

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Objek dan Subjek Penelitian**

Objek dalam penelitian ini merujuk pada penerapan *Data-Driven Marketing* dalam menganalisis *market-driven strategy* dengan memanfaatkan *big data analytic tools* melalui proses klusterisasi, evaluasi akurasi model mengklasifikasi dan memetakan karakteristik klaster untuk optimalisasi strategi pemasaran di Soulja Coffee. Penelitian ini menitikberatkan pada bagaimana strategi pemasaran berbasis data digunakan untuk mengelompokkan konsumen berdasarkan pola perilaku aktual guna meningkatkan efektivitas pemasaran. Sementara itu, subjek penelitian ini adalah konsumen Soulja Coffee, dengan data perilaku pembelian mereka yang akan dianalisis menggunakan teknik *K-Means Clustering* untuk mendapatkan segmentasi pasar yang lebih optimal dan relevan. Hasil segmentasi tersebut kemudian diproses menggunakan algoritma *Decision Tree* untuk mengidentifikasi akurasi model klusterisasi dan memetakan karakteristik klaster.

##### **3.1.1 Lokasi Penelitian**

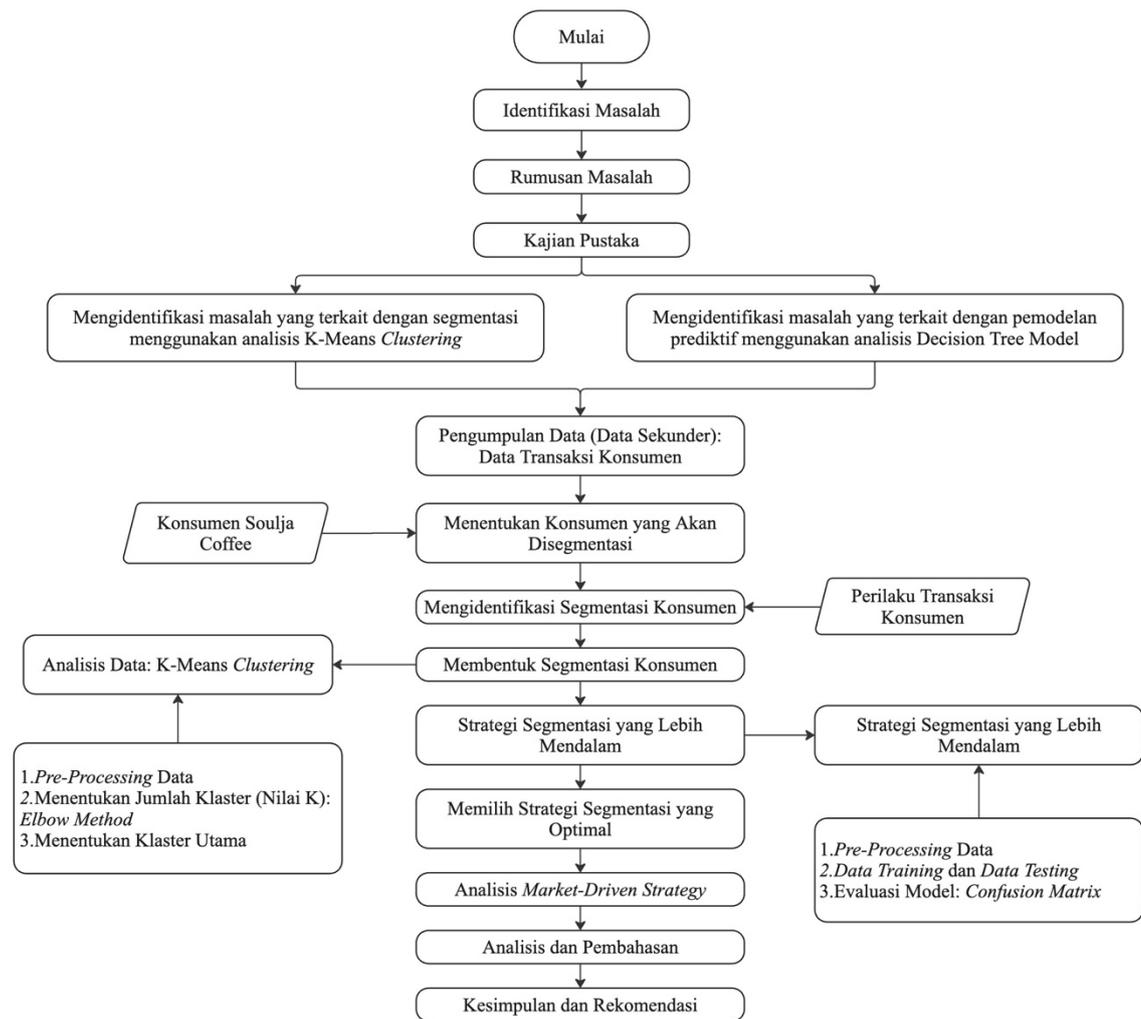
Lokasi penelitian adalah tempat di mana penelitian dilakukan serta menjadi lokasi utama bagi peneliti untuk mengungkap kondisi sebenarnya dari subjek yang diteliti. Lokasi ini berperan sebagai sumber utama bagi peneliti dalam memperoleh data yang diperlukan guna mendukung analisis dalam penelitian. Dalam penelitian ini, peneliti memilih lokasi di Soulja Coffee, yang beralamat di Jalan Galunggung No.21, Tawang, Kota. Tasikmalaya, Jawa Barat.

