



## ***Anhesa Gym Customer Segmentation Based on RFM, CLV, and K-Means for Segmentation, Targeting, and Positioning Strategies***

### **Segmentasi Pelanggan Anhesa Gym Berbasis RFM, CLV, dan Algoritma K-Means untuk Strategi Segmentation, Targeting, and Positioning**

**Ricky Mangihut Rajagukguk<sup>1\*</sup>, Budi Prasetyo<sup>2</sup>**

<sup>1,2</sup>Program Studi Sistem Informasi, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Negeri Semarang, Indonesia

E-Mail: <sup>1</sup>kyy@students.unnes.ac.id, <sup>2</sup>bprasetyo@mail.unnes.ac.id

Received May 14th 2026; Revised Jun 05th 2025; Accepted Jun 12th 2026; Available Online Jun 25th 2026

Corresponding Author: Ricky Mangihut Rajagukguk

Copyright © 2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

#### **Abstract**

*Anhesa Gym faces challenges retaining customer loyalty and optimizing its marketing strategy due to a lack of in-depth understanding of each customer segment's characteristics. This study aims to design a targeted Segmentation, Targeting, and Positioning (STP) marketing strategy. Unlike previous studies that generally rely on traditional calculations, this study contributes by integrating the Recency, Frequency, Monetary (RFM) model and the K-Means Clustering algorithm, validated using the Analytic Hierarchy Process (AHP) weighting, to calculate Customer Lifetime Value (CLV). This combination is applied to minimize managerial-assumption bias and to provide objective weights for each behavioral variable. The analysis results indicate the formation of three optimal clusters, with a Davies-Bouldin Index (DBI) of 0.436. Through AHP, the Monetary variable was identified as the most influential (0.650). Cluster 2 was designated as the Loyal segment (CLV score: 1.000), followed by Cluster 0 as Potential (0.327), and Cluster 1 as Lost/Churn (0.217). Based on these results, an STP strategy was formulated comprising exclusive premium service offerings for the Loyal segment and a price reactivation program for the Lost/Churn segment. This study practically facilitates management in making data-driven marketing decisions*

**Keywords:** CLV, Customer Segmentation, K-Means algorithm, RFM, STP strategy

#### **Abstrak**

Anhesa Gym menghadapi tantangan dalam mempertahankan loyalitas pelanggan dan mengoptimalkan strategi pemasaran karena kurangnya pemahaman mendalam terhadap karakteristik tiap segmen pelanggan. Penelitian ini bertujuan untuk merancang strategi pemasaran *Segmentation, Targeting, dan Positioning* (STP) yang tepat sasaran. Berbeda dari penelitian terdahulu yang umumnya mengandalkan perhitungan tradisional, penelitian ini memberikan kontribusi melalui integrasi model *Recency, Frequency, Monetary* (RFM) dan algoritma *K-Means Clustering* yang divalidasi dengan pembobotan *Analytic Hierarchy Process* (AHP) dalam menghitung *Customer Lifetime Value* (CLV). Kombinasi ini diterapkan untuk meminimalisir bias asumsi manajerial dan memberikan bobot yang objektif pada tiap variabel perilaku. Hasil analisis menunjukkan terbentuknya tiga klaster optimal dengan nilai *Davies-Bouldin Index* (DBI) 0,436. Melalui AHP, variabel *Monetary* diidentifikasi paling berpengaruh (0,650). Klaster 2 ditetapkan sebagai segmen *Loyal* (skor CLV: 1,000), diikuti Klaster 0 sebagai *Potential* (0,327), dan Klaster 1 sebagai *Lost/Churn* (0,217). Berdasarkan hasil tersebut, dirumuskan strategi STP berupa penawaran layanan premium eksklusif bagi segmen *Loyal* serta program reaktivasi harga bagi segmen *Lost/Churn*. Penelitian ini secara praktis memfasilitasi pihak manajemen dalam mengambil keputusan pemasaran berbasis data.

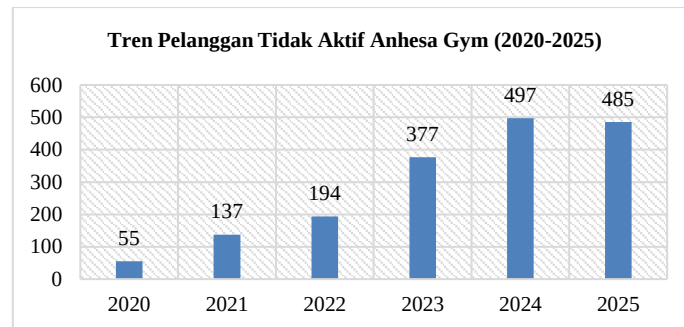
**Kata Kunci:** Algoritma K-Means, CLV, RFM, Segmentasi Pelanggan, Strategi STP

## **1 PENDAHULUAN**

Pertumbuhan industri kebugaran (*fitness industry*) menunjukkan tren yang positif pascapandemi COVID-19, sehingga menciptakan iklim kompetisi yang sangat ketat antar penyedia layanan atau *gym* [1].



Berdasarkan laporan riset pasar global, industri pusat kebugaran diproyeksikan terus berekspansi dengan tingkat pertumbuhan tahunan majemuk atau *Compound Annual Growth Rate* (CAGR) mencapai 9,3%, yang secara langsung memicu bertambahnya jumlah penyedia layanan dan menciptakan persaingan pangsa pasar yang sangat ketat [2]. Di tengah iklim kompetisi ini, tantangan utama bisnis telah bergeser dari sekadar akuisisi anggota baru menjadi upaya retensi pelanggan, di mana tingginya tingkat *churn* (berhentinya keanggotaan) menjadi ancaman serius bagi profitabilitas jangka panjang. Fenomena penurunan retensi ini juga terjadi secara nyata pada operasional Anhesa Gym. Berdasarkan data historis kunjungan terakhir (*last visit*) pelanggan selama periode 2022 hingga awal 2025 yang ditunjukkan pada Gambar 1, terlihat adanya tren peningkatan jumlah pelanggan yang berhenti aktif setiap tahunnya.



