

Pengaruh Analisis Customer Lifetime Value Menggunakan RFM Dan K-Means Clustering Terhadap Strategi Pemasaran Industri E-Commerce

Juhairiah^{1*}, Taufik Zulfikar²

^{1,2}S2 Magister Manajemen, Universitas Sangga Buana, Bandung

Correspondence: juhairiah@hamzanwadi.ac.id

Received: 17 Juni 2026 | Revised: 20 Juli 2026 | Accepted: 25 Agustus 2026

Keywords: *Customer*

Lifetime Value;

Customer

Segmentation; E-

Commerce; K-Means

Clustering; RFM

Abstract

The rapid growth of the e-commerce industry has intensified competition in customer retention and business profitability. Therefore, companies need to understand customer characteristics to develop targeted marketing strategies. This study aims to analyze Customer Lifetime Value (CLV) using the Recency, Frequency, Monetary (RFM) approach and K-Means Clustering to support marketing decision-making. The research methodology follows the CRISP-DM framework utilizing the Online Retail Dataset consisting of 4,338 customers. The K-Means algorithm successfully classified customers into five main segments: Potential Customers, Hibernating Customers, Champions, Loyal Customers, and VIP Customers. The VIP Customers segment achieved the highest average CLV of 572,425.61, while Hibernating Customers had the lowest average CLV of 1,435.94. Furthermore, Potential Customers contributed the largest share of revenue due to their dominant population size. The integration of RFM, K-Means, and CLV provides deeper insights into customer behavior to support more effective marketing strategies

Kata Kunci:

Customer Lifetime

Value; K-Means

Clustering; RFM;

Segmentasi

Pelanggan;

Strategi

Pemasaran

Abstract

Perkembangan industri e-commerce yang semakin pesat menyebabkan meningkatnya persaingan dalam mempertahankan pelanggan dan meningkatkan profitabilitas bisnis. Oleh karena itu, perusahaan perlu memahami karakteristik pelanggan untuk menyusun strategi pemasaran yang lebih efektif dan tepat sasaran. Penelitian ini bertujuan untuk menganalisis Customer Lifetime Value (CLV) melalui pendekatan Recency, Frequency, Monetary (RFM) dan K-Means Clustering guna menghasilkan segmentasi pelanggan yang dapat mendukung pengambilan keputusan pemasaran pada industri e-commerce. Metode penelitian yang digunakan mengacu pada kerangka kerja CRISP-DM dengan menggunakan Online Retail Dataset yang terdiri dari 4.338 pelanggan. Hasil penelitian menunjukkan bahwa algoritma K-Means berhasil mengelompokkan pelanggan menjadi lima segmen utama, yaitu Potential Customers, Hibernating Customers, Champions, Loyal Customers, dan VIP Customers. Segmen VIP Customers memiliki nilai CLV tertinggi dengan rata-rata sebesar 572.425,61, sedangkan Hibernating Customers memiliki nilai CLV terendah sebesar 1.435,94. Selain itu, Potential Customers memberikan kontribusi pendapatan terbesar karena jumlah pelanggan yang dominan. Integrasi analisis RFM, K-Means Clustering, dan CLV mampu memberikan pemahaman yang lebih mendalam mengenai perilaku pelanggan serta mendukung penyusunan strategi pemasaran yang lebih efektif

PENDAHULUAN

Industri e-commerce baik di Indonesia maupun global mengalami pertumbuhan yang sangat pesat, ditandai oleh peningkatan jumlah transaksi digital setiap tahun yang menggambarkan peningkatan penetrasi belanja online dan adopsi teknologi pembayaran [1], [2]. Pertumbuhan ini memicu persaingan bisnis yang lebih ketat di antara perusahaan e-commerce, sehingga perusahaan tidak cukup hanya fokus pada akuisisi pelanggan baru; retensi pelanggan menjadi kunci diferensiasi kompetitif yang berkelanjutan karena biaya memperoleh pelanggan baru (CAC) biasanya lebih tinggi dibandingkan biaya menjaga pelanggan yang sudah ada [3], [4], [5]. Seiring dengan meningkatnya kompleksitas perilaku konsumen online, diperlukan pendekatan analitik yang dapat menggabungkan segmentasi pelanggan dengan evaluasi nilai ekonomi pelanggannya (CLV) untuk mengoptimalkan alokasi sumber daya pemasaran dan memperkuat posisi persaingan perusahaan [6], [7] .