#### **3.2 Metode Penelitian**

##### **3.2.1 Jenis Penelitian dan Metode Yang Digunakan**

Penelitian ini menggunakan pendekatan kuantitatif yang menekankan pada aspek objektivitas dengan pengolahan data numerik yang akurat, lengkap, dan sesuai yang diperoleh menggunakan teknik statistik tertentu (Bougie & Sekaran, 2019). Penerapan Metode penelitian kuantitatif ini memanfaatkan *machine learning* yang mampu untuk mempelajari pola dari data. *Machine learning* memiliki beberapa kategori dua diantaranya terdapat *unsupervised learning* (*K-*

*Means Clustering*) dan *supervised learning (Decision Tree)*. Dari kedua jenis algoritma *machine learning* tersebut dapat menggali pola tersembunyi dalam perilaku konsumen dan mengidentifikasi segmentasi konsumen dengan potensi tinggi dalam meningkatkan retensi serta profitabilitas bisnis. Adapun rancangan penelitian yang terlampir pada gambar 3.1.



Gambar 3. 1 Rancangan Penelitian

Rancangan penelitian ini dirancang menggunakan pendekatan kuantitatif dengan metode eksploratif dan deskriptif. Proses dimulai mengidentifikasi masalah, pengumpulan data transaksi melalui sistem POS, dilanjutkan dengan tahap Exploratory Data Analysis (EDA). Data yang telah diproses kemudian digunakan dalam dua tahap analisis utama yaitu segmentasi pelanggan menggunakan

Melani Defina Maharani, 2025

**MARKET-DRIVEN STRATEGY DENGAN PEMANFAATAN BIG DATA ANALYTIC TOOLS: KLASTERISASI UNTUK OPTIMALISASI STRATEGI PEMASARAN DI SOULJA COFFEE**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

algoritma *K-Means* serta evaluasi model dan memetakan karakteristik perilaku pelanggan menggunakan *Decision Tree*. Hasil akhir dari analisis ini digunakan untuk merancang strategi pemasaran berbasis data.

### 3.2.2 Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder, yaitu data yang telah tersedia dan diperoleh dari pihak lain, bukan dikumpulkan langsung oleh peneliti (Bougie & Sekaran, 2019). Data sekunder dalam penelitian ini berupa riwayat transaksi pelanggan Soulja Coffee selama periode Desember 2024 hingga Maret 2025.

Data tersebut diperoleh dari sistem *Customer Relationship Management* (CRM) operasional berbasis *platform* Majoo, yang berfungsi mencatat dan mengelola aktivitas transaksi pelanggan secara otomatis melalui sistem *Point of Sale* (POS). Sistem ini mencatat berbagai aspek penting dalam perilaku konsumen, seperti waktu pembelian, jenis produk, metode pembayaran, dan jenis pemesanan. Informasi tersebut digunakan sebagai dasar dalam menyusun segmentasi pelanggan dan perancangan strategi pemasaran berbasis data.

### 3.2.3 Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini menggunakan data sekunder yang diperoleh dari sistem *Point of Sale* (POS) Soulja Coffee yang diekspor dari sistem POS dalam format *Excel*. Data yang diperoleh pada tabel 3.1 mencakup berbagai variabel atribut transaksi yang dijelaskan sebagai berikut.

