

CUSTOMER LOYALTY SEGMENTATION IN ONLINE STORE USING LRFM AND MLRFM IN COMBINATION WITH RM K-MEANS ALGORITHM

Angelina Caroline Utomo¹, Andreas Handoyo^{*2}, Tanti Octavia³

^{1,2}Business Information System, Industrial Technology Faculty, Universitas Kristen Petra, Indonesia

³Industrial Engineering, Industrial Technology Faculty, Universitas Kristen Petra, Indonesia

Email: ¹c14190219@john.petra.ac.id, ²handoyo@petra.ac.id, ³tanti@petra.ac.id

(Article received: October 31, 2023; Revision: November 27, 2023; published: April 15, 2024)

Abstract

The rapid development of online business in recent years has driven Store X to embark on a digital transformation. By the end of 2020, Store X relocate their conventional business to online business. The greatest obstacle and key to success for online business operators, such as Store X, is gaining and retaining consumer loyalty in the face of an increasing number of competitors. Therefore, the company must be able to identify the character (behavior) of its clients to provide appropriate treatment. Each customer's behavior is unique, which means they must all be treated differently. However, all this time, Online Store X has provided the same treatment (as much of a discount) to all its customers due to the lack of information regarding their customers' characteristics. Therefore, in this study, customers of Online Store X were segmented based on their transactional behavior using online transaction history data from March 2021 to March 2023. Two customer analysis models, LRFM and MLRFM, will be combined with RM K-Means to find the best combination through Silhouette Coefficient values. The optimal number of clusters (k) is then determined using the Elbow Method. The results indicate that the optimal number of clusters for both combinations is K=3, with the combination of MLRFM and RM K-Means is the best combination. The finest combination has a silhouette coefficient value of 0.8609. Based on this combination, it is also known that 2,053 customers in cluster 3 are loyal customers, while 2,339 customers in cluster 1 and 2 are lost customers. The results of this study were also implemented on websites built for X Store using Python programming languages and MySQL databases, making it easier for companies to see data visualization.

Keywords: customer segmentation, LRFM, MLRFM, RM k-means algorithm.

SEGMENTASI LOYALITAS PELANGGAN MENGGUNAKAN LRFM DAN MLRFM DENGAN KOMBINASI ALGORITMA RM K-MEANS PADA TOKO ONLINE

Abstrak

Pesatnya perkembangan bisnis *online* dalam beberapa tahun terakhir mendorong Toko X untuk melakukan transformasi digital. Toko X mengalihkan bisnis konvensional menjadi bisnis *online* pada akhir 2020. Tantangan terbesar dan kunci kesuksesan bagi pelaku bisnis *online*, seperti Toko X adalah memperoleh dan mempertahankan loyalitas pelanggan dari banyaknya pesaing yang bermunculan. Sehingga perusahaan harus mampu mengenali karakter (perilaku) pelanggannya agar dapat memberikan perlakuan yang sesuai. Setiap pelanggan memiliki perilaku yang berbeda sehingga perlu diberikan perlakuan yang berbeda pula. Namun, selama ini Toko *Online* X memberikan perlakuan (besaran diskon) yang sama kepada seluruh pelanggannya karena belum memiliki informasi terkait karakteristik pelanggannya. Oleh karena itu, pada penelitian ini dilakukan segmentasi pelanggan untuk mengelompokkan pelanggan Toko *Online* X berdasarkan kesamaan karakteristiknya melalui data riwayat transaksi *online* periode Maret 2021 hingga Maret 2023. Dua model analisa pelanggan, LRFM dan MLRFM, akan dikombinasikan dengan RM K-Means untuk mengetahui kombinasi yang terbaik melalui nilai *silhouette coefficient*. Kemudian, metode *elbow* digunakan untuk menentukan jumlah *cluster* (k) optimal. Hasil penelitian menunjukkan bahwa jumlah *cluster* optimal dari kedua kombinasi berada K=3 dengan kombinasi MLRFM dan RM K-Means adalah kombinasi terbaik. Nilai *silhouette coefficient* dari kombinasi terbaik sebesar 0,8609. Berdasarkan kombinasi tersebut, juga diketahui bahwa sebanyak 2,053 pelanggan adalah pelanggan loyal (*loyal customers*) berada pada *cluster* 3. Selanjutnya, pelanggan yang berada di *cluster* 1 dan 2 dikategorikan sebagai *lost customer* dengan beranggotakan 2,339 pelanggan. Hasil penelitian juga diimplementasikan pada *website* yang dibangun untuk Toko X menggunakan bahasa pemrograman *Python* dan *database* MySQL sehingga memudahkan perusahaan dalam melihat visualisasi data.