Permasalahan bisnis utama terletak pada tingginya CAC relatif terhadap biaya retensi, serta adanya keragaman kontribusi pendapatan antar pelanggan yang tidak selalu sejalan dengan strategi pemasaran massal saat ini. Banyak perusahaan menerapkan mass marketing untuk seluruh basis pelanggan, sehingga strategi pemasaran menjadi kurang efektif dan terjadinya pemborosan anggaran promosi serta kesulitan dalam mengidentifikasi pelanggan yang paling bernilai [3], [5]. Dampak dari hal ini antara lain ketidakpastian dalam prioritas pelanggan bernilai tinggi, serta rendahnya efisiensi promosi dan ROI (return on investment) pemasaran yang berujung pada rendahnya retensi jangka panjang [4], [7] . Beberapa studi sebelumnya menyoroti bahwa bidang analisis RFM, K-Means, atau CLV secara terpisah belum cukup untuk memberikan rekomendasi strategi yang komprehensif; penelitian yang mengintegrasikan ketiga metodologi tersebut masih terbatas, sehingga terdapat celah penelitian yang dapat diisi dengan pendekatan terintegrasi [4], [5], [8] .

Solusi utama adalah adopsi dan operasionalisasi Customer Lifetime Value (CLV) sebagai ukuran nilai ekonomi pelanggan sepanjang hubungan bisnis, yang memungkinkan perusahaan menentukan prioritas pelanggan, mengoptimalkan alokasi sumber daya pemasaran, serta meningkatkan pengambilan keputusan strategis dan efektivitas promosi [1], [3], [8]. Integrasi CLV dengan analisis perilaku transaksi pelanggan akan lebih optimal jika dipadukan dengan teknik analisis transaksi, seperti RFM, sehingga pelaku bisnis dapat melihat tidak hanya kapan pembelian terakhir terjadi (recency) dan seberapa sering pembelian (frequency), tetapi juga berapa besar pendapatan yang dihasilkan oleh pelanggan dalam rentang waktu tertentu (monetary) dan bagaimana nilai ekonomi pelanggan tersebut berkontribusi terhadap profitabilitas jangka panjang [4], [5], [9], [10]. Dengan demikian, kombinasi RFM-K-Means-CLV berpotensi menghasilkan segmentasi pelanggan yang lebih komprehensif, serta memberikan dasar untuk strategi pemasaran yang lebih presisi dan berbasis nilai ekonomis pelanggan ([3], [7], [11], [12] .

Pendekatan inti terdiri dari tiga komponen utama: RFM untuk analisis perilaku transaksi

berdasarkan Recency, Frequency, dan Monetary; algoritma K-Means clustering untuk membentuk cluster pelanggan secara otomatis menjadi kelompok homogen yang memudahkan identifikasi karakteristik; dan perhitungan CLV untuk mengukur nilai ekonomi pelanggan dalam tiap cluster. RFM menyediakan tiga dimensi yang telah lama digunakan dalam segmentasi pelanggan, sedangkan K-Means berfungsi sebagai teknik clustering yang umum dipakai untuk membentuk segmen tanpa pengawasan (unsupervised). Integrasi CLV dengan RFM-K-Means memungkinkan penilaian nilai pelanggan secara lebih holistik, yakni memadukan perilaku pembelian dengan kontribusi pendapatan yang dihasilkan, sehingga rekomendasi pemasaran dapat disesuaikan secara spesifik untuk setiap segmen (misalnya fokus pada cluster dengan CLV tinggi dan RFM kuat) [1], [4], [9], [13]. Beberapa literatur juga menegaskan bahwa kombinasi RFM dengan K-Means dan CLV meningkatkan akurasi segmentasi dan efektivitas promosi dibandingkan penggunaan RFM atau clustering secara terpisah [3], [5], [14].

Penelitian ini menonjolkan kebaruan dengan mengintegrasikan tiga komponen utama: RFM untuk analisis perilaku transaksi, K-Means untuk segmentasi pelanggan, dan CLV untuk mengukur nilai ekonomi pelanggan, sehingga menghasilkan segmentasi yang lebih komprehensif dan penilaian nilai pelanggan yang berlandaskan ekonomi. Keunggulan pendekatan ini mencakup identifikasi pelanggan bernilai tinggi dan rendah berdasarkan kombinasi perilaku dan nilai ekonominya, serta rekomendasi strategi pemasaran yang lebih tepat sasaran dan efektif, sehingga mendukung retensi pelanggan dan keunggulan kompetitif bagi perusahaan e-commerce [1], [7], [15], [16]. Tujuan penelitian mencakup pengelompokan pelanggan berdasarkan perilaku transaksi menggunakan RFM dan K-Means, pengukuran nilai pelanggan via CLV, identifikasi kelompok bernilai tinggi dan rendah, serta penyusunan rekomendasi kebijakan pemasaran berbasis karakteristik dan nilai pelanggan. Manfaat teoretis meliputi pengembangan integrasi metodologi RFM-K-Means-CLV dalam literatur customer analytics, sedangkan manfaat praktis meliputi pemahaman karakteristik pelanggan, dukungan keputusan berbasis data, peningkatan efektivitas promosi dan retensi, serta kontribusi pada keunggulan bersaing melalui manajemen pelanggan yang lebih terkelola dengan baik [1], [7], [17], [18]