Tabel 3. 1 Atribut Dataset Transaksi Konsumen

No.	Nama Atribut	Deskripsi Atribut
1.	No Transaksi	Kode unik untuk setiap transaksi
2.	Outlet	Lokasi tempat transaksi dilakukan.
3.	Produk	Nama produk yang dibeli dalam transaksi tersebut.
4.	Jenis Order	Kategori jenis pesanan seperti <i>outlet</i> (ditempat) atau melalui layanan <i>ojek online</i> (GoFood, GrabFood, Shopeefood)
5.	Sisa Tagihan	Jumlah sisa tagihan yang belum dibayar untuk transaksi tersebut.
6.	Total Penjualan	Total nilai moneter dari transaksi (penjualan) yang dilakukan.

Melani Defina Maharani, 2025

*MARKET-DRIVEN STRATEGY DENGAN PEMANFAATAN BIG DATA ANALYTIC TOOLS: KLASTERISASI UNTUK OPTIMALISASI STRATEGI PEMASARAN DI SOULJA COFFEE*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

No.	Nama Atribut	Deskripsi Atribut
7.	Metode Pembayaran	Jenis pembayaran (Cash, QRIS, dll)
8.	Tanggal Order	Tanggal ketika pesanan dibuat atau dicatat dalam sistem.
9.	Jam Order	Jam ketika pesanan dibuat atau dicatat dalam sistem.
10.	Tanggal Bayar	Tanggal ketika pelanggan membayar pesanan
11.	Jam Bayar	Jam ketika pelanggan membayar pesanan

Data pada tabel tersebut dikumpulkan mencakup informasi transaksi pelanggan. Kemudian setelah data diekspor, dilakukan proses verifikasi untuk memastikan kelengkapan data, dilanjutkan dengan seleksi variabel yang relevan dan penyesuaian manual sederhana seperti standarisasi penamaan. Data akhir yang telah dibersihkan dan disesuaikan digunakan untuk membentuk dataset analisis perilaku konsumen.

### 3.3 Instrumen Penelitian

Instrumen penelitian yang digunakan dalam penelitian ini mencakup sejumlah alat bantu teknis dan metode analisis yang dirancang untuk mengolah dan menginterpretasikan data transaksi pelanggan secara sistematis. Mengingat pendekatan penelitian ini bersifat kuantitatif berbasis data sekunder, maka instrumen yang digunakan berupa *big data analytic tools* yang mendukung proses pengolahan secara menyeluruh. Sistem *Point of Sale* (POS) yang digunakan di Soulja Coffee menjadi sumber utama data, yang menghasilkan kumpulan informasi transaksi dalam jumlah besar dan terus bertambah seiring waktu yang disebut sebagai *Big Data*.

Data tersebut akan dikelola menggunakan Microsoft Excel untuk keperluan validasi struktur dan seleksi variabel yang relevan. Seluruh proses analisis data dalam penelitian ini dilakukan melalui pendekatan *data mining*, yaitu rangkaian proses eksplorasi data besar untuk menemukan pola tersembunyi yang bernilai dalam mendukung pengambilan keputusan. Tahapan *data mining* yang dilakukan meliputi *data preprocessing*, segmentasi pelanggan menggunakan algoritma *K-*

*Means Clustering*, dan akurasi model *cluster* menggunakan *Decision Tree* yang dijalankan melalui Google Colab berbasis Python.

Instrumen-instrumen tersebut sebagai *big data analytic tools* memungkinkan peneliti melakukan klasifikasi dan segmentasi pelanggan dengan lebih sistematis dan objektif. Adapun tahapan teknis dari penggunaan instrumen tersebut yang akan dijelaskan secara lebih rinci sebagai berikut.