**Gambar 1.** Tren Pelanggan Tidak Aktif Anhesa Gym (2020-2025)

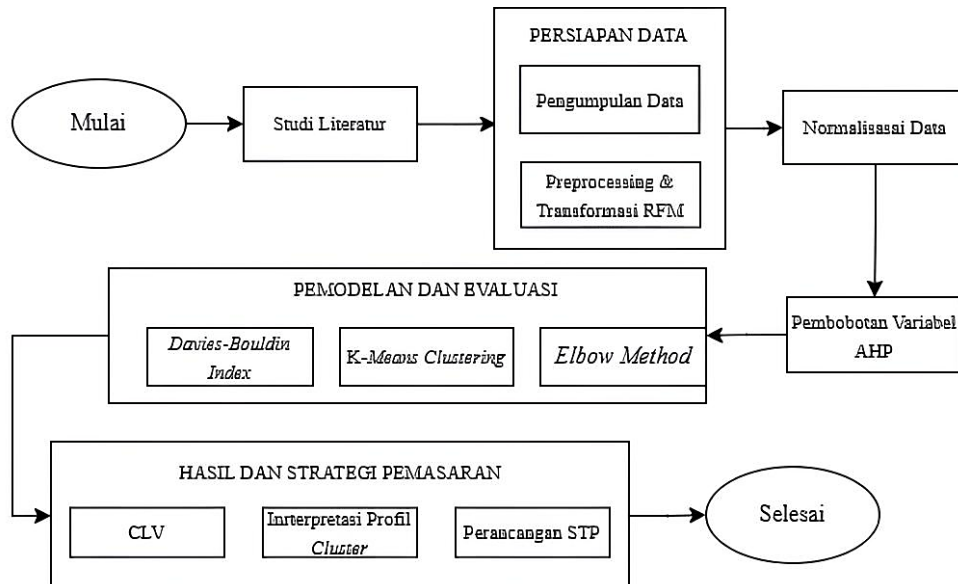
Tingginya angka *drop-out* pada Gambar 1 menjadi bukti empiris bahwa pendekatan pemasaran "satu strategi untuk semua" (*mass marketing*) sudah tidak lagi efektif dalam mempertahankan loyalitas dan menjawab heterogenitas kebutuhan anggota [3]. Untuk merancang strategi retensi yang presisi, diperlukan sebuah alur analitik terpadu yang mampu menerjemahkan data riwayat transaksi menjadi keputusan manajerial. Tahap analitik awal dilakukan dengan mengekstraksi perilaku pelanggan menggunakan model *Recency, Frequency, Monetary* (RFM) [4]. Namun, variabel RFM murni hanyalah metrik perilaku dasar; agar memiliki signifikansi bisnis, metrik tersebut harus dikonversi menjadi *Customer Lifetime Value* (CLV) guna menakar estimasi nilai ekonomi tiap individu. Setelah valuasi kelayakan tiap pelanggan diketahui, tahap selanjutnya adalah mengelompokkan mereka secara matematis menggunakan algoritma *K-Means Clustering*. Pada tahap akhir, hasil pengelompokan (*clustering*) ini tidak dibiarkan mengambang sebagai analisis teoretis, melainkan diterjemahkan secara langsung menjadi langkah operasional yang nyata melalui kerangka *Segmentation, Targeting, dan Positioning* (STP) [5], [6]. Melalui alur ini, konsep RFM, CLV, dan STP terintegrasi menjadi satu kesatuan solusi pemecahan masalah yang berkelanjutan.

Berbagai penelitian terdahulu telah berupaya memecahkan persoalan segmentasi pelanggan menggunakan analitik data. Mayoritas literatur, seperti penelitian [7] dan [8], membuktikan bahwa model RFM dan algoritma K-Means efektif untuk pemetaan profil pelanggan. Di sektor pusat kebugaran secara spesifik, pendekatan *clustering* juga telah diimplementasikan oleh [9], sementara [10] serta [11] kembali menegaskan efektivitas teknis gabungan RFM dan K-Means. Namun, kajian kritis terhadap literatur tersebut mengungkap dua keterbatasan utama. Pertama, penelitian tersebut umumnya berhenti pada tahap pembentukan kluster matematis tanpa mengonversinya menjadi taktik manajerial yang konkret. Padahal, [5] serta [6] menekankan bahwa penyusunan strategi operasional seperti STP sangat krusial agar hasil segmentasi dapat dieksekusi oleh manajemen. Kedua, pengelompokan berbasis RFM konvensional sering kali mengasumsikan ketiga variabel (R, F, dan M) memiliki bobot kepentingan yang sama. Pendekatan ini mengabaikan fakta bahwa nilai ekonomi CLV sangat dipengaruhi oleh prioritas bisnis yang berbeda-beda, terutama dalam industri jasa kebugaran.

Untuk menjembatani kesenjangan tersebut, penelitian ini mengusulkan sebuah kerangka kerja terintegrasi yang menjadi kontribusi dan kebaruan utamanya. Perbedaan mendasar penelitian ini dibandingkan riset sebelumnya terletak pada integrasi metode *Analytic Hierarchy Process* (AHP) untuk menyuntikkan bobot prioritas berbasis penilaian pakar secara objektif ke dalam setiap variabel RFM, guna mengeliminasi bias asumsi sebelum perhitungan CLV dieksekusi. Penelitian ini secara teoretis menyajikan evaluasi algoritma pengelompokan berbasis K-Means yang divalidasi oleh pembobotan pakar ahli, dan secara praktis berhasil menjembatani luaran data sains tersebut menjadi rancangan strategi pemasaran STP yang aplikatif bagi pihak manajemen. Berdasarkan landasan tersebut, tujuan penelitian ini adalah menghitung nilai RFM dan CLV pelanggan menggunakan bobot AHP, menerapkan algoritma K-Means untuk segmentasi, mengevaluasi kualitas kluster, serta merumuskan rekomendasi pemasaran yang spesifik untuk Anhesa Gym.

## 2 METODOLOGI PENELITIAN

Dalam penelitian ini, metode *Data Mining* digunakan untuk melakukan segmentasi pelanggan dengan memanfaatkan model *Recency, Frequency, Monetary* (RFM) dan algoritma *K-Means Clustering* sehingga dapat mengelompokkan pelanggan berdasarkan karakteristik perilaku transaksi mereka [12]. Prosedur penelitian dalam penelitian ini dilakukan melalui beberapa tahapan yang disusun secara sistematis sebagaimana ditunjukkan pada Gambar 2.