METODE

Penelitian ini mengikuti alur CRISP-DM (Cross Industry Standard Process for Data Mining) sebagai kerangka kerja utama untuk melakukan analisis perilaku transaksi pelanggan dan menurunkan strategi pemasaran berbasis Customer Lifetime Value (CLV) pada e-commerce di Indonesia. Pendekatan Data Mining dipilih untuk memanfaatkan pola-pola transaksi pelanggan yang tersebar pada dataset ritel online dan untuk menghasilkan segmentasi pelanggan yang informatif secara ekonomis. CRISP-DM membagi proses menjadi enam fase: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Struktur ini konsisten dengan literatur praktik CRISP-DM yang luas dan relevan untuk konteks penelitian serupa [19], [20], [21]. Pada penelitian ini, data yang digunakan adalah Online Retail

Dataset dengan jumlah pelanggan sebanyak 4.338, serta variabel yang dipertimbangkan yaitu CustomerID, InvoiceDate, Quantity, UnitPrice, dan Revenue, sesuai dengan karakteristik dataset transaksi ritel online yang lazim dipakai untuk analisis RFM dan klasterisasi berbasis K-Means [19], [20], [22], [23]. Langkah-langkahnya diuraikan secara menyeluruh di bawah ini, dengan penekanan bahwa setiap klaim dan metodologi didasarkan pada literatur yang relevan dan konsisten dengan tujuan penelitian.

Fase Business Understanding dimulai dengan merumuskan tujuan bisnis: memahami perilaku transaksi pelanggan untuk mendukung strategi pemasaran berbasis CLV pada platform e-commerce Indonesia. Tujuan ini sejalan dengan praktik data mining yang menekankan pentingnya merumuskan tujuan bisnis secara eksplisit sebelum proses teknis berlangsung [19], [21]. Dalam fase Data Understanding, data Online Retail Dataset diperoleh dan diperiksa kualitasnya, termasuk identifikasi variabel kunci (CustomerID, InvoiceDate, Quantity, UnitPrice, Revenue) serta ukuran sampel (4.338 pelanggan). Konteks penggunaan variabel-variabel tersebut selaras dengan paket praktik RFM dasar dan perhitungan CLV berbasis transaksi [19], [23]. Perbedaan pandangan antara studi sebelumnya terkait detail dataset (misalnya tanggal, negara, atau atribut lain) dicatat jika relevan; dalam hal ini fokus pada variabel inti yang diperlukan untuk RFM, klastering K-Means, dan perhitungan CLV.

Fase Data Preparation meliputi pembersihan data, transformasi variabel, dan normalisasi. Proses pembersihan mencakup:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

- Penanganan missing value dan duplikasi data,
- Penghapusan transaksi tidak valid yang dapat memedasikan analisis lebih lanjut.

Setelah pembersihan, variabel-variabel RFM dibentuk melalui konfigurasi data transaksi yang ada. Normalisasi dilakukan menggunakan Min-Max Scaling dengan rumus $X_{norm} = (X - X_{min}) / (X_{max} - X_{min})$ untuk menyamakan skala atribut RFM sebelum dilakukan clustering [19], [23], [24]. Rumus Min-Max yang diberikan diinstruksikan juga untuk diterapkan pada data relatif RFM dalam tahapan pra-pemodelan.

2.1. Analisis RFM

Analisis perilaku pelanggan dilakukan dengan memanfaatkan model RFM (Recency, Frequency, Monetary). Rumus-rumus yang digunakan adalah sebagai berikut:

$$R = T_{analisis} - T_{terakhir}$$

$$F = \sum Invoice\ Unique$$

$$M = \sum (Quantity \times UnitPrice)$$

- Recency (R) diinterpretasi sebagai selisih tanggal analisis dengan tanggal transaksi terakhir pelanggan: $R = T_{analisis} - T_{terakhir}$.
- Frequency (F) dihitung sebagai jumlah transaksi unik per pelanggan; secara eksplisit $F =$

sum of Invoices; Unique.

- Monetary (M) dihitung sebagai total nilai transaksi pelanggan: $M = \text{sum} (\text{Quantity} \times \text{UnitPrice})$ [20], [23]. RFM ini kemudian menjadi dasar untuk segmentasi menggunakan K-Means, serta sebagai input untuk perhitungan CLV di tingkat segmen, sebagaimana praktik umum dalam literatur RFM-K-Means [19], [20], [25].

2.2. K-Means Clustering

Untuk segmentasi pelanggan dilakukan klustering dengan algoritma K-Means. Penentuan jumlah kluster optimal dilakukan melalui Elbow Method, dengan mengevaluasi SSE (Sum of Squared Errors) sebagai fungsi kualitas kluster. Rumus Euclidean Distance yang dipakai untuk jarak antar titik data dan centroid adalah:

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

dan fungsi objektif SSE:

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

K-Means dipilih karena kemampuannya dalam membentuk kluster berbasis atribut numerik (R, F, M) yang telah dinormalisasi, serta adanya basis empiris yang banyak dipakai dalam studi segmentasi pelanggan berbasis RFM dan CLV [19], [20], [21].