### 3.3.1 *Exploratory Data Analysis (EDA)*

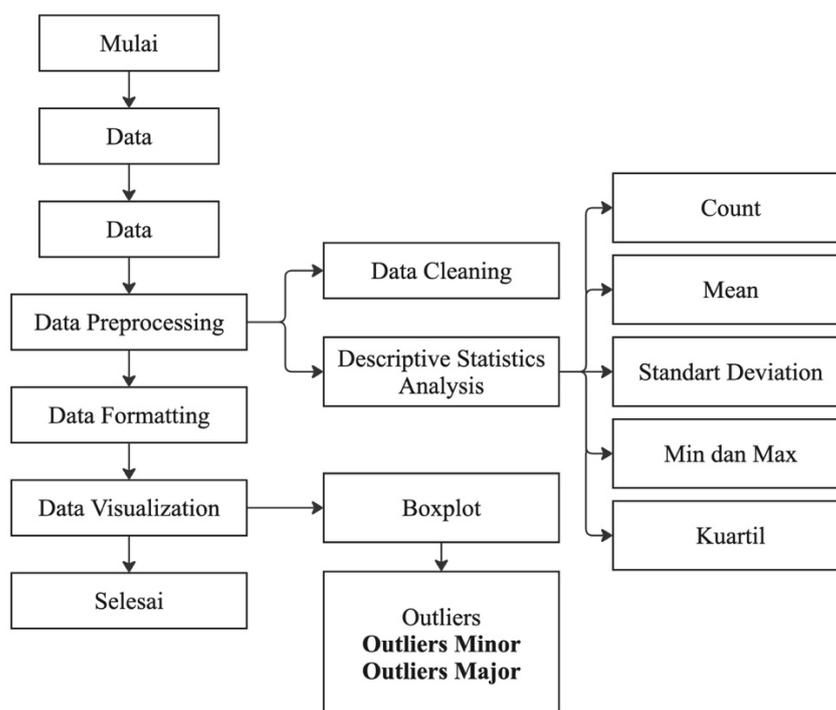
*Exploratory Data Analysis (EDA)* merupakan tahapan penting dalam proses analisis data yang dilakukan sebelum menerapkan metode statistik lanjutan maupun algoritma *machine learning*. Tahap ini bertujuan untuk memahami struktur, karakteristik, dan distribusi data secara menyeluruh, serta membantu peneliti mengidentifikasi pola, tren, hubungan antarvariabel, maupun anomali seperti *outlier*. Dengan menggali informasi tersembunyi dari data, EDA mempermudah pemilihan teknik analisis yang tepat dan menjadi landasan awal dalam pengembangan model yang akurat dan relevan. Dalam konteks penelitian berbasis data, EDA tidak hanya berperan sebagai tahap pendahuluan, tetapi juga sebagai proses eksploratif yang mendukung arah analisis secara strategis dan ilmiah (Maresti dkk., 2024).

Sejalan dengan penjelasan sebelumnya, tahapan perancangan proses analisis dalam penelitian ini mengacu pada metode *Exploratory Data Analysis (EDA)*. EDA digunakan untuk menggali informasi awal dari data sebelum diterapkan ke dalam algoritma analitik. Menurut Angela, Islamiyah, dan Irsyad (2023), adapun tahapan yang dilakukan dalam proses EDA ini meliputi:

1. *Data Understanding*, setelah dilakukan penyusunan dataset dengan format excel yaitu dilakukan *data understanding*, penelaahan terhadap atribut-atribut dalam dataset, seperti jenis data (kategorikal atau numerik), jumlah entri, distribusi awal data, serta keterkaitan antar variabel, untuk menentukan variabel mana yang relevan digunakan dalam proses analisis.
2. *Data Preprocessing*, pada *data preprocessing* dilakukan pembersihan data (*data cleaning*), *Data cleaning* merupakan tahapan untuk mendeteksi dan

memperbaiki, atau menghapus data yang tidak konsisten, salah, duplikat, berformat tidak sesuai, atau mengandung kesalahan lain dalam dataset yang dapat mengganggu jalannya analisis dan menghasilkan output yang kurang akurat. Proses ini meliputi tindakan seperti penggabungan atau penyederhanaan, mengganti, menghapus nilai kosong (*missing value*), atau memperbaiki data yang dianggap tidak relevan atau bermasalah.

3. *Formatting* dan Perbaikan Tipe Data, melakukan koreksi tipe data agar sesuai dengan kebutuhan analisis (misalnya: mengubah kolom tanggal ke format datetime, memastikan kolom numerik dalam format float/integer, dan encoding data kategorikal).
4. *Data Visualitation*, menyusun visualisasi eksploratif terhadap variabel-variabel yang ditentukan.



Gambar 3. 2 Alur *Exploratory Data Analysis* (EDA).