Gambar 2. Alur Penelitian

### 2.1 Studi Literatur

Tahapan identifikasi masalah dan studi literatur krusial dilakukan agar manajemen dapat merumuskan strategi berbasis data analitis [3]. Kajian literatur memfasilitasi evaluasi parameter algoritma dari penelitian terdahulu guna menghasilkan desain *clustering* yang mutakhir [13]. Relevansi K-Means dan pemodelan RFM terus terbukti dalam riset terkini. Penelitian [14] mendemonstrasikan bahwa integrasi K-Means dan RFM pada data transaksi sangat akurat memetakan karakteristik pelanggan untuk menentukan strategi pemasaran. Dalam konteks industri spesifik, pendekatan *clustering* RFM juga terbukti tangguh sebagai dasar pengambilan keputusan manajerial [15]. Literatur-literatur terbaru tersebut memperkuat landasan bahwa integrasi metode ini sangat relevan untuk mendukung perancangan strategi bisnis seperti STP.

### 2.2 Persiapan Data

#### 2.2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini bersumber dari dua pendekatan, yaitu data primer dan sekunder. Data primer dikumpulkan melalui kuesioner matriks perbandingan berpasangan dan wawancara terstruktur bersama Pemilik Anhesa Gym yang bertindak sebagai pakar. Data ini nantinya digunakan untuk perhitungan bobot *Analytic Hierarchy Process* (AHP) dan validasi profil pelanggan (*expert judgment*). Di sisi lain, data sekunder diperoleh melalui ekstraksi riwayat transaksi pelanggan dari sistem pencatatan gym, yang mencakup data awal operasional pada tahun 2020 hingga periode penarikan data pada Februari 2026.

#### 2.2.2 Preprocessing & Transformasi RFM

Proses *preprocessing* dilakukan untuk memastikan dataset bebas dari anomali. Dari 1.000 riwayat transaksi awal, dilakukan *purposive filtering* ketat untuk mengeliminasi *missing values*, data duplikat, dan profil dengan variabel demografis tidak utuh. Data yang tereliminasi memiliki pola spesifik, yakni didominasi oleh pelanggan insidental (*daily pass*) dan anggota lama dengan pencatatan administrasi yang belum lengkap. Penyusutan signifikan menjadi 150 data final ini tidak merusak representativitas sampel, melainkan justru mengeliminasi bias (*noise*). Sebaliknya, 150 data yang tersisa secara akurat merepresentasikan populasi anggota inti (*core members*) yang aktivitas transaksinya relevan untuk penghitungan *Customer Lifetime Value* (CLV). Dataset bersih ini kemudian ditransformasikan ke dalam tiga dimensi utama: *Recency, Frequency, dan Monetary* (RFM):

##### 1. *Recency* (R)

Dihitung berdasarkan selisih waktu (dalam hari) antara tanggal acuan penelitian (*snapshot date* pada akhir Februari 2026) dengan tanggal transaksi atau kunjungan terakhir pelanggan.

2. *Frequency* (F)

Dihitung dari akumulasi total kedatangan atau transaksi pelanggan selama periode pengamatan.

3. *Monetary* (M).

Merupakan total akumulasi uang yang dibelanjakan pelanggan untuk berbagai layanan gym (iuran bulanan, pendaftaran, dll) sepanjang riwayat keanggotaan mereka.

### 2.3 Normalisasi Data

Metode ini dipilih karena mampu mempertahankan distribusi asli data sekaligus mengurangi dampak perbedaan skala secara proporsional [16]. Tanpa proses normalisasi, atribut dengan rentang nilai numerik yang jauh lebih besar akan secara otomatis mendominasi perhitungan jarak *Euclidean*, yang pada akhirnya mendistorsi akurasi hasil pengelompokan algoritma *K-Means* [17]. Penelitian ini menggunakan teknik *Min-Max Normalization* untuk mentransformasikan data asli ke dalam rentang skala baru antara 0 dan 1. Transformasi dilakukan dengan menerapkan persamaan 1 sebagai berikut.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Di mana  $X'$  merupakan nilai hasil normalisasi,  $X$  merupakan nilai data asli,  $X_{min}$  merupakan nilai minimum dari variabel tersebut, dan  $X_{max}$  merupakan nilai maksimum dari variabel tersebut. Implementasi normalisasi ini sangat krusial untuk meningkatkan stabilitas komputasi dan memastikan setiap dimensi RFM memberikan kontribusi yang seimbang dalam pembentukan kluster pada algoritma *K-Means* [18].

### 2.4 Pembobotan Variabel Dengan AHP

Penilaian pakar ditransformasikan ke dalam matriks perbandingan berpasangan menggunakan skala intensitas 1–9. Nilai perbandingan kriteria  $j$  terhadap  $i$  menggunakan rumus kebalikan (*reciprocal*) yang dituliskan pada persamaan 2.

$$a_{ji} = \frac{1}{a_{ij}} \quad (2)$$

Matriks tersebut dinormalisasi dengan membagi setiap elemen dengan jumlah total pada kolom yang bersangkutan ( $X_{ij}$ ) kemudian bobot prioritas didapatkan dengan menghitung nilai rata-rata dari setiap baris menggunakan persamaan 3.

$$W_i = \frac{\sum_{j=1}^n X_{ij}}{n} \quad (3)$$

Sebelum menghitung bobot akhir, perlu dilakukan perhitungan konsistensi menggunakan indeks inkonsistensi untuk memastikan bahwa penilaian yang diberikan rasional. Pengujian ini dilakukan dengan menghitung *Consistency Index* ( $C_i$ ) dan *Consistency Ratio* ( $C_r$ ) menggunakan persamaan 4 dan 5:

$$C_i = \frac{(\lambda_{max} - \text{Jumlah Kriteria})}{\text{Jumlah Kriteria} - 1} \quad (4)$$

$$C_r = \frac{C_i}{RI} \quad (5)$$

$RI$  Adalah *Random Index*. Nilai  $RI$  ditetapkan secara empiris oleh Saaty berdasarkan ukuran matriks untuk penelitian ini  $n = 3$  yang digunakan adalah 0,58 [19]. Pembobotan dinyatakan sah dan konsisten apabila nilai  $C_r \leq 0,1$  [20].

### 2.5 Pemodelan dan Evaluasi Kluster

#### 2.5.1 Penentuan Jumlah Kluster (*Elbow Method*)

Penentuan nilai  $K$  optimal dilakukan menggunakan *Elbow Method* dengan mengevaluasi perubahan nilai *Sum of Squared Error* (SSE) atau *Within Cluster Sum of Squares* (WCSS) pada berbagai jumlah kluster. Nilai SSE dihitung berdasarkan total jarak kuadrat antara titik data dan centroid pada setiap kluster, melalui persamaan 6.

