

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Algoritma**

Menurut Munir (2012:176) algoritma adalah urutan logis langkah-langkah penyelesaian masalah yang disusun secara sistematis. Metode pengurutan digambarkan dalam sejumlah langkah terbatas yang mengarah pada solusi permasalahan.

#### **2.2 Sistem Operasi**

Ariyus dan Pangera (2010 : 57) menguraikan bahwa Sistem Operasi merupakan sebuah penghubung antara pengguna mesin dengan perangkat keras yang dimiliki mesin tersebut. Sebelum ada sistem operasi, orang hanya menggunakan komputer dengan menggunakan sinyal analog dan digital. Seiring dengan berkembangnya pengetahuan dan teknologi, pada saat ini terdapat berbagai sistem operasi dengan keunggulan masing-masing. Sistem operasi bertindak sebagai antarmuka antara program aplikasi dengan perangkat keras komputer, level dari pengguna setiap lapisan juga berbeda-beda. Program aplikasi hanya digunakan oleh pemakai terakhir (*End User*), sedangkan sistem operasi dan perangkat keras merupakan tugas pemrogram dan pendesain sistem operasi tersebut.

#### **2.3 K-means**

Menurut Vuldari (2017:54) k-means merupakan algoritma yang menetapkan nilai-nilai cluster ( $k$ ) secara random, untuk sementara nilai tersebut menjadi pusat dari cluster yang biasa disebut centroid. Kemudian menghitung jarak setiap data yang ada terhadap masing-masing centroid menggunakan rumus Euclidean hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid hingga nilai centroid tidak berubah (stabil).

Menurut Suyanto (2017:262) k-means merupakan algoritma klusterisasi yang memiliki ide dasar sederhana dengan cara meminimalkan Sum of Squared Error (SSE) antara objek-objek data dengan sejumlah  $k$  centroid.

Berikut ini adalah langkah-langkah algoritma k-means :

- a) Penentuan cluster awal

Dalam menentukan buah pusat cluster awal dilakukan pembangkitan bilangan random yang merepresentasikan urutan data input. Pusat awal cluster didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan random pusat awal dari data.

b) Perhitungan jarak dengan pusat cluster

Untuk mengukur jarak antar data dengan pusat dengan cluster digunakan euclidian distance, algoritma perhitungan jarak data dengan pusat cluster :

1. Pilih nilai data dan nilai pusat cluster.
2. Hitung euclidian distance data dengan tiap pusat cluster
3.  $d(x_i, \mu_j) = \sqrt{(x_i - \mu_j)^2} \dots (1)$
4. Penjelasan :

$x_i$  : Data kriteria

$\mu_j$  : Centroid pada cluster ke j

c) Pengelompokan data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat cluster jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster terdekat. Adapun cara pengelompokan data tersebut adalah :

1. Pilih nilai jarak tiap pusat cluster dengan data.
2. Cari nilai jarak terkecil.
3. kelompokkan data dengan pusat cluster yang memiliki jarak terkecil.

d) Penentuan pusat cluster baru

Untuk mendapatkan pusat cluster baru bisa dihitung dari rata-rata nilai anggota cluster dan pusat cluster. Pusat cluster yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh user atau hasil yang dicapai sudah konvergen (pusat cluster baru sama dengan pusat cluster lama).

1. Algoritma penentuan pusat cluster :

Cari jumlah anggota tiap cluster

2. Hitung pusat baru dengan rumus

$$\mu_j^{(t+1)} = \frac{1}{N_j} \sum_{x_j \in S_j} x_j \dots (2)$$

Penjelasan :

$\mu_j^{(t+1)}$  : Centroid baru pada iterasi ke 1

$N_{sj}$  : Banyak data pada cluster  $s_j$

Hasil dari operasi *clustering* yang terbentuk selanjutnya akan di evaluasi menggunakan Davies bouldin index yang dihitung dengan persamaan :

$$DBI = \frac{1}{K} \sum_{i=1}^k R_i$$

$$R_i = \max_{j=1 \dots k, i \neq j} R_{ij}$$

$$R_{ij} = \frac{\text{var}(C_i) + \text{var}(C_j)}{\|c_i - c_j\|}$$

Dimana:

$C_i$  = Cluster  $i$  dan cluster  $c_i$  adalah centroid dari cluster  $i$

## 2.4 Clustering

Menurut Suyanto (2017:260) Clustering adalah proses pengelompokan satu set objek data (into multiple groups) atau cluster sehingga benda-benda dalam suatu kelompok memiliki kesamaan yang tinggi, tetapi sangat berbeda dengan objek di kelompok lain.

Menurut Han,dkk (2012:445) clustering adalah proses mempartisi sekumpulan objek data (pengamatan) kedalam himpunan bagian yang dapat digunakan untuk mengatur hasil pencarian ke dalam kelompok dan menyajikan hasil dengan cara yang ringkas dan mudah diakses. *Clustering* banyak digunakan dalam berbagai bidang dengan beragam aplikasi yang sangat penting diantaranya riset pasar, sistem rekomendasi, sistem keamanan dan mesin pencarian.

## 2.5 Data Mining

### 2.5.1 Pengertian Data Mining

Perkembangan teknologi informasi telah memberikan kontribusi pada cepatnya pertumbuhan jumlah data yang dikumpulkan dan disimpan dalam basis data berukuran besar (big data). Big data adalah istilah yang menggambarkan volume data yang besar, baik data yang terstruktur maupun data yang tidak terstruktur. Big data memiliki potensi tinggi untuk mengumpulkan wawasan kunci dari informasi bisnis. Big data dapat dianalisis untuk wawasan yang mengarah pada pengambilan keputusan dan strategi bisnis yang lebih baik. Sebuah metode atau teknik diperlukan untuk dapat merubah data tersebut menjadi sebuah informasi berharga atau pengetahuan yang bermanfaat untuk mendukung pengambilan keputusan.

Menurut Stanton (2013:173) data mining adalah bidang penelitian dan praktik yang berfokus pada penemuan pola-pola baru dalam data yang mengacu pada penggunaan algoritma dan komputer untuk menemukan pola baru dan menarik dalam data.

Menurut Suyatno (2017:2) data mining ditujukan untuk mengekstrak pengetahuan dari sekumpulan data sehingga didapatkan struktur yang dapat dimengerti manusia serta meliputi basis data dan manajemen data, prapemrosesan data, pertimbangan model dan inferensi, ukuran ketertarikan, pertimbangan kompleksitas, pascapemrosesan terhadap struktur yang ditemukan, visualisasi dan online updating. Sebagai teknologi umum, data mining dapat diterapkan ke semua jenis data selama data bermakna untuk aplikasi target. Bentuk data paling dasar untuk penambangan aplikasi adalah database, data warehouse dan data transaksional. Data mining juga dapat diterapkan ke bentuk data lain (misalnya, aliran data, data urutan, grafik atau data jaringan, data spesial, data teks, data multimedia).

### 2.5.2 Fungsi Data Mining

Secara umum, kegunaan data mining terbagi menjadi dua yaitu deskriptif dan prediktif. Deskriptif memiliki arti untuk mencari pola-pola yang dapat dipahami manusia yang menjelaskan karakteristik data sedangkan prediktif digunakan untuk membentuk sebuah model pengetahuan guna melakukan prediksi. Berdasarkan fungsionalitasnya, tugas-tugas data mining bisa dikelompokkan menjadi enam kelompok yaitu :



**Gambar 2.1** Fungsi Data Mining

Adapun penjelasan rinci dari enam kelompok tersebut sebagai berikut:

a) Klasifikasi (classification)

Proses generalisasi struktur yang diketahui untuk diaplikasikan pada data-data baru.

b) Klasterisasi (clustering)

Mengelompokan data yang belum diketahui label kelasnya ke dalam sejumlah kelompok tertentu sesuai dengan ukuran kemiripannya.