### 3.3.2 *K-Means Clustering*

Algoritma ini adalah untuk menemukan sejumlah kelompok atau kluster dalam data, di mana jumlah kluster ditentukan berdasarkan nilai  $k$  yang ditetapkan sebelumnya. Proses kerja algoritma ini bersifat iteratif, di mana setiap data akan

Melani Defina Maharani, 2025

*MARKET-DRIVEN STRATEGY DENGAN PEMANFAATAN BIG DATA ANALYTIC TOOLS: KLASTERISASI UNTUK OPTIMALISASI STRATEGI PEMASARAN DI SOULJA COFFEE*

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

dikelompokkan ke dalam salah satu dari k kluster berdasarkan kemiripan karakteristik atau fitur yang dimilikinya. Setiap objek dimasukkan ke dalam kluster dengan pusat (*centroid*) yang paling dekat. Algoritma ini berupaya untuk membentuk k kluster yang memiliki perbedaan maksimum antar kelompok dan kesamaan maksimum dalam kelompok. Pemilihan nilai k yang optimal tidak ditentukan secara langsung, melainkan perlu dianalisis terlebih dahulu dari data, salah satunya menggunakan metode *Elbow*. Tujuan akhir dari *K-Means* adalah untuk meminimalkan total variasi dalam kluster (*intra-cluster variance*), dengan mengukur jarak antara titik data dan pusat kluster yang menaunginya (Sarker et al., 2018). Jika diberikan sekumpulan data (misalnya  $x_1, x_2, \dots, x_n$ ), algoritma akan membagi n data tersebut ke dalam k kluster (dengan  $k \leq n$ ) secara optimal, dengan meminimalkan fungsi objektif tertentu, yaitu total kesalahan kuadrat antara data dan *centroid*-nya.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|(x_i^j - c_j)\|^2$$

Di mana J adalah fungsi objek, k jumlah

kluster, n jumlah kasus, dan  $\|x_j^i - c_j\|^2$  adalah ukuran jarak yang dipilih antara titik data  $x_j^i$  dan pusat kluster  $c_j$ . Menurut Sarker dkk., (2018) Tahapan pelaksanaan algoritma *K-Means* dijelaskan sebagai berikut:

1. Menentukan Jumlah Kluster (K)

Menetapkan jumlah kluster yang akan digunakan berdasarkan tujuan analisis atau hasil evaluasi visual seperti metode *Elbow*.

2. Menentukan k Centroid

- a. Memilih k titik awal secara acak sebagai centroid awal.
- b. Mengukur kualitas *clustering* dengan menghitung *sum of squared distances* (SSD) dari setiap titik data ke centroidnya.

3. Pengelompokan data berdasarkan kedekatan dengan centroid.

Menghitung jarak antara setiap titik data ke masing-masing pusat kluster menggunakan *Euclidean Distance*. *Euclidean Distance* adalah metode pengukuran jarak terpendek antara dua titik dalam ruang multidimensi. Semakin kecil nilai *Euclidean Distance*, semakin dekat titik

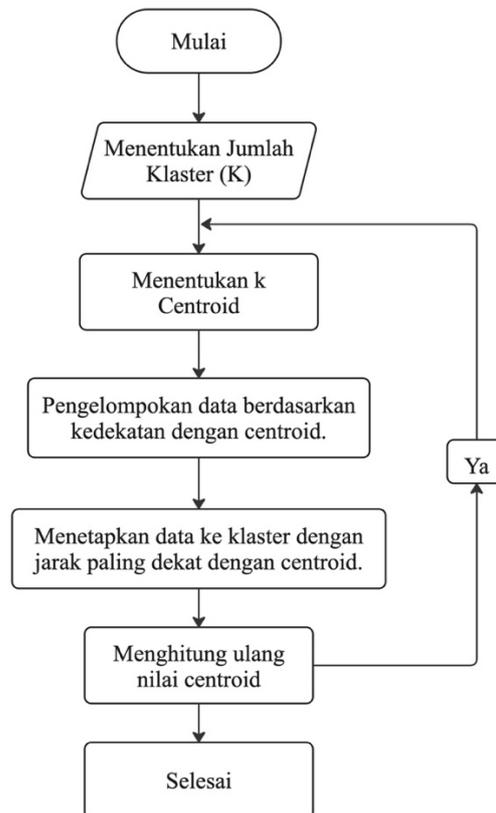
tersebut dengan *centroid*, sehingga lebih mungkin menjadi bagian dari kluster tersebut.

4. Menetapkan data ke kluster dengan jarak paling dekat dengan centroid.
5. Menghitung ulang pusat kluster

Menghitung ulang nilai centroid untuk setiap kluster berdasarkan rata-rata dari seluruh anggota kluster.

6. Iterasi Sampai Konvergen

Proses pengelompokan dan perhitungan ulang terus diulang hingga posisi pusat kluster stabil (tidak berubah signifikan), atau iterasi maksimum tercapai.