2.3. Customer Lifetime Value (CLV)

Nilai CLV dihitung sebagai ukuran ekonomi rata-rata pelanggan berdasarkan perilaku historisnya. Rumus yang diadopsi adalah:

$$CLV = \text{Average Transaction Value} \times \text{Purchase Frequency} \times \text{Customer Lifespan}$$

- $CLV = \text{Average Transaction Value} \times \text{Purchase Frequency} \times \text{Customer Lifespan}$ Rumus ini selaras dengan formula CLV yang banyak digunakan dalam literatur praktisi CRM dan CRM berbasis RFM-K-Means, untuk memberikan ranking prioritas pada segmen pelanggan dan merumuskan strategi pemasaran yang tepat bagi setiap segmen [19], [21], [24], [26].

2.4. Evaluasi dan Strategi Pemasaran

Hasil clustering dianalisis dengan memanfaatkan profil RFM yang terdefinisi untuk setiap cluster, serta nilai CLV yang dihitung per cluster. Analisis ini bertujuan mengidentifikasi karakteristik perilaku pelanggan tiap segmen dan mengaitkannya dengan potensi nilai ekonomis (CLV).

Dengan demikian, rekomendasi strategi pemasaran yang spesifik untuk setiap segmen dapat dirumuskan, misalnya strategi retensi untuk klaster High-Value Loyal, up-selling untuk segmen Potential Loyalists, atau reaktivasi untuk segmen Lost Customers, sesuai praktik umum dalam literatur CLV-RFM-K-Means dan studi-studi serupa [19], [22], [26], [27]. Hasil evaluasi ini akan menjadi dasar untuk rekomendasi operasional bagi tim pemasaran e-commerce di Indonesia, termasuk alokasi sumber daya, penentuan tawaran promosi, dan penyesuaian pengalaman pelanggan.

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Hasil

4.1.1. Analisis RFM

Tahap awal pengolahan data menghasilkan nilai deskriptif umum untuk variabel transaksional. Berdasarkan analisis terhadap total keseluruhan sampel pelanggan, didapatkan pemahaman awal mengenai pola distribusi dan tendensi sentral data transaksi ritel online sebelum dilakukan pengelompokan menggunakan algoritma clustering.

Tabel 2. Statistik Deskriptif Variabel RFM

Variabel	Rata-rata
Recency	92,54
Frequency	4,27
Monetary	2.048,69
CLV	6.146,06

Berdasarkan Tabel 2, terlihat bahwa rata-rata jarak transaksi terakhir pelanggan (Recency) adalah 92,54 hari, yang mengindikasikan adanya jeda waktu yang cukup signifikan sejak aktivitas belanja terakhir dari sebagian besar populasi. Rata-rata frekuensi pembelian berada pada angka 4,27 kali transaksi, sementara akumulasi nilai transaksi (Monetary) mencatatkan nilai rata-rata sebesar 2.048,69 per pelanggan. Nilai baseline ekonomi dasar ini (CLV rata-rata sebesar 6.146,06) kemudian akan ditelaah lebih dalam pada variasi performa antarkluster.

4.1.2. Hasil K-Means Clustering

Hasil segmentasi menggunakan algoritma K-Means Clustering berhasil mengelompokkan pelanggan ke dalam lima cluster dengan karakteristik perilaku transaksi yang berbeda berdasarkan nilai Recency, Frequency, dan Monetary (RFM). Cluster 1 memiliki nilai Recency tertinggi (248,47 hari) serta nilai Frequency (1,55) dan Monetary (478,65) terendah, sehingga dapat dikategorikan sebagai pelanggan tidak aktif (Hibernating Customers) yang berisiko tinggi mengalami churn. Sebaliknya, Cluster 2 menunjukkan tingkat aktivitas transaksi tertinggi dengan Frequency sebesar 120,50 kali dan Recency hanya 6,50 hari, sehingga termasuk segmen

Champions yang sangat loyal dan aktif. Cluster 4 memiliki nilai Monetary tertinggi sebesar 190.808,54 dengan Recency yang rendah (7,67 hari), menunjukkan bahwa segmen ini merupakan VIP Customers yang memberikan kontribusi finansial terbesar bagi perusahaan. Sementara itu, Cluster 3 terdiri dari pelanggan loyal dengan aktivitas dan nilai transaksi yang stabil, sedangkan Cluster 0 merupakan Potential Customers yang masih memiliki peluang untuk dikembangkan menjadi pelanggan yang lebih loyal. Hasil ini menunjukkan bahwa K-Means mampu mengidentifikasi heterogenitas perilaku pelanggan secara efektif sehingga dapat digunakan sebagai dasar penyusunan strategi pemasaran yang lebih tepat sasaran dan berbasis nilai pelanggan.