c) Regresi (regression)

Menemukan suatu fungsi yang memodelkan data dengan kesalahan prediksi seminimal mungkin.

d) Deteksi anomali (anomaly detection)

Mengidentifikasi data yang tidak umum, berupa outlier (pencilan), perubahan atau deviasi yang mungkin sangat penting dan perlu investigasi lebih lanjut.

e) Pemodelan kebergantungan (Dependency modeling)

Mencari relasi antar tabel.

f) Perangkuman (summarization)

Menyediakan representasi data yang lebih sederhana, meliputi visualisasi dan pembuatan laporan.

## 2.6 Perangkat Lunak Yang Digunakan

### 2.6.1 HTML (*Hyper Text Markup Language*)

Diar Puji Oktavian (2010 : 13) menguraikan HTML (*Hyper Text Markup Language*) adalah suatu bahasa yang dikenali web browser untuk menampilkan informasi dengan lebih menarik dibandingkan dengan tulisan teks biasa”. Sedangkan web browser adalah program komputer yang digunakan untuk membaca HTML, kemudian menerjemahkan dan dan menampilkan hasilnya secara visual ke layar komputer. Contoh program web browser antara lain seperti Mozilla Firefox, Google Chrome, Internet Explore, Microsoft Edge, dan sebagainya.

### 2.6.2 MySQL

Eko dan Angga (2019 : 5) MYSQL merupakan database engine atau server database yang mendukung bahasa database pencarian SQL. MYSQL adalah sebuah perangkat lunak sistem manajemen basis data SQL atau DBMS yang *Multithread* dan *Multiuser*. MSQl AB membuat MYSQL tersedia sebagai perangkat lunak gratis dibawah lisensi GNU *Generated Public License* (GPL). Mereka juga menjual dibawah lisensi komersial untuk kasus-kasus dimana penggunaanya tidak cocok dengan pengguna GPL.

### 2.6.3 Hypertext Preprocessor (PHP)

Menurut Hikmah, dkk (2015:1) "PHP merupakan kependekan dari Hypertext Preprocessor. PHP tergolong sebagai perangkat lunak open source yang diatur dalam aturan general purpose licences (GPL). Bahasa pemrograman PHP sangat cocok dikembangkan dalam lingkungan web, karena PHP bisa diletakkan pada script HTML atau sebaliknya.

Menurut Sibero (2013:49) "PHP adalah pemrograman interpreter yaitu proses penerjemahan baris kode sumber menjadi kode mesin yang dimenegerti komputer secara langsung pada saat baris kode dijalankan".

## **2.7 Metode Perangkat Lunak Yang Digunakan**

### **2.7.1 Metode Prototype**

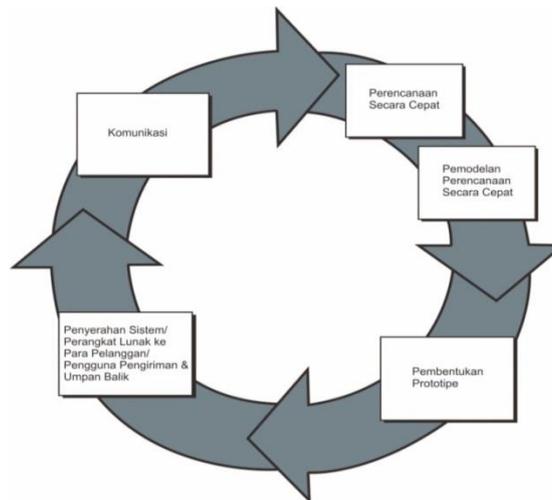
Metode pengembang perangkat lunak yang digunakan pada aplikasi ini adalah metode *prototype*. Dalam melakukan perancangan system yang akan dikembangkan dapat menggunakan metode *prototype*. Metode ini cocok digunakan untuk mengembangkan sebuah perangkat yang akan dikembangkan kembali. Kemudian membuat sebuah rancangan kilat yang selanjutnya akan dievaluasi kembali sebelum diproduksi secara benar.