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} (x - c_i)^2 \quad (6)$$

Di mana  $K$  adalah jumlah kluster dan  $C_i$  merupakan *centroid* kluster ke  $i$ . Titik di mana penurunan nilai SSE mulai melambat secara signifikan dan membentuk pola menyerupai siku (*elbow*) diidentifikasi sebagai jumlah kluster yang paling optimal [21].

### 2.5.2 Algoritma K-Means Clustering

*K-Means* merupakan algoritma pengelompokan non-hierarki yang mempartisi dataset ke dalam  $K$  kelompok berdasarkan kemiripan karakteristik perilaku pelanggan. Algoritma ini bekerja secara iteratif dengan menghitung jarak terkecil antara setiap data terhadap *centroid* menggunakan *Euclidean Distance* melalui persamaan 7.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (7)$$

Di mana  $d(x, y)$  adalah jarak antara data pelanggan  $x$  dengan pusat kluster  $y$ . Proses pembaruan posisi *centroid* dilakukan secara berulang hingga mencapai kondisi konvergen atau tidak terjadi perubahan signifikan pada pusat kluster [21].

### 2.5.3 Evaluasi Kualitas Kluster (Davies-Bouldin Index)

*Davies-Bouldin Index* (DBI) merupakan metrik evaluasi internal pengelompokan yang mengukur rasio perbandingan antara tingkat kepadatan di dalam kluster (*within-cluster scatter*) terhadap tingkat separasi atau jarak antar-pusat kluster (*between-cluster separation*) yang dihitung berdasarkan persamaan 8.

$$DBI = \frac{1}{k} \sum_{i=1}^k \max \left( \frac{s_i + a_j}{d_{ij}} \right) \quad (8)$$

Indikator kualitas pada DBI bersifat terbalik, yakni nilai indeks yang semakin kecil (mendekati nol) menunjukkan kualitas pengelompokan yang semakin optimal, karena merepresentasikan kluster yang padat di bagian dalam namun terpisah jauh dari kluster lainnya [22].

## 2.6 Hasil Dan Strategi Pemasaran

### 2.6.1 Perhitungan Customer Lifetime Value (CLV)

Perhitungan dilakukan dengan mengalikan nilai variabel RFM yang telah dinormalisasi dengan bobot prioritas yang diperoleh dari metode AHP melalui persamaan 9.

$$CLV_i = (w_R \times NR_i) + (w_f \times Nf_i) + (w_m \times NM_i) \quad (9)$$

Di mana  $CLV_i$  adalah nilai pelanggan ke  $i$  sedangkan  $w_R$ ,  $w_f$ ,  $w_m$  merupakan bobot kepentingan relatif untuk setiap dimensi RFM. Pendekatan ini memungkinkan pemeringkatan kluster berdasarkan profitabilitas dan loyalitas secara lebih objektif.

### 2.6.2 Interpretasi Profil Kluster

Pada tahap ini, karakteristik setiap kluster dianalisis berdasarkan nilai rata-rata RFM dan CLV yang dihasilkan. Proses interpretasi bertujuan untuk mengidentifikasi kelompok pelanggan bernilai tinggi (*high-value*), menengah, hingga pelanggan yang berisiko lepas (*churn*). Profiling kluster ini kemudian divalidasi melalui *expert judgment* bersama pihak manajemen gym guna memastikan temuan data selaras dengan realitas operasional di lapangan.

### 2.6.3 Perancangan Strategi Pemasaran (STP)

Berdasarkan profil kluster yang telah divalidasi, dirumuskan rekomendasi strategi pemasaran menggunakan pendekatan *Segmentation, Targeting, dan Positioning* (STP). Berikut adalah perinciannya:

1. *Segmentation* menggunakan hasil klusterisasi sebagai dasar pembagian pasar berdasarkan perilaku transaksi nyata.

2. *Targeting* Menentukan segmen prioritas yang memiliki nilai CLV tertinggi untuk diberikan alokasi sumber daya pemasaran yang lebih besar.
3. *Positioning* Merancang program promosi dan layanan yang personal sesuai dengan kebutuhan spesifik masing-masing segmen, seperti program loyalitas untuk pelanggan tetap atau penawaran khusus untuk menarik kembali pelanggan yang tidak aktif.

### 3. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil eksperimen pengelompokan pelanggan Anhesa Gym menggunakan integrasi model RFM, algoritma K-Means, dan analisis nilai ekonomi melalui CLV.

#### 3.1. Pengumpulan dan Pembersihan Data

Data mentah yang ditarik dari sistem informasi transaksi Anhesa Gym mencakup riwayat aktivitas pelanggan dari tahun 2020 hingga Februari 2026 dengan total mencapai 1.000 entitas. Namun, setelah melalui tahap *data cleaning* untuk menangani *missing values* pada atribut demografis (usia, pendapatan, dan pekerjaan) serta penghapusan data duplikat, diperoleh 150 data pelanggan yang valid dan siap diproses. Penggunaan dataset yang bersih dan lengkap ini bertujuan untuk menjamin stabilitas komputasi dan mencegah bias pada hasil klusterisasi. Berikut adalah atribut data yang disajikan dalam Tabel 1.

**Tabel 1.** Atribut Data

| Atribut                | Keterangan                                                        | Skala Data  |
|------------------------|-------------------------------------------------------------------|-------------|
| ID Pelanggan           | Berisi ID unik masing-masing pelanggan                            | Kategorikal |
| <i>Last Visit</i>      | Tanggal terakhir pelanggan berkunjung                             | Rasio       |
| <i>Total Of Orders</i> | Jumlah pesanan yang dilakukan pelanggan                           | Rasio       |
| <i>Amount Lifetime</i> | Jumlah uang yang dikeluarkan pelanggan selama bertransaksi di gym | Rasio       |

#### 3.2. Transformasi Variabel RFM

Dataset yang telah dibersihkan kemudian ditransformasikan menjadi variabel *Recency* (R), *Frequency* (F), dan *Monetary* (M). Tabel 2 menunjukkan cuplikan hasil perhitungan variabel RFM untuk lima pelanggan pertama sebelum tahap normalisasi dilakukan.