Kata kunci: algoritma RM K-Means, LRFM, MLRFM, segmentasi pelanggan.

1. PENDAHULUAN

Pesatnya perkembangan teknologi informasi yang terjadi pada sektor bisnis saat ini memberikan berbagai kemudahan. Pelanggan dapat dengan mudah membeli sesuatu melalui internet tanpa harus datang ke toko [1]. Akibatnya, banyak pelanggan yang beralih untuk belanja *online* [2]. Bank Indonesia (BI) menyatakan nilai transaksi *e-commerce* di Indonesia pada 2022 meningkat 18,8% dari tahun sebelumnya, yaitu sebesar 476,3 triliun rupiah [3]. Secara tidak langsung, para pelaku bisnis konvensional dituntut untuk melakukan transformasi digital agar dapat bertahan di era ini. Toko X merupakan perusahaan yang menjual berbagai macam sepatu dan sandal produksi lokal sejak 2005. Sejalan dengan maraknya bisnis *online*, Toko X mentransformasikan bisnisnya secara *online* di beberapa platform *e-commerce* pada akhir 2020.

Sebagai pelaku bisnis *online*, mendapatkan dan mempertahankan loyalitas pelanggan merupakan kunci keberhasilan sekaligus tantangan terbesar bagi Toko X [4], [5]. Toko *Online* X harus mampu memahami perbedaan karakteristik pelanggannya agar dapat mengenali pelanggan potensial hingga menentukan strategi pemasaran yang sesuai. Hal ini penting karena setiap pelanggan perlu mendapatkan perlakuan yang berbeda karena memiliki kebutuhan yang berbeda pula [6]. Selain itu, pemberian perlakuan yang sama kepada semua pelanggan menyebabkan pelanggan yang tidak potensial menghilangkan nilai daripada menghasilkan nilai bagi perusahaan [7]. Hingga saat ini, tim penjualan Toko *Online* X masih memberikan perlakuan (dalam hal ini besaran diskon) yang sama kepada seluruh pelanggannya, baik yang jarang ataupun sering bertransaksi dalam nominal kecil hingga besar. Ini terjadi karena tim penjualan Toko *Online* X tidak memiliki informasi dan belum memahami bagaimana pola pelanggannya dalam bertransaksi. Oleh karena itu, segmentasi pelanggan diperlukan untuk mengetahui segmen pelanggan potensial dengan cara mengelompokkan pelanggan yang memiliki kesamaan pola bertransaksi [8], [9].

Pada penelitian sebelumnya, segmentasi pelanggan dilakukan pada perusahaan aksesoris otomotif untuk mengenali pelanggan potensial menggunakan metode RFM dan K-Means tradisional. Metode wawancara digunakan dalam penelitian tersebut untuk menguji validitas hasil *cluster*. Penelitian tersebut menghasilkan 4 *cluster* pelanggan dengan tipe *ordinary consumers*, *big company*, dan *top class customers* [10]. Selanjutnya terdapat penelitian yang mengelompokkan pelanggan *e-commerce* berdasarkan kesamaannya menggunakan kombinasi LRFM dan K-Means tradisional. Sebanyak 3 *cluster* terbentuk dan hasil *cluster* dianalisis menggunakan LRFM *index* sehingga