(Pressman 2012) *Prototype* bukanlah merupakan sesuatu yang lengkap, tetapi sesuatu yang harus dievaluasi dan dimodifikasi kembali. Segala perubahan dapat terjadi pada saat *prototype* dibuat untuk memenuhi kebutuhan pengguna dan pada saat yang sama memungkinkan pengembang untuk lebih memahami kebutuhan pengguna secara lebih baik.

Berikut adalah tahapan dalam metode *prototype*:

- a) Komunikasi dan pengumpulan data awal, yaitu wawancara terhadap pihak yang terkait dalam penelitian dan analisis terhadap kebutuhan pengguna  
( wawancara terhadap owner atau karyawan salon)
- b) Perencanaan Secara Cepat yaitu pembuatan desain secara umum untuk selanjutnya dikembangkan kembali.
- c) Pembentukan *Prototype* yaitu pembuatan perangkat prototype termasuk pengujian dan penyempurnaan.
- d) Evaluasi terhadap *Prototype*, yaitu mengevaluasi prototype dan memperhalus analisis terhadap kebutuhan pengguna.
- e) Perbaikan *Prototype* , yaitu pembuatan tipe yang sebenarnya berdasarkan hasil dari evaluasi prototype.

- f) Penyerahan sistem perangkat lunak ke para pelanggan pengguna pengiriman dan umpan balik



**Gambar 2.2** Metode *Prototype* (Pressman, 2012)

## 2.8 Pengujian *Black-box Testing*

Metode pengujian pada aplikasi ini menggunakan *Black Box Testing* yaitu yang berfokus pada spesifikasi fungsional dari perangkat lunak. Tester dapat mendefinisikan kumpulan kondisi input dan melakukan pengetesan pada spesifikasi fungsional program.

*Black Box Testing* bukanlah solusi alternatif dari *White Box Testing* tapi lebih merupakan pelengkap untuk menguji hal-hal yang tidak dicakup oleh *White Box Testing* (Pressman 2012).

*Black Box Testing* cenderung untuk menemukan hal-hal berikut:

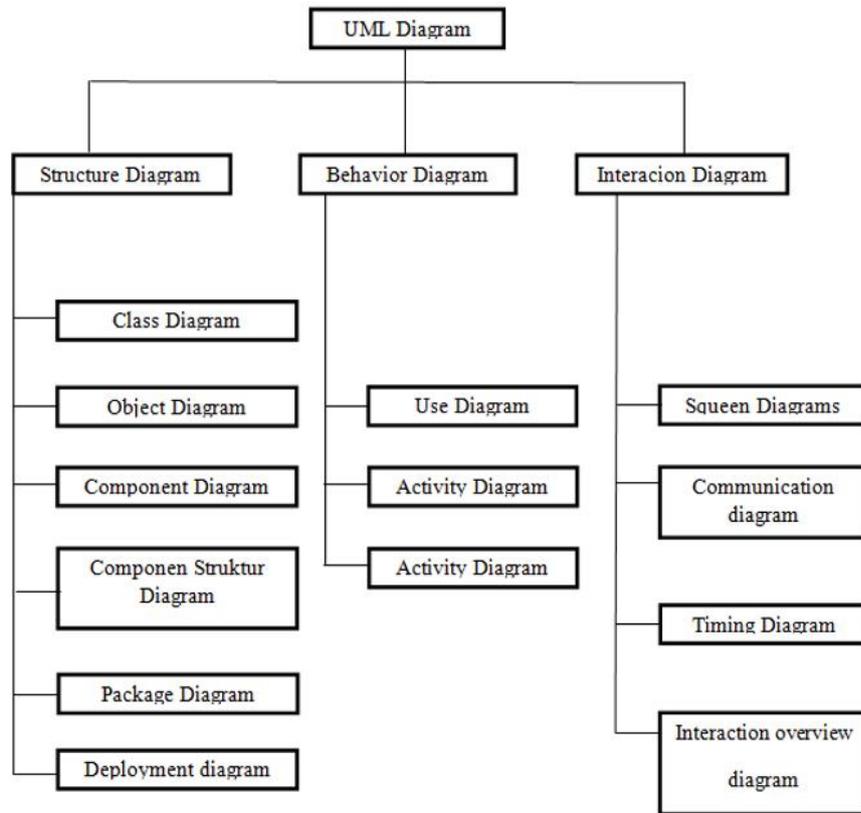
- Fungsi yang tidak benar atau tidak ada.
- Kesalahan antarmuka (*interface errors*).
- Kesalahan pada struktur data dan akses basis data.