Cluster	Recency	Frequency	Monetary
0	43,92	3,65	1.333,13
1	248,47	1,55	478,65
2	6,50	120,50	55.099,49
3	15,67	21,29	12.813,94
4	7,67	42,83	190.808,54

Tabel 3. Profil Cluster Pelanggan

Karakteristik antarkluster menunjukkan variasi ekstrem. Cluster 1 ditandai dengan nilai Recency yang sangat tinggi (248,47 hari) dengan aktivitas minimal, sementara Cluster 2 menunjukkan interaksi yang sangat intens dengan rata-rata frekuensi mencapai 120,50 kali pembelian. Cluster 4 menonjol sebagai segmen dengan nilai pengeluaran finansial yang luar biasa besar (Monetary 190.808,54) meskipun nilai frekuensinya tidak setinggi Cluster 2.

4.1.3. Segmentasi Pelanggan

Berdasarkan karakteristik orientasi nilai transaksi pada tabel profil kluster, dilakukan pemetaan nama segmen bisnis konkrit untuk mempermudah perumusan keputusan taktis.

Cluster	Segmen
0	Potential Customers
1	Hibernating Customers
2	Champions
3	Loyal Customers
4	VIP Customers

Tabel 4. Nama dan Karakteristik Segmen

Setiap kelompok didefinisikan secara unik: Cluster 0 diidentifikasi sebagai Potential Customers karena memiliki retensi transaksi menengah. Cluster 1 merupakan Hibernating Customers yang berisiko tinggi mengalami churn. Cluster 2 dikukuhkan sebagai Champions akibat tingginya loyalitas dan keaktifan. Cluster 3 dipetakan sebagai Loyal Customers yang bertransaksi stabil,

sedangkan Cluster 4 adalah VIP Customers yang menyumbangkan nilai margin keuntungan individual tertinggi bagi korporasi.

4.1.4. Analisis Customer Lifetime Value

Tahapan evaluasi selanjutnya adalah mengukur performa ekonomi jangka panjang riil dari masing-masing segmen pelanggan e-commerce melalui perhitungan nilai Customer Lifetime Value (CLV).

Segmen	CLV
VIP Customers	572.425,61
Champions	165.298,47
Loyal Customers	38.441,82
Potential Customers	3.999,40
Hibernating Customers	1.435,94

Tabel 5. Nilai CLV Setiap Segmen

Hasil menunjukkan disparitas nilai ekonomi yang sangat mencolok antar segmen. Segmen VIP Customers menduduki peringkat teratas dengan raihan nilai rata-rata CLV individu yang luar biasa tinggi sebesar 572.425,61, disusul oleh Champions sebesar 165.298,47. Hal ini membuktikan bahwa kelompok pelanggan VIP memberikan kontribusi nilai jangka panjang terbesar per kapita bagi keberlangsungan bisnis perusahaan ritel online tersebut.

4.1.5. Kontribusi Pendapatan Setiap Segmen

Selain melihat nilai rata-rata per individu melalui CLV, analisis total kontribusi pendapatan akumulatif dari masing-masing segmen perlu dilakukan untuk melihat gambaran makro performa bisnis.

Segmen	Revenue (%)
Potential Customers	45,72
Loyal Customers	30,71
VIP Customers	12,88
Hibernating Customers	5,73
Champions	4,96

Tabel 6. Kontribusi Revenue per Segmen

Terjadi fenomena menarik di mana segmen Potential Customers yang memiliki nilai CLV per individu relatif rendah justru memberikan kontribusi akumulatif pendapatan terbesar bagi perusahaan, yaitu mencapai 45,72%. Hal ini dipicu oleh dominasi volume populasi dari segmen tersebut yang sangat masif dibandingkan kelompok VIP atau Champions yang jumlah anggotanya

sangat sedikit.

4.1.6. Strategi Pemasaran Berdasarkan Segmentasi

Langkah akhir dari kerangka kerja CRISP-DM adalah penyusunan rekomendasi bauran program aksi pemasaran yang disesuaikan secara personal (personalized marketing) berdasarkan karakteristik perilaku unik masing-masing segmen.

Segmen	Strategi
VIP	VIP Membership, Priority Service
Champions	Loyalty Program, Referral Program
Loyal	Upselling, Bundling
Potential	Cross Selling, Email Marketing
Hibernating	Reactivation Promotion

Tabel 7. Strategi Pemasaran per Segmen

4.1.7. Implikasi Strategi Pemasaran Berbasis Segmen

Hasil integrasi pemodelan RFM, K-Means Clustering, dan CLV memberikan implikasi praktis berupa panduan penyusunan strategi pemasaran berbasis data yang presisi dan personal untuk masing-masing kluster:

1. **VIP Customers:** Strategi diprioritaskan pada program retensi eksklusif seperti pemberian VIP Membership, layanan prioritas (Priority Service), serta penawaran produk yang dipersonalisasi secara khusus demi menjaga loyalitas jangka panjang.
2. **Champions:** Rekomendasi strateginya adalah melibatkan mereka dalam Loyalty Program yang agresif, pemberian Reward Point, serta Referral Program untuk memanfaatkan loyalitas kuat mereka sebagai promotor merek.
3. **Loyal Customers:** Perusahaan dapat mengoptimalkan nilai transaksi segmen ini melalui strategi penjualan Upselling dan pemberian rekomendasi produk otomatis (Product Recommendation) yang relevan.
4. **Potential Customers:** Mengingat jumlahnya yang sangat besar, segmen ini perlu didorong menjadi pelanggan loyal menggunakan taktik promosi silang (Cross Selling) dan Personalized Marketing agar frekuensi belanja mereka meningkat.
5. **Hibernating Customers:** Fokus utama pada segmen ini adalah reaktivasi melalui Retention Campaign, pemberian kupon Cashback Program, dan promosi khusus via email marketing guna menekan angka churn rate.

Secara keseluruhan, integrasi metode ini berhasil membuktikan bahwa pengambilan keputusan berbasis data mining mampu mengoptimalkan penggunaan sumber daya pemasaran secara jauh lebih efisien dibandingkan pendekatan konvensional.

4.2. Pembahasan

Hasil analisis RFM menunjukkan bahwa perilaku transaksi pelanggan e-commerce memiliki karakteristik yang beragam. Nilai rata-rata Recency sebesar 92,54 hari menunjukkan bahwa sebagian besar pelanggan tidak melakukan transaksi dalam waktu yang relatif dekat, sedangkan nilai Frequency sebesar 4,27 kali transaksi dan Monetary sebesar 2.048,69 menunjukkan tingkat aktivitas dan kontribusi transaksi yang bervariasi antar pelanggan. Selanjutnya, penerapan algoritma K-Means Clustering berhasil mengelompokkan 4.338 pelanggan ke dalam lima segmen utama, yaitu Potential Customers, Hibernating Customers, Champions, Loyal Customers, dan VIP Customers. Setiap segmen memiliki karakteristik yang berbeda berdasarkan kombinasi nilai Recency, Frequency, dan Monetary sehingga memberikan gambaran yang lebih jelas mengenai pola perilaku pelanggan.

Hasil perhitungan Customer Lifetime Value (CLV) menunjukkan adanya perbedaan nilai ekonomi yang signifikan antar segmen pelanggan. Segmen VIP Customers memiliki nilai CLV tertinggi sebesar 572.425,61, diikuti oleh Champions sebesar 165.298,47 dan Loyal Customers sebesar 38.441,82. Temuan ini menunjukkan bahwa pelanggan VIP memberikan kontribusi nilai jangka panjang terbesar bagi perusahaan sehingga perlu menjadi prioritas utama dalam strategi retensi pelanggan. Namun demikian, analisis kontribusi pendapatan menunjukkan fenomena yang menarik, di mana segmen Potential Customers justru memberikan kontribusi revenue terbesar sebesar 45,72% karena jumlah anggotanya yang jauh lebih banyak dibandingkan segmen lainnya. Kondisi ini menunjukkan bahwa nilai pelanggan secara individual tidak selalu sejalan dengan kontribusi pendapatan secara agregat.

Integrasi metode RFM, K-Means Clustering, dan CLV terbukti mampu menghasilkan segmentasi pelanggan yang lebih komprehensif dibandingkan penggunaan satu metode secara terpisah. Informasi yang diperoleh tidak hanya menggambarkan perilaku transaksi pelanggan, tetapi juga menunjukkan nilai ekonomi yang dihasilkan oleh setiap segmen. Berdasarkan hasil tersebut, perusahaan e-commerce dapat menerapkan strategi pemasaran yang lebih tepat sasaran, seperti program VIP Membership dan layanan prioritas untuk VIP Customers, loyalty program untuk Champions, strategi upselling dan bundling untuk Loyal Customers, cross-selling untuk Potential Customers, serta program reaktivasi bagi Hibernating Customers. Dengan demikian, penerapan segmentasi pelanggan berbasis data mining dapat meningkatkan efektivitas pemasaran, efisiensi alokasi anggaran promosi, serta mendukung peningkatan profitabilitas dan keunggulan bersaing perusahaan e-commerce.

KESIMPULAN

Penelitian ini berhasil mengintegrasikan analisis RFM, algoritma K-Means Clustering, dan penilaian Customer Lifetime Value (CLV) untuk merancang segmentasi pelanggan berbasis data pada industri e-commerce. Berdasarkan analisis terhadap 4.338 pelanggan, ditarik kesimpulan

bahwa perilaku transaksi pelanggan e-commerce bersifat sangat heterogen, namun berhasil dikelompokkan secara optimal menjadi lima segmen homogen, yaitu: Potential Customers, Hibernating Customers, Champions, Loyal Customers, dan VIP Customers.