**Tabel 2.** Transformasi RFM

| ID Pelanggan | R   | F   | M       |
|--------------|-----|-----|---------|
| Cust 001     | 111 | 10  | 227000  |
| Cust 002     | 70  | 45  | 566500  |
| Cust 003     | 1   | 272 | 4407000 |
| ...          | ... | ... | ...     |
| Cust 150     | 64  | 8   | 209000  |

#### 3.3. Normalisasi Data

Dengan mengimplementasikan metode *Min-Max Normalization*, seluruh data ditransformasikan ke dalam rentang 0 hingga 1. Proses ini menjamin stabilitas komputasi algoritma K-Means sehingga hasil pengelompokan pelanggan menjadi lebih akurat dan representatif. Berikut adalah hasil normalisasi data yang disajikan pada Tabel 3.

**Tabel 3.** Normalisasi Data

| ID Pelanggan | R     | F     | M     |
|--------------|-------|-------|-------|
| Cust 001     | 0,071 | 0,09  | 0,018 |
| Cust 002     | 0,045 | 0,043 | 0,048 |
| Cust 003     | 0,01  | 0,266 | 0,385 |
| ...          | ...   | ...   | ...   |
| Cust 150     | 0,041 | 0,007 | 0,016 |

#### 3.4. Bobot AHP

Pendekatan ini memastikan bahwa segmentasi tidak hanya berbasis data kuantitatif, tetapi juga selaras dengan prioritas strategi bisnis. Berikut adalah rincian langkah-langkah nya:

1. Matriks Perbandingan Berpasangan

Tahap pertama adalah mengonversi hasil penilaian kuesioner ke dalam matriks perbandingan berpasangan. Tabel 4 menunjukkan hasil perbandingan tingkat kepentingan antarvariabel *Recency* (R), *Frequency* (F), dan *Monetary* (M).

**Tabel 4.** Matriks Perbandingan Berpasangan

| Kriteria | R     | F  | M     |
|----------|-------|----|-------|
| R        | 1     | 7  | 1/3   |
| F        | 1/7   | 1  | 1/9   |
| M        | 3     | 9  | 1     |
| Total    | 4,142 | 17 | 1,444 |

## 2. Normalisasi dan Bobot Prioritas

Matriks perbandingan kemudian dinormalisasi dengan membagi setiap nilai sel dengan total kolomnya. Nilai rata-rata dari setiap baris pada matriks normalisasi ini akan menjadi bobot prioritas (*Eigenvector*) yang digunakan sebagai pengali dalam perhitungan *Customer Lifetime Value* [20]. Bobot prioritas masing-masing variabel ditunjukkan dalam Tabel 5.

**Tabel 5.** Bobot Prioritas

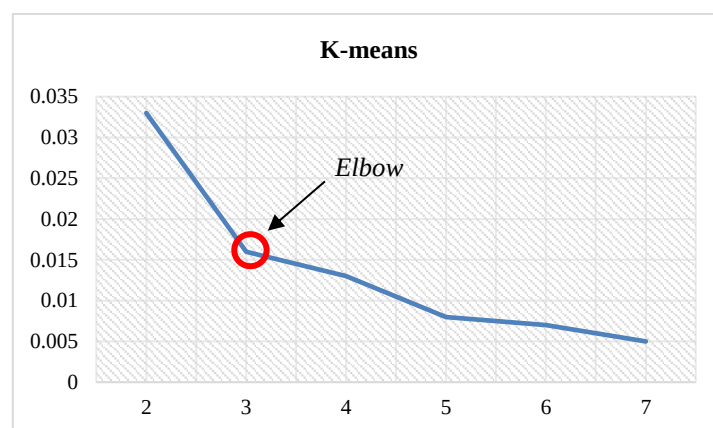
| Kriteria | Norm R | Norm F | Norm M | Bobot (W) |
|----------|--------|--------|--------|-----------|
| R        | 0,2413 | 0,411  | 0,23   | 0,294     |
| F        | 0,034  | 0,058  | 0,076  | 0,056     |
| M        | 0,724  | 0,529  | 0,692  | 0,648     |
| Total    |        |        |        | 1         |

Berdasarkan perkalian matriks baris dan perhitungan nilai eigen maksimum, diperoleh nilai ( $\lambda_{max}$ ) = 3,081, Tahap akhir adalah menguji rasionalitas penilaian pakar melalui *Consistency Index* (CI) dan *Consistency Ratio* (CR) dengan jumlah kriteria ( $n = 3$ ) dan *Random Indeks* (IR = 0,58) Hasil komputasi menunjukkan nilai CI sebesar 0,0406 dan CR sebesar 0,07. Mengingat nilai  $CR \leq 0,1$ , penilaian perbandingan matriks berpasangan disimpulkan konsisten. Merujuk pada bobot akhir di Tabel 5, variabel *Monetary* memperoleh bobot paling dominan (0,648). Dominasi ini sangat rasional karena industri layanan kebugaran memiliki struktur biaya operasional tetap yang tinggi (perawatan alat, penyusutan aset, dan utilitas). Oleh karena itu, kemampuan pelanggan dalam menghasilkan aliran kas (*cash flow*) secara langsung menjadi indikator paling krusial bagi stabilitas finansial Anhesa Gym dibandingkan sekadar parameter seberapa sering mereka berkunjung (*Frequency*).

## 3.5. Pemodelan dan Evaluasi

### 3.5.1 Penentuan Kluster Optimal (*Elbow Method*)

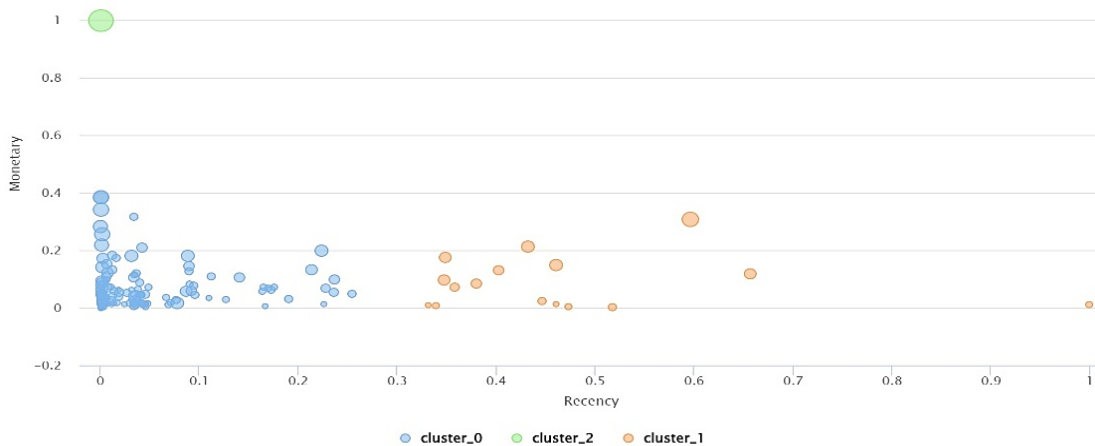
Pengujian dilakukan pada rentang  $k = 2$  hingga  $k = 7$  dengan mengukur perubahan nilai *Sum of Squared Error* (SSE). Berdasarkan visualisasi grafik pada Gambar 3, terlihat bahwa penurunan nilai SSE yang awalnya drastis mulai melambat dan membentuk patahan siku (*elbow*) yang signifikan pada titik ke 3 yang berarti  $K = 3$ . Oleh karena itu  $K = 3$  ditetapkan sebagai parameter jumlah kluster optimal untuk penelitian ini yang ditunjukkan pada Gambar 3.