Gambar 3. 3 *Flowchart K-Means Clustering*

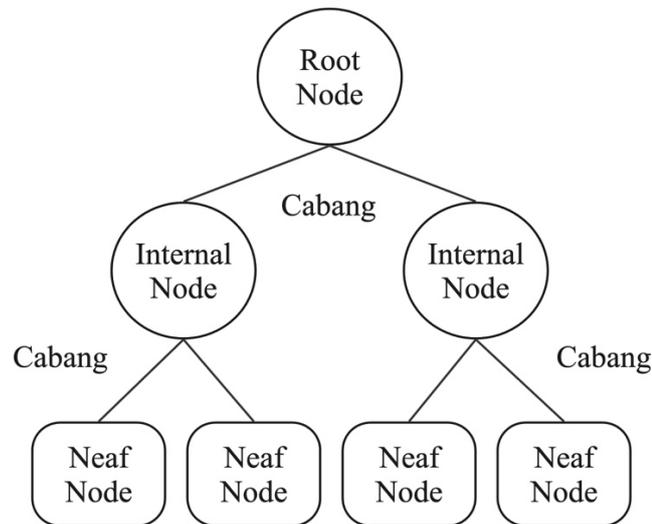
### 3.3.3 Decision Tree Model

Alur pada *Decision Tree* di telusuri dari simpul akar ke simpul daun yang membentuk struktur. Diagram alir yang berbentuk seperti struktur pohon yang mana setiap *internal node* menyatakan pengujian terhadap suatu atribut, setiap

cabang menyatakan output dari pegujian tersebut dan *leaf node* menyatakan kelas-kelas/ distribusi kelas. *Node* yang paling atas disebut sebagai *root node* atau *node* akar, memiliki beberapa *edge* keluar tetapi tidak memiliki *edge* masuk. *Internal node* akan memiliki satu *edge* masuk dan beberapa *edge* keluar, sedangkan *leaf node* hanya akan memiliki satu *edge* masuk tanpa memiliki *edge* keluar. *Leaf node* adalah hasil akhir yang mewakili label kelas dari kombinasi atribut yang terbentuk menjadi *rule* (Kasih, 2019)

Dalam membangun sebuah *Decision Tree* secara *top-down* (dari atas ke bawah), tahap awal yang dilakukan adalah mengevaluasi semua atribut yang ada menggunakan suatu ukuran statistik (yang biasa digunakan adalah *information gain*) untuk mengukur efektifitas suatu atribut dalam mengklasifikasikan suatu kumpulan sampel data. Atribut yang diletakkan pada *root node* adalah atribut yang memiliki *information gain* terbesar. Semua atribut adalah bersifat kategori yang bernilai diskrit. Atribut dengan nilai *continuous* harus didiskritkan (Kasih, 2019)

Nilai *Gain* adalah *Information Gain* yang digunakan untuk mencari satu *variable/* atribut dari dataset (S) untuk dijadikan *root/ node* dan *branch node*, yaitu satu atribut yang mempunyai nilai *gain* tertinggi. Untuk pencarian nilai *gain (information Gain)* dapat dilakukan dengan konsep *Entropy*, *Gini Index* dan *Clasification Error*. Nilai *gain* terbesar yang di dapat dari atribut pada dataset (data pembelajaran) pertama kali digunakan untuk mencari atribut yang layak menjadi *root* (akar) pohon keputusan (*decision tree*). Selanjutnya proses akan diulang untuk mencari atribut yang menjadi cabang hingga menemukan *leaf* yang merupakan label kelas (Kasih, 2019).



Gambar 3. 4 Contoh bentuk *Decision Tree*

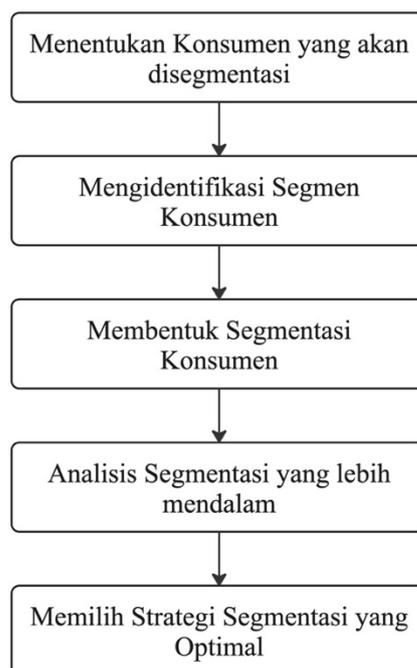
Layaknya sebuah pohon, *Decision Tree* mempunyai akar (*root/ node*), batang/cabang (*branch node*), dan daun (*leaf*). Menurut Kasih (2019), Strategi yang dapat digunakan untuk pembentukan pohon keputusan dengan *Decision Tree* adalah:

- a. Pohon dimulai sebagai *node* tunggal (akar/ *root*) yang merepresentasikan semua data.
- b. Sesudah *node root* dibentuk, maka data pada *node* akar akan diukur dengan *information gain* untuk dipilih atribut mana yang akan dijadikan atribut pembaginya.
- c. Sebuah cabang dibentuk dari atribut yang dipilih menjadi pembagi dan data akan didistribusikan ke cabang masing-masing.
- d. Algoritma ini akan terus menggunakan proses yang sama (rekursif) untuk dapat membentuk sebuah *Decision Tree* . Ketika sebuah atribut telah dipilih menjadi *node* pembagi/ cabang, maka atribut tersebut tidak diikuti lagi dalam penghitungan nilai *information gain*.
- e. Proses pembagian rekursif akan berhenti jika salah satu dari kondisi dibawah ini terpenuhi:
  1. Semua data dari anak cabang telah termasuk dalam kelas yang sama.

2. Semua atribut telah dipakai, tetapi masih tersisa data dalam kelas yang berbeda. Dalam kasus ini, diambil data yang mewakili kelas yang terbanyak untuk menjadi label kelas pada *node* daun.
3. Tidak terdapat data pada anak cabang yang baru. Dalam kasus ini, *node* daun akan dipilih pada cabang sebelumnya dan diambil data yang mewakili kelas terbanyak untuk dijadikan label kelas.

### 3.4 Rancangan Analisis Data

Metode analisis data dalam penelitian ini menggunakan adaptasi pendekatan segmentasi konsumen yang telah dikembangkan oleh Cravens dan Piercy. Penelitian ini berbasis data transaksi yang diterapkan pada Soulja Coffee. Berikut adalah langkah-langkah dalam mengolah data penelitian berdasarkan metode yang digunakan (Cravens & Piercy, 2013)



Gambar 3. 5 Fase Analisis Data

#### 1. Menentukan Konsumen yang Akan Disegmentasi

Langkah pertama dalam segmentasi konsumen adalah mengidentifikasi variasi kebutuhan dan preferensi konsumen berdasarkan transaksi mereka di Soulja Coffee. Variabel yang digunakan dalam penelitian ini mencakup pola

pembelian dan konsumsi. Dalam penelitian ini, peneliti menggunakan riwayat transaksi pembelian konsumen sebagai sumber data utama untuk segmentasi.

## 2. Mengidentifikasi Segmentasi Konsumen

Setelah konsumen yang akan disegmentasi ditentukan, variabel perilaku transaksi digunakan untuk mengidentifikasi segmentasi konsumen potensial. Variabel ini mencakup frekuensi pembelian, jenis menu yang sering dibeli, serta jumlah pengeluaran per transaksi, dan lain-lain. Analisis ini membantu dalam mengelompokkan konsumen yang relevan dengan strategi pemasaran Soulja Coffee.

## 3. Membentuk Segmentasi Konsumen

Untuk membentuk segmentasi konsumen, penelitian ini menggunakan analisis berbasis *K-Means Clustering*. Metode ini memungkinkan pengelompokan konsumen berdasarkan kesamaan pola transaksi mereka. Dalam analisis kluster ini, variabel yang digunakan mencerminkan perilaku konsumsi konsumen.

## 4. Analisis Segmentasi yang Lebih Mendalam

Variabel perilaku transaksi seperti jam bayar, jenis produk, metode pembayaran, dan hari transaksi digunakan sebagai fitur masukan dalam pengembangan model klasifikasi berbasis *Decision Tree*. Model ini bertujuan untuk mengidentifikasi karakteristik khas dari masing-masing kluster hasil *K-Means*, yang nantinya dapat digunakan sebagai dasar dalam optimalisasi strategi pemasaran. Untuk mengevaluasi sejauh mana model dapat memprediksi kluster pelanggan secara akurat, digunakan pendekatan *confusion matrix* dan metrik evaluasi klasifikasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*.

Mengingat model klasifikasi yang digunakan bersifat multi-kelas, evaluasi model tidak cukup hanya menggunakan *accuracy*, tetapi juga mempertimbangkan nilai *macro average* dan *weighted average* untuk memahami kinerja model terhadap semua kelas secara adil. Menurut Han, Kamber, dan Pei (2011), *accuracy* paling efektif ketika distribusi kelas relatif

seimbang, sedangkan metrik seperti sensitivitas (*recall*), *precision*, dan *F-score* lebih tepat untuk data dengan distribusi kelas tidak merata.

