0	2	6
5	3	5	2.619	0	0	8
6	1	10	3.670	4	0	10
7	2	8	4.420	3	1	8
8	2	3	4.505	7	5	9
9	2	4	4.774	8	0	10
10	1	2	5.718	6	9	0

Gambar 2.10 Selisih Jarak Pada Agglomeration Schedule

Selisih jarak terbesar terdapat pada antara *stage* 5 dan *stage* 6, yang menunjukkan solusi jumlah *cluster* terbaik adalah sebanyak 6 *cluster*. Hal yang sama dapat dilihat juga dari dendrogram yang dihasilkan dari proses *clustering* menggunakan metode *hierarki.*

2.2.10. Unified Modeling Language (UML)

Unified modeling language adalah bahasa pemodelan untuk sistem atau perangkat lunak yang berorientasi objek [13]. Beberapa model yang akan

digunakan untuk menggambarkan sistem yang akan dibangun dalam penelitian ini adalah :

1. *Use Case Diagram*

   Use case diagram digunakan untuk memodelkan fungsionalitas-fungsionalitas sistem atau perangkat lunak dilihat dari pengguna yang ada di luar sistem (yang sering dinamakan sebagai *actor*). Kegunaan dari *use case* diagram adalah untuk mendaftarkan aktor-aktor dan *use case* serta memperlihatkan aktor mana yang berpartisipasi dalam masing-masing *use case*. Aktor merupakan representasi dari orang yang berada di luar sistem dan akan berinteraksi dengan sistem. *Use case* adalah fungsionalitas sistem atau perangkat lunak.

2. *Class Diagram*

   Class diagram merupakan diagram yang menggambarkan hubungan antara *class* yang satu dengan *class* yang lain. *Class* didefinisikan dengan tiga area pokok, yaitu nama *class*, atribut, dan operasi atau *method*. Pada *class*, atribut dan *method* akan dilakukan *encapsulation*, jenis-jenis *encapsulation* adalah *public*, *private*, dan *protected*.

3. *Activity Diagram*

   Activity diagram bertujuan memodelkan komputasi-komputasi dan aliran-aliran kerja yang terjadi dalam sistem atau perangkat lunak. Suatu diagram aktivitas memuat didalamnya *activity state* di mana suatu *activity state* merepresentasikan eksekusi pernyataan dalam suatu prosedur atau kinerja suatu aktivitas dalam satu aliran kerja.

4. *Sequence Diagram*

   Sequence diagram digambarkan secara *vertical* dan horizontal. Pada bagian *vertical* digambarkan sumbu waktu (*lifeline*), waktu bertambah dari atas ke bawah. Pada bagian horizontal memperlihatkan objek-objek yang terlibat. Selama objek ada, peran digambarkan menggunakan garis tegas. Selama aktivitas prosedur pada objek aktif, garis waktu digambarkan dengan garis ganda. Pesan-pesan digambarkan sebagai suatu tanda panah dari garis waktu suatu objek ke garis waktu objek lainnya.