**Gambar 3.** *Elbow Method*

### 3.5.2 Hasil K-Means Clustering

Berdasarkan nilai  $k=3$ , algoritma mempartisi 150 data pelanggan ke dalam kelompok yang memiliki karakteristik serupa. Kualitas pemisahan struktur kluster ini telah dievaluasi menggunakan *Davies-Bouldin Index* (DBI) dan menghasilkan nilai absolut sebesar 0,436. Mengingat indikator DBI yang semakin mendekati nol menunjukkan separasi kluster yang semakin baik, nilai 0,436 mengonfirmasi bahwa algoritma

telah berhasil mengelompokkan karakteristik pelanggan secara optimal. Visualisasi sebaran titik pelanggan pada masing-masing kluster divisualisasikan melalui *scatter plot* pada Gambar 4.



**Gambar 4.** Visualisasi Scatter Plot Klaster Pelanggan

### 3.6. Analisis CLV dan Perancangan Strategi STP

#### 3.6.1 Penghitungan *Customer Lifetime Value* (CLV)

Untuk mempermudah interpretasi, centroid hasil klastering disintesis dengan bobot AHP guna menghasilkan skor CLV. Rangkuman profil statistik dan skor akhir CLV untuk setiap kluster disajikan pada Tabel 6.

**Tabel 6.** Profil Statistik dan Skor CLV Pelanggan

| Klster    | Jml Anggota | Rata-rata Recency | Rata-rata Frequency | Rata-rata Monetary | Skor CLV | Label Segmen      |
|-----------|-------------|-------------------|---------------------|--------------------|----------|-------------------|
| Klaster 2 | 1           | 0                 | 1020                | 11.402.000         | 1        | <i>Loyal/VIP</i>  |
| Klaster 0 | 133         | 69,5              | 39                  | 786.469            | 0,326    | <i>Potential</i>  |
| Klaster 1 | 16          | 732,8             | 73                  | 1.026.750          | 0,217    | <i>Lost/Churn</i> |

Pelanggan dapat diinterpretasikan secara mendalam sebagai berikut. Klaster 2 (*Loyal/VIP*) adalah segmen paling bernilai dengan skor CLV absolut (1,000). Meskipun hanya beranggotakan 1 entitas, pelanggan ini mencatatkan transaksi historis mencapai Rp 11,4 juta dengan jarak kunjungan terakhir 0 hari (sangat aktif), menandakan dedikasi dan kontribusi finansial maksimal. Klaster 0 (*Potential*) adalah kelompok mayoritas (133 anggota) dengan skor CLV 0,326. Segmen ini merepresentasikan pelanggan reguler yang masih aktif berkunjung (rata-rata *Recency* 69 hari) namun tingkat *spending*-nya belum maksimal, sehingga memiliki ruang pertumbuhan yang luas. Sebaliknya, Klaster 1 (*Lost/Churn*) memiliki skor CLV terendah (0,217). Walaupun secara historis mereka pernah melakukan transaksi finansial yang cukup besar, jarak kunjungan terakhir mereka sangat buruk (rata-rata 732 hari atau absen >2 tahun), mengindikasikan secara kuat bahwa mereka adalah pelanggan yang telah putus hubungan atau beralih ke kompetitor.

#### 3.6.2 Perancangan Strategi Pemasaran (STP)

Tahap akhir dari penelitian ini adalah menerjemahkan peringkat loyalitas pelanggan ke dalam strategi bisnis melalui kerangka kerja *Segmentation, Targeting, dan Positioning* (STP). Langkah ini bertujuan untuk mempersonalisasi pendekatan pemasaran tidak hanya berdasarkan riwayat transaksi, tetapi juga profil demografis pelanggan. Berdasarkan hasil penggabungan analisis RFM dan skor CLV, pelanggan Anhesa Gym dibagi menjadi tiga segmen utama. Indikator perilaku pelanggan pada tiap segmen ditandai dengan simbol panah (↑↓) untuk mempermudah identifikasi karakteristik

##### 1. *Segmentation* (Segmentasi)

Berdasarkan hasil penggabungan analisis RFM, skor CLV, dan data kuesioner, pelanggan Anhesa Gym dikelompokkan menjadi tiga segmen utama dengan profil demografis pada Tabel 7.

**Tabel 7.** Segmentasi Pelanggan

| Klaster   | Karakteristik RFM | Karakteristik Demografis                         |
|-----------|-------------------|--------------------------------------------------|
| Klaster 2 | R↓, F↑, M↑        | Usia 30-35 Tahun dengan pendapatan lebih dari Rp |

| Klaster   | Karakteristik RFM                                          | Karakteristik Demografis                                                                                                                     |
|-----------|------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------|
|           |                                                            | 4.000.000, yang didominasi oleh Pegawai Swasta / Karyawan / PNS                                                                              |
| Klaster 0 | $R \leftrightarrow$ , $F \downarrow$ , $M \leftrightarrow$ | Usia 17-40 tahun dengan rata-rata pendapatan Rp 2.000.000, yang didominasi oleh mahasiswa/pelajar                                            |
| Klaster 1 | $R \uparrow$ , $F \leftrightarrow$ , $M \leftrightarrow$   | Rentang usia 24-30 tahun dengan pendapatan lebih dari Rp 3.000.000 yang didominasi oleh <i>freelance</i> dan Pegawai Swasta / Karyawan / PNS |

Keterangan: Simbol  $R \downarrow$  menunjukkan nilai hari yang rendah/baru berkunjung, yang secara matematis memberikan kontribusi positif pada perhitungan CLV.