Analisis finansial menunjukkan bahwa segmen VIP Customers memegang nilai rata-rata CLV tertinggi sebesar 572.425,61, sementara kontribusi akumulasi pendapatan terbesar justru disumbangkan oleh segmen Potential Customers (45,72%) karena faktor volume populasi yang dominan. Penemuan ini membuktikan bahwa integrasi model data mining ini mampu menjadi instrumen pengambilan keputusan pemasaran yang efektif.

Berdasarkan hasil tersebut, disarankan bagi perusahaan e-commerce untuk mengabaikan strategi pemasaran massal dan mulai menerapkan program pemasaran terdiferensiasi (seperti pemberian layanan eksklusif untuk VIP dan program reaktivasi cashback untuk pelanggan hibernating) demi meningkatkan efisiensi biaya serta profitabilitas jangka panjang. Untuk penelitian selanjutnya, disarankan mengembangkan model prediktif CLV masa depan dengan membandingkan algoritma clustering lain seperti DBSCAN atau Hierarchical Clustering.

DAFTAR RUJUKAN

- [1] M. Ray and B. K. Mangaraj, "AHP Based Data Mining for Customer Segmentation Based on Customer Lifetime Value," *Int. J. Data Min. Tech. Appl.*, vol. 5, no. 1, pp. 28–34, 2016, doi: 10.20894/ijdmata.102.005.001.007.
- [2] D. W. Maulida and M. Dachyar, "Customer Loyalty Strategy Using Customer Lifetime Value (A Case Study of Jamu and Herbal Products SME)", doi: 10.46254/an12.20220195.
- [3] Y. Shih and C.-Y. Liu, "A Method for Customer Lifetime Value Ranking — Combining the Analytic Hierarchy Process and Clustering Analysis," *J. Database Mark. Cust. Strategy Manag.*, vol. 11, no. 2, pp. 159–172, 2003, doi: 10.1057/palgrave.dbm.3240216.
- [4] G. M. Díaz and D. Gómez, "RFID: A Fuzzy Linguistic Model to Manage Customers From the Perspective of Their Interactions With the Contact Center," *Mathematics*, vol. 9, no. 19, p. 2362, 2021, doi: 10.3390/math9192362.
- [5] F. Safari, N. Safari, and G. A. Montazer, "Customer Lifetime Value Determination Based on RFM Model," *Mark. Intell. Plan.*, vol. 34, no. 4, pp. 446–461, 2016, doi: 10.1108/mip-03-2015-0060.
- [6] B. C. Laksono and I. Y. Wulansari, "Estimating Customer Lifetime Value in the E-Commerce Industry Using Multivariate Analysis," *Proc. Int. Conf. Data Sci. Off. Stat.*, vol. 2021, no. 1, pp. 507–518, 2022, doi: 10.34123/icdsos.v2021i1.161.
- [7] C. Hsu, Y. Chang, T. Chen, T. Chang, and Y. Lin, "Who Donates on Line? Segmentation Analysis and Marketing Strategies Based on Machine Learning for Online Charitable Donations in Taiwan," *Ieee Access*, vol. 9, pp. 52728–52740, 2021, doi: 10.1109/access.2021.3066713.