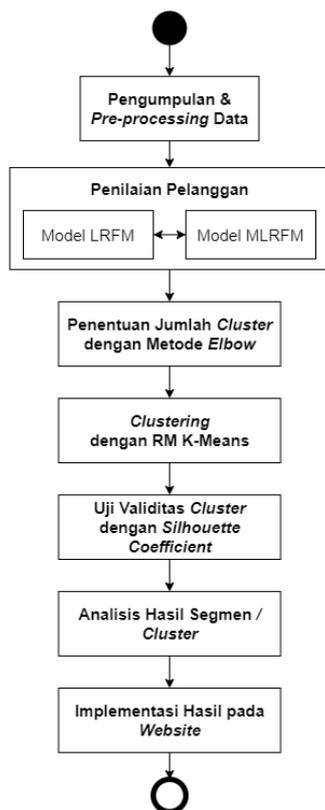
didapatkan kelompok *new customers*, *lost customer*, dan *core customer* [11]. Penelitian lainnya juga menggunakan metode K-Means tradisional untuk menentukan kelompok siswa yang mengalami hambatan belajar. Jumlah *cluster* ditentukan secara manual oleh peneliti dengan jumlah 2 *cluster*, di mana terdiri dari *cluster* tinggi dan rendah [12].

Berangkat dari penelitian sebelumnya, metode-metode pengembangan yang berkaitan dengan proses segmentasi pelanggan ditemukan. Model LRFM dan MLRFM merupakan contoh pengembangan dari kelemahan model analisis pelanggan RFM. Sedangkan, algoritma RM K-Means adalah pengembangan dari metode K-Means tradisional. Oleh karena itu, penelitian ini akan membandingkan kombinasi model penilaian pelanggan LRFM dan MLRFM dengan algoritma RM K-Means menggunakan nilai *silhouette coefficient*. Model *Length, Recency, Frequency, Monetary* (LRFM) dinilai lebih efektif dalam menilai umur hidup (loyalitas) pelanggan melalui variabel baru, yaitu *Length (L)* [13]. Sedangkan, model *Multi Layer Recency, Frequency, Monetary* (MLRFM) menjawab ketidakadilan model RFM dalam menilai *frequency* dan *monetary* pelanggan ketika pelanggan lama mengurangi atau bahkan tidak lagi bertransaksi dan munculnya pelanggan baru yang aktif bertransaksi, di mana fenomena ini biasanya terjadi dalam transaksi *online* [14]. Selanjutnya, *Repetitive Median* K-Means (RM K-Means) merupakan metode *clustering* untuk membentuk segmen pelanggan. Metode pengembangan dari K-Means tradisional ini mampu mengurangi iterasi dengan penentuan *centroid* awal berdasarkan perhitungan median [15].

Penelitian ini bertujuan untuk membantu Toko *Online* X dalam mengetahui kelompok mana yang merupakan pelanggan loyal dan potensial berdasarkan kombinasi metode terbaik. Sehingga hasil segmentasi nantinya dapat dijadikan sebagai acuan oleh tim penjualan untuk menetapkan strategi pemasaran yang sesuai guna meningkatkan loyalitas, tingkat *repurchase*, dan pendapatan perusahaan kedepannya.

2. METODE PENELITIAN

Tahapan yang dilakukan dalam penelitian ini dimulai dari pengumpulan dan *pre-processing* data, analisis pelanggan menggunakan model LRFM dan MLRFM, penentuan jumlah *cluster* yang optimal, segmentasi dengan RM K-Means, pengujian validitas terhadap hasil *cluster*, penentuan tipe segmen pelanggan, dan implementasi hasil ke dalam aplikasi berbasis *website*. *Flowchart* pada Gambar 1 menunjukkan tahapan yang telah disebutkan.



Gambar 1. Flowchart Metodologi Penelitian.