## 2. Targeting (Penargetan)

Penentuan target berpedoman pada pemeringkatan segmen yang memberikan kontribusi *lifetime value* terbesar untuk menjaga stabilitas operasional dan finansial perusahaan [23]. Berikut adalah strategi penargetan yang difokuskan pada optimalisasi sumber daya berdasarkan potensi ekonomi dan profil status pelanggan:

- Prioritas Utama (Retensi): Diarahkan pada Klaster 2 (*Loyal/VIP*). Karena segmen ini memiliki daya beli tinggi dan merupakan penggerak utama *revenue*, menjaga kepuasan mereka adalah prioritas mutlak.
- Prioritas Sekunder (Pengembangan): Diarahkan pada Klaster 0 (*Potential*). Segmen ini memiliki potensi pertumbuhan yang besar jika didorong dengan program loyalitas yang sesuai dengan gaya hidup mereka.
- Efisiensi (Reaktivasi): Diarahkan pada Klaster 1 (*Low-Value*). Mengingat keterbatasan *budget* mahasiswa pada segmen ini, pendekatan dilakukan melalui efisiensi biaya operasional pemasaran.

## 3. Positioning (Penempatan)

Penempatan posisi merek Anhesa Gym disesuaikan dengan kebutuhan psikografis dan kemampuan finansial tiap segmen sebagaimana dirinci pada Tabel 8. Strategi ini bertujuan untuk mengevaluasi dan merancang kampanye pemasaran yang relevan serta efektif dalam memperpanjang siklus hidup pelanggan [24].

**Tabel 8.** Penerapan Strategi Pemasaran

| Segmen           | Strategi Pemasaran ( <i>Positioning</i> )                                                                                                                                                                                                                  |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Loyal/VIP</i> | Eksklusivitas & Layanan Premium: Menekankan pada nilai kenyamanan dan prestise. Implementasi: Fasilitas loker VIP, prioritas akses alat saat jam sibuk, atau apresiasi berupa <i>merchandise</i> khusus sebagai pengikat komunitas elit.                   |
| <i>Potential</i> | Keterikatan & Gaya Hidup: Menawarkan paket yang mendukung gaya hidup aktif mahasiswa/pekerja muda. Implementasi: Promo perpanjangan jangka menengah (3-6 bulan) atau diskon <i>bundling personal trainer</i> dengan harga menengah.                        |
| <i>Low-Value</i> | Aksesibilitas & Fleksibilitas: Memposisikan gym sebagai fasilitas olahraga yang terjangkau secara harga. Implementasi: Penawaran diskon <i>Daily Pass</i> untuk mahasiswa atau promo kelompok ("Daftar Berdua Lebih Murah") untuk memacu volume kunjungan. |

## 3.7. Diskusi Hasil

Temuan dari penelitian ini secara signifikan memperluas kajian dari riset terdahulu [8] dan [11] yang umumnya hanya berfokus pada evaluasi matematis K-Means, dengan cara mengintegrasikan hasil klaster langsung ke dalam kerangka STP serta menggunakan AHP untuk meminimalkan bias asumsi manajerial. Meski demikian, terdapat empat keterbatasan utama yang perlu dipertimbangkan. Pertama, ukuran sampel final yang menyusut menjadi 150 pelanggan, meskipun esensial untuk membuang anomali data (*noise*) dapat membatasi representasi populasi secara makro. Kedua, pembobotan AHP yang bertumpu pada pakar tunggal (Pemilik Anhesa Gym) rentan terhadap bias subjektivitas, kendati secara manajerial valid karena mewakili otoritas visi bisnis [19]. Ketiga, fokus eksklusif pada satu cabang gym membuat luaran riset bersifat kontekstual, sehingga generalisasinya pada pusat kebugaran lain harus dikaji secara hati-hati. Terakhir, terdapat potensi bias data karena pemodelan saat ini murni bergantung pada metrik transaksi finansial (RFM). Penggabungan riwayat transaksi dengan analitik data perilaku non-finansial (seperti durasi latihan atau preferensi kelas) sangat direkomendasikan untuk penelitian mendatang guna menghasilkan profil yang lebih holistik.

#### 4. KESIMPULAN

Penerapan algoritma K-Means Clustering berbasis model RFM terbukti efektif dalam memetakan profil pelanggan di Anhesa Gym secara objektif, menghasilkan tiga kelompok pelanggan valid dengan nilai *Davies-Bouldin Index* (DBI) sebesar 0,436. Integrasi metode *Analytic Hierarchy Process* (AHP) turut memberikan bobot yang akurat pada variabel *Monetary* (0,650), *Recency* (0,294), dan *Frequency* (0,056), yang kemudian digunakan dalam analisis *Customer Lifetime Value* (CLV) untuk mengurutkan prioritas pelanggan secara adil melalui teknik normalisasi terbalik pada variabel *Recency*. Hasil perhitungan CLV tersebut menetapkan Klaster 2 sebagai segmen Loyal/VIP (skor mutlak 1,000), disusul Klaster 0 sebagai segmen Potential (skor 0,327), dan Klaster 1 sebagai segmen Low-Value/Churn (skor 0,217). Sebagai tindak lanjut dari pemeringkatan tersebut, dirancang strategi pemasaran *Segmentation, Targeting, dan Positioning* (STP) yang dipersonalisasi sesuai profil demografis, dengan fokus utama pada retensi segmen VIP (yang didominasi kalangan profesional) melalui layanan eksklusif, serta pengembangan segmen potensial (mayoritas mahasiswa dan pekerja muda) melalui penawaran paket keanggotaan yang kompetitif guna mengoptimalkan pengelolaan hubungan pelanggan dan meminimalkan risiko *churn* di masa mendatang. Melalui integrasi ini, penelitian ini berkontribusi secara praktis dalam menerjemahkan luaran matematis algoritma menjadi rancangan keputusan manajerial yang konkret. Meskipun demikian, penelitian ini memiliki keterbatasan karena pemodelannya hanya bergantung pada riwayat transaksi finansial. Oleh karena itu, penelitian lanjutan direkomendasikan untuk menguji metode *clustering* lain, seperti DBSCAN atau K-Medoids, serta mengintegrasikan data perilaku pelanggan nontransaksional, seperti preferensi alat atau durasi latihan, guna menghasilkan segmentasi yang lebih komprehensif.