Tabel 3. 2 *Confusion Matrix*

Actual \ Predicted	Cluster 0	Cluster 1	Cluster 2	...	Cluster n	Total
Cluster 0	TP (C0)	FP (C1 on C0)	FP (C2 on C0)	...	FP (Cn on C0)	P <sub>0</sub>
Cluster 1	FN (C0 from C1)	TP (C1)	FP (C2 on C1)	...	FP (Cn on C1)	P <sub>1</sub>
Cluster 2	FN (C0 from C2)	FN (C1 from C2)	TP (C2)	...	FP (Cn on C2)	P <sub>2</sub>
...	...	...	...	...	...	...
Cluster n	FN (C0 from Cn)	FN (C1 from Cn)	FN (C2 from Cn)	...	TP (Cn)	P <sub>n</sub>

Han, Kamber, dan Pei (2011) juga menyebutkan beberapa metrik evaluasi kinerja model klasifikasi, yaitu:

- Accuracy:** Proporsi dari total prediksi yang benar terhadap jumlah total sampel. Semakin mendekati 1 model semakin sempurna (*The Perfect Classifier*). **Formula:**  $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
- Precision (per kelas):** Ketepatan prediksi terhadap suatu kelas; berapa banyak yang diprediksi benar dari semua yang diprediksi sebagai kelas tersebut. Jika Precision = 1 maka semua prediksi kelas itu benar (*no false positives*). **Formula:**  $Precision = \frac{TP}{TP+FP}$
- Recall / Sensitivity (per kelas):** Kemampuan model dalam menemukan semua data dari suatu kelas. Jika Recall = 1 maka semua anggota kelas berhasil terdeteksi (*no false negatives*). **Formula:**  $Recall = \frac{TP}{TP+FN}$
- F1-Score (per kelas):** Rata-rata harmonik dari precision dan recall, mencerminkan keseimbangan antara keduanya. Jika F1 = 1 maka keseimbangan sempurna antara *precision* dan *recall*. **Formula:**  $F1 = \frac{2 \times Precision \times Recall}{Precision+Recall}$
- Macro Average:** Rata-rata dari *precision*, *recall*, dan *f1-score* dari seluruh kelas, tanpa memperhitungkan jumlah data tiap kelas.

- f. **Weighted Average:** Rata-rata dari metrik di semua kelas yang diberi bobot sesuai jumlah data per kelas.

## 5. Memilih Strategi Segmentasi yang Optimal

Analisis konsumen digunakan untuk memilih segmentasi yang memiliki karakteristik serupa dengan konsumen yang memiliki kecenderungan tinggi untuk merespons strategi pemasaran. Tujuan dari analisis ini adalah menemukan karakteristik deskriptif dari konsumen berdasarkan riwayat transaksi yang memiliki korelasi tinggi dengan variabel penting yang digunakan dalam pembentukan segmentasi dalam model. Dari segmentasi yang optimal, strategi pemasaran dapat disesuaikan dengan kebutuhan dan pola transaksi konsumen, sehingga meningkatkan peluang keberhasilan strategi pemasaran yang diterapkan.

Dalam merancang strategi pemasaran yang optimal dan relevan dengan karakteristik tiap segmen. Pendekatan yang digunakan dalam penelitian ini berfokus pada *market-driven strategy*, yaitu strategi yang disusun berdasarkan pemahaman menyeluruh terhadap kebutuhan, perilaku, dan preferensi pelanggan di masing-masing klaster. Dalam mengimplementasikan *market-driven strategy*, salah satu pendekatan yang efektif adalah dengan menerapkan strategi *positioning*.

Menurut Kotler dan Keller (2016), Penerapan strategi *positioning* dalam penelitian ini mengacu pada kerangka *marketing mix* 4P, yaitu:

1. *Product* (Produk): Menyesuaikan jenis produk unggulan yang ditawarkan sesuai dengan selera dominan pada masing-masing klaster.
2. *Price* (Harga): Mengatur strategi harga berdasarkan sensitivitas harga tiap segmen.
3. *Place* (Distribusi): Mengoptimalkan saluran distribusi dan metode pemesanan yang sering digunakan oleh masing-masing klaster.
4. *Promotion* (Promosi): Merancang komunikasi pemasaran dan promosi sesuai dengan waktu transaksi dan perilaku belanja tiap klaster.