## 2.9 Sistem Pemodelan

### 2.9.1 *Unified Modeling Language (UML)*

*Unified Modeling Language (UML)* adalah sebuah bahasa yang berdasarkan gambar untuk memvisualisasikan, menspesifikasikan, membangun dan pendokumentasian dari sebuah sistem pengembangan perangkat lunak berbasis Objek. *Unified Modeling Language (UML)* bukanlah merupakan bahasa pemrograman tetapi model-model yang tercipta berhubungan langsung dengan berbagai macam bahasa pemrograman berorientasi obyek, seperti Java (Syafitri 2016). UML tersusun atas sejumlah elemen grafis membentuk diagram-diagram. Dalam

penelitian ini melakukan desain hanya 2 diagram yaitu *Use Case Diagram* dan *Activity Diagram*.



**Gambar 2.3** Diagram UML

Berikut penjelasan singkat dari pembagian kategori tersebut :

- a) *Structure* diagram yaitu sekumpulan diagram yang digunakan untuk menggambarkan suatu struktur statis dari sistem yang dimodelkan.
- b) *Behavior* diagram yaitu kumpulan diagram yang digunakan untuk menggambarkan kelakuan sistem atau rangkaian perubahan yang terjadi pada sebuah sistem.
- c) *Interaction* diagram yaitu diagram yang digunakan untuk menggambarkan interaksi sistem dengan sistem ini maupun interaksi antar subsistem pada suatu sistem.

### 2.9.2 Use Case Diagram

Munawar (2018 : 89) menguraikan bahwa *Use Case* adalah deskripsi fungsi sebuah system dari perspektif pengguna. *Use Case* bekerja dengan cara mendeskripsikan tipikal interaksi antara pengguna sebuah system dengan sistemnya sendiri melalui sebuah cerita bagaimana sebuah

system dipakai. Urutan langkah-langkah yang menerangkan antara pengguna dan system disebut sebagai *Scenario*. Setiap *Scenario* menggambarkan urutan kejadian. Setiap urutan di inialisasi oleh orang, system yang lain, perangkat keras atau urutan waktu. Dengan demikian, secara singkat bias dikatakan *Use Case* adalah serangkaian *Scenario* yang digabungkan bersaa-sama oleh tujuan umum pengguna. *Use Case* dibuat berdasarkan kebutuhan Aktor. *Use Case* harus merupakan 'apa' yang dikerjakan software aplikasi, bukan 'bagaimana' software aplikasi mengerjakannya. Tabel 2.1 pada halaman berikut ini adalah Simbol-simbol yang digunakan dalam *Use Case Diagram* :

**Tabel 2.1** Simbol *Use Case Diagram*.

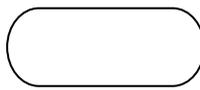
GAMBAR	NAMA	KETERANGAN
	<i>Actor</i>	Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan <i>use case</i> .
	<i>Note</i>	Elemen fisik yang eksis saat aplikasi dijalankan dan mencerminkan suatu sumber daya komputasi
	<i>Collaboration</i>	Interaksi aturan-aturan dan elemen lain yang bekerja sama untuk menyediakan perilaku yang lebih besar dari jumlah dan elemen-elemennya (sinergi).
	<i>Dependency</i>	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri ( <i>independent</i> ) akan mempengaruhi elemen yang bergantung padanya elemen yang tidak mandiri ( <i>independent</i> ).
	<i>Generalization</i>	Hubungan dimana objek anak ( <i>descendent</i> ) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk ( <i>ancestor</i> ).
	<i>Include</i>	Menspesifikasikan bahwa <i>use case</i> sumber secara <i>eksplisit</i> .