#### REFERENSI

- [1] A. M. Newsome, R. Reed, J. Sansone, A. Batrakoulis, C. McAvoy, and M. W. Parrott, "2025 ACSM Worldwide Fitness Trends: Future Directions of the Health and Fitness Industry," *ACSMs Health Fit. J.*, vol. 28, no. 6, pp. 11–25, 2024, doi: 10.1249/FIT.0000000000001017.
- [2] Fortune Business Insights, "Health and Fitness Club Market Size, Share & Industry Analysis." Accessed: Jun. 04, 2026. [Online]. Available: <https://www.fortunebusinessinsights.com/health-and-fitness-club-market-108652>
- [3] P. Kotler, K. L. Keller, and A. Chernev, *Marketing management, global edition*, vol. 832. New York: Pearson, 2022.
- [4] S. Supangat and Y. Mulyani, "Customer Loyalty Analysis Using Recency, Frequency, Monetary (RFM) and K-means Cluster for Labuan Bajo Souvenirs in Online Store," *Journal of Information Systems and Informatics*, vol. 5, no. 1, pp. 285–299, 2023, doi: 10.51519/journalisi.v5i1.421.
- [5] F. Handayani, J. Kadang, and I. Syrifuddin, "Penerapan Strategi Pemasaran STP (Segmenting, Targeting, Positioning) pada Usaha Toreko," *Empiricism Journal*, vol. 4, no. 1, pp. 208–212, 2023, doi: 10.36312/ej.v4i1.1170.
- [6] H. Kusuma Ningrum and M. Indira Hasmarini, "Marketing Strategy Analysis Using STP Strategy (Segmentation, Targeting, Positioning) at PT. Berkah Ridho Cinta Indonesia," *Income Journal of Economics Development*, vol. 4, no. 12, pp. 15–19, 2024, doi: 10.54065/ijed.4.1.2024.205.
- [7] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, "RFM ranking – An effective approach to customer segmentation," *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 10, pp. 1251–1257, 2021, doi: 10.1016/j.jksuci.2018.09.004.
- [8] P. Anitha and M. M. Patil, "RFM model for customer purchase behavior using K-Means algorithm," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 5, pp. 1785–1792, May 2022, doi: 10.1016/J.JKSUCI.2019.12.011.
- [9] Z. Niu, "Customer segmentation and management strategy optimization for gyms using k-means clustering," in *Proceedings of the 1st International Conference on E-commerce and Artificial Intelligence - Volume 1: ECAI*, SciTePress, 2024, pp. 239–242. doi: 10.5220/0013214000004568.
- [10] C. Rungruang, P. Riyapan, A. Intarasit, K. Chuarkham, and J. Muangprathub, "RFM model customer segmentation based on hierarchical approach using FCA," *Expert Syst. Appl.*, vol. 237, p. 121449, 2024, doi: 10.1016/j.eswa.2023.121449.
- [11] D. S. Nugraha, I. Thoib, N. Sururi, and B. P. Candra, "Segmentasi Pelanggan Berbasis Analisis RFM Menggunakan Algoritma K-Means Clustering," *Jurnal Komputer Teknologi Informasi Sistem Informasi (JUKTISI)*, vol. 4, no. 2, pp. 1361–1369, 2025, doi: 10.62712/juktisi.v4i2.641.
- [12] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2011.
- [13] Sugiyono, *Pendekatan Kuantitatif, Kualitatif dan R & D*. Bandung: Alfabeta, 2008.
- [14] B. Ilham, R. Rahmadden, A. Putra, D. Najario, and M. A. Fakhrizal, "Optimasi Performa K-Means melalui Hybrid Feature Engineering RFM dan Behavioral Analytics untuk Segmentasi Pelanggan: Optimizing K-Means Performance with Hybrid RFM and Behavioral Analytics for Customer

- Segmentation,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 4, pp. 1377–1386, 2025, doi: 10.57152/malcom.v5i4.2227.
- [15] M. Shodikin and N. Taniasari, “Analisis Komparasi Algoritma DBSCAN dan K-Means dalam Pemetaan Segmentasi Pasien Rawat Inap Menggunakan Model RFMT: Comparative Analysis of DBSCAN and K-Means Algorithms in Inpatient Patient Segmentation Using the RFMT Model Approach,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 4, pp. 1231–1241, 2025, doi: 10.57152/malcom.v5i4.2218.
- [16] M. A. Fikri, I. D. W. Prastya, and A. Y. Kartini, “Comparison of K-Means and K-Medoids for PKH Priority Clustering,” *Journal of Applied Informatics and Computing*, vol. 10, no. 2, pp. 1347–1358, 2026, doi: 10.30871/jaic.v10i2.12267.
- [17] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. O'Reilly Media, Inc., 2022.
- [18] C. M. Bishop and H. Bishop, *Deep learning: Foundations and concepts*. Springer Cham, 2023. doi: 10.1007/978-3-031-45468-4.
- [19] T. L. Saaty, “How to make a decision: The analytic hierarchy process,” *Eur. J. Oper. Res.*, vol. 48, no. 1, pp. 9–26, Sep. 1990, doi: 10.1016/0377-2217(90)90057-1.
- [20] F. Rusly, R. Purba, and M. F. Pasha, “New Approach: Customer Segmentation Using RFM Model and Demand Classification,” *Jurnal Teknologi dan Manajemen Informatika*, vol. 11, no. 2, pp. 165–176, 2025, doi: 10.26905/jtmi.v11i2.16208.
- [21] A. Alamsyah *et al.*, “Customer Segmentation Using the Integration of the Recency Frequency Monetary Model and the K-Means Cluster Algorithm,” *Scientific Journal of Informatics*, vol. 9, no. 2, pp. 189–196, Nov. 2022, doi: 10.15294/sji.v9i2.39437.
- [22] D. L. Davies and D. W. Bouldin, “A cluster separation measure,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, 1979, doi: 10.1109/TPAMI.1979.4766909.
- [23] D. N. Aini, I. Ramadhani, and D. Erlangga, “Analisis Segmentasi Pelanggan Menggunakan Model Lrfm Dan Teknik Clustering,” *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 9, no. 2, pp. 110–117, 2023, doi: 10.24014/rmsi.v9i2.24311.
- [24] H. Andy, A. Fachmi, R. R. Lintang, M. I. Nuur, S. Aji, and R. J. Nila, “Analisis segmentasi pelanggan berbasis RFM dan evaluasi efektivitas kampanye pemasaran untuk meningkatkan retensi,” *Neptunus*, vol. 2, no. 4, pp. 28–42, 2024, doi: 10.61132/neptunus.v2i4.400.