- [8] C. Yan, H. Sun, W. Liu, and J. Chen, “An Integrated Method Based on Hesitant Fuzzy Theory and RFM Model to Insurance Customers’ Segmentation and Lifetime Value Determination,” *J. Intell. Fuzzy Syst.*, vol. 35, no. 1, pp. 159–169, 2018, doi: 10.3233/jifs-169577.
- [9] B. C. Laksono and I. Y. Wulansari, “Pemodelan Dan Penerapan Metode RFM Pada Estimasi Nilai Konsumen (Customer Lifetime Value) Menggunakan K-Means Clustering Machine Learning,” *Semin. Nas. Off. Stat.*, vol. 2020, no. 1, pp. 1277–1285, 2021, doi: 10.34123/semnasoffstat.v2020i1.689.
- [10] M. A. A. Kazemi, N. Estemdad, and A. Poorebrahimi, “Customer Behavior in Banking Industry: Comparison of Data Mining Techniques,” *J. Adv. Manag. Sci.*, pp. 13–16, 2015, doi: 10.12720/joams.3.1.13-16.
- [11] G. ASLANTAŞ, M. GENÇGÜL, M. RUMELLİ, M. ÖZSARAÇ, and G. BAKIRLI, “Customer Segmentation Using K-Means Clustering Algorithm and RFM Model,” *Dokuz Eylül Üniversitesi Mühendis. Fakültesi Fen Ve Mühendis. Derg.*, vol. 25, no. 74, pp. 491–503, 2023, doi: 10.21205/deufmd.2023257418.
- [12] D. Aslan and M. Tolon, “Comparing Customer Segmentation With CLV Using Data Mining and Statistics: A Case Study,” *J. Bus. Res. - Turk*, vol. 10, no. 4, pp. 887–900, 2018, doi: 10.20491/isarder.2018.554.
- [13] R. Riyanto, “Modelling Customers Lifetime Value for Non-Contractual Business,” *Ijiis Int. J. Inform. Inf. Syst.*, vol. 4, no. 1, pp. 55–62, 2021, doi: 10.47738/ijiis.v4i1.77.
- [14] S. M. H. Hasheminejad and M. Khorrami, “Clustering of Bank Customers Based on Lifetime Value Using Data Mining Methods,” *Intell. Decis. Technol.*, vol. 14, no. 4, pp. 507–515, 2021, doi: 10.3233/idt-190176.
- [15] K. Kumaran, P. Pathak, R. Haque, and P. Stynes, “A Machine And Deep Learning Framework To Retain Customers Based On Their Lifetime Value,” pp. 91–103, 2022, doi: 10.1007/978-3-031-24094-2_6.
- [16] Fitri Marisa, S. S. S. Ahmad, Z. I. M. Yusof, F. Fachrudin, and T. M. Akhriza, “Segmentation Model of Customer Lifetime Value in Small and Medium Enterprise (SMEs) Using K-Means Clustering and LRFM Model,” *Int. J. Integr. Eng.*, vol. 11, no. 3, 2019, doi: 10.30880/ijie.2019.11.03.018.
- [17] N. A. Ejgerdi and M. Kazerooni, “A Stacked Ensemble Learning Method for Customer Lifetime Value Prediction,” *Kybernetes*, vol. 53, no. 7, pp. 2342–2360, 2023, doi: 10.1108/k-12-2022-1676.
- [18] S. Chavhan, R. C. Dharmik, S. Jain, and K. Kamble, “RFM Analysis for Customer Segmentation Using Machine Learning: A Survey of a Decade of Research,” *3c Tic Cuad. Desarro. Apl. Las Tic*, vol. 11, no. 2, pp. 166–173, 2022, doi: 10.17993/3ctic.2022.112.166-173.
- [19] E. F. L. Awalina and W. I. Rahayu, “Optimalisasi Strategi Pemasaran Dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering Pada Transaksi Online Retail,” *J. Teknol. Dan Inf.*, vol. 13, no. 2, pp. 122–137, 2023, doi: 10.34010/jati.v13i2.10090.

- [20] R. Siagian, P. S. P. Sirait, and A. Halima, “E-Commerce Customer Segmentation Using K-Means Algorithm and Length, Recency, Frequency, Monetary Model,” *J. Inform. Telecommun. Eng.*, vol. 5, no. 1, pp. 21–30, 2021, doi: 10.31289/jite.v5i1.5182.
- [21] T. A. Afifah, R. Novita, T. K. Ahsyar, and Z. Zarnelly, “Penerapan Algoritma K-Means Menggunakan Model LRFM Dalam Klasterisasi Nilai Hidup Pelanggan,” *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 1010, 2024, doi: 10.30865/mib.v8i2.7605.
- [22] A. P. Pramudiansyah, “Segmentasi Pelanggan Menggunakan Algoritma K-Means Berdasarkan Model Recency Frequency Monetary,” *J. Ilm. Ilmu Komput.*, vol. 7, no. 2, pp. 06–19, 2021, doi: 10.35329/jiik.v7i2.201.
- [23] I. Maskanah, “Segmentasi Pelanggan Toko Purnama Dengan Algoritma K-Means Dan Model RFM Untuk Perancangan Strategi Pemasaran,” *Inovtek Polbeng - Seri Inform.*, vol. 5, no. 2, p. 218, 2020, doi: 10.35314/isi.v5i2.1443.
- [24] A. A. D. Sulistyawati and M. Sadikin, “Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan,” *Sistemasi*, vol. 10, no. 3, p. 516, 2021, doi: 10.32520/stmsi.v10i3.1332.
- [25] F. Juniati, S. Monalisa, R. Zafa, and M. Muslim, “KLAUSTERISASI CUSTOMER LIFETIME VALUE DENGAN MODEL LRFM MENGGUNAKAN ALGORITMA Fuzzy C-Means,” *J. Ilm. Rekayasa Dan Manaj. Sist. Inf.*, vol. 6, no. 1, p. 38, 2020, doi: 10.24014/rmsi.v6i1.8635.
- [26] B. T. Kristanti, A. Junaidi, and E. P. Mandyartha, “Implementasi K-Means Clustering Dalam Segmentasi Pelanggan Berdasarkan Usia, Pendapatan, Dan Model RFM (Studi Kasus: Lantikya Store Jombang),” *J. Inform. Dan Tek. Elektro Terap.*, vol. 12, no. 3, 2024, doi: 10.23960/jitet.v12i3.4677.
- [27] T. L. Afandi, B. Warsito, and R. Santoso, “Implementasi K-Medoids Dan Model Weighted-Length Recency Frequency Monetary (W-Lrfm) Untuk Segmentasi Pelanggan Dilengkapi Gui R,” *J. Gaussian*, vol. 11, no. 3, pp. 429–438, 2022, doi: 10.14710/j.gauss.11.3.429-438.