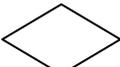
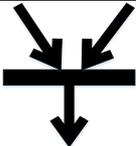
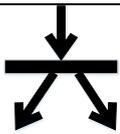
	<i>Extend</i>	Menspesifikasikan bahwa <i>use case</i> target memperluas perilaku dari <i>use case</i> sumber pada suatu titik yang diberikan.
	<i>Association</i>	Apa yang menghubungkan antara objek satu dengan objek lainnya.
	<i>Use Case</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu aktor

### 2.9.3 Activity Diagram

Munawar (2018 : 127) menguraikan bahwa *Activity Diagram* adalah bagian penting dari UML yang menggambarkan aspek dinamis dari Sistem. Logika Prosedural, proses bisnis dan aliran kerja suatu bisnis bisa dengan mudah di deskripsikan dalam *Activity Diagram*. *Activity Diagram* mempunyai peran seperti halnya *Flowchart*, akan tetapi perbedaannya dengan *Flowchart* adalah `; *Activity Diagram* bisa mendukung perilaku paralel sedangkan *Flowchart* tidak bisa. Simbol-simbol yang digunakan dalam *Activity diagram* dapat dilihat pada tabel 2.2 dibawah ini :

**Tabel 2.2** Simbol *Activity Diagram*.

Simbol	Keterangan
	<i>Activity</i> : Memperlihatkan bagaimana masing-masing kelas antarmuka saling berinteraksi satu sama lain.
	<i>Initial Node</i> : Bagaimana objek dibentuk atau diawali
	<i>Activity Final Node</i> : Bagaimana objek dibentuk dan diakhiri.

	<i>Decision</i> : Asosiasi percabangan dimana jika ada pilihan aktifitas lebih dari satu.
	<i>Swimlane</i> : Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktifitas yang terjadi.
	<i>Join</i> : Digunakan untuk menunjukkan kegiatan yang digabungkan.
	<i>Fork</i> : Digunakan untuk menunjukkan kegiatan yang dilakukan secara paralel

## 2.10 Penelitian Terkait

Dalam penyusunan skripsi ini, penulis terinspirasi dan mereferensi dari penelitian-penelitian sebelumnya yang berkaitan dengan skripsi ini. Daftar penelitian terkait sebagai berikut :

**Tabel 2.3** Penelitian Terkait

No.	Nama	Judul	Tahun	Uraian
1.	Bagas Prayogo Haryaji.	Penerapan Algoritma K-Means Untuk Memetakan Garis Kemiskinan Menurut Provinsi Di Indonesia	2018	Penerapan algoritma k-means membagi dataset menjadi tiga kelompok yaitu sangat rendah, rendah dan cukup sesuai dengan kemiripan tingkat pendapatan/ kapita/ bulan.
2.	Baginda Harahap.	Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus	2019	Berdasarkan hasil yang diperoleh dari penelitian ini dengan metode K-Means clustering.

		Pada UD. Toko Bangunan YD Indarung)		Hasilnya mendapatkan 3 kategori penjualan bahan bangunan, yaitu kurang laris, sangat laris dan laris.
3.	Ade Bastian, Harun Sujadi, Dan Gigin Febrianto	Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)	2018	Dari hasil penelitian yang telah dilakukan dapat disimpulkan bahwa K-means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok. Kelebihan Algoritma K-means diantaranya adalah mampu mengelompokkan objek besar dan pencilan obyek dengan sangat cepat sehingga mempercepat

				proses pengelompokan.
--	--	--	--	--------------------------