2.1. Pengumpulan dan Pre-processing Data

Data dikumpulkan melalui observasi dan wawancara bersama pemilik dan kepala tim penjualan Toko Online X. Data kemudian akan dibersihkan dan dipilah sebelum diproses. Pembersihan data dilakukan dengan menghilangkan baris yang tidak lengkap atau bernilai kosong. Kemudian, tahap pemilihan data merupakan proses memilah atribut (kolom) data sesuai yang diperlukan.

2.2. Length, Recency, Frequency, Monetary (LRFM)

Model ini dikembangkan oleh Chang dan Tsay untuk menyempurnakan model RFM sehingga lebih baik dalam mengidentifikasi lama hubungan (loyalitas) pelanggan dengan perusahaan [16]. Hal ini dilakukan dengan menambahkan variabel *Length* (L) untuk mengukur interval pembelian pertama dan terakhir pelanggan. Selanjutnya, atribut *Recency* (R) mengidentifikasi interval antara transaksi terakhir hingga dilakukannya penelitian, *Frequency* (F) menunjukkan berapa kali seorang pelanggan melakukan transaksi, dan *Monetary* (M) adalah total uang yang dibelanjakan dalam suatu periode waktu oleh pelanggan. Langkah-langkah analisis LRFM pelanggan adalah:

- Hitung ringkasan nilai LRFM pelanggan.
- Normalisasi data dengan metode *min-max* agar berada pada skala yang sama, terutama untuk data yang memiliki perbedaan cukup signifikan

[13]. Persamaan (1) merupakan formula untuk melakukan normalisasi *min-max*.

$$V' = \frac{(v_i - v_{min})}{v_{max} - v_{min}} \quad (1)$$

Pada persamaan (1), V' adalah variabel model LRFM, di mana v_i merupakan nilai dari data ke- i yang dinormalisasi, v_{min} merupakan nilai terkecil dan v_{max} adalah nilai terbesar dalam kumpulan data [9].

- Tentukan skor LRFM pelanggan berdasarkan daftar skor pada suatu rentang nilai normalisasi pada Tabel 1.

Tabel 1. Rentang Skor LRFM / MLRFM

Rentang Nilai Normalisasi	Skor LRFM / MLRFM
0,70 – 1	5
0,50 – 0,69	4
0,30 – 0,49	3
0,10 – 0,29	2
0 – 0,09	1

- Skor setiap variabel LRFM dikalikan dengan bobot yang diinginkan hingga didapatkan skor akhir variabel L, R, F, M.

2.3. Multi Layer Recency, Frequency, Monetary (LRFM)

Handojo *et al.* menciptakan MLRFM yang membagi penilaian *frequency* dan *monetary* dalam beberapa *layer* dan setiap *layer* diberi bobot sesuai keinginan perusahaan. MLRFM memberikan keadilan dalam menilai *frequency* dan *monetary* pelanggan jika digunakan pada pola transaksi *online*, di mana pelanggan lama dapat dengan mudah beralih dan pelanggan baru datang silih berganti dengan cepat. Sebagai contoh, ada pelanggan baru yang terus-menerus melakukan transaksi ke perusahaan selama 6 bulan terakhir. Sedangkan, pelanggan lama yang satu tahun lalu aktif bertransaksi, kini jarang bertransaksi bahkan menghilang atau pindah ke perusahaan lain, sehingga pelanggan baru mungkin berhak mendapat peringkat yang lebih tinggi [14]. Langkah-langkah analisis MLRFM pelanggan adalah:

- Tentukan *layer* periode dan bobot setiap periode.
- Hitung ringkasan nilai RFM pelanggan untuk setiap *layer* waktu.
- Normalisasi data dengan persamaan (1) yang telah disebutkan sebelumnya.
- Tentukan skor RFM berdasarkan skor rentang nilai normalisasi yang terdapat pada Tabel 1.
- Khusus untuk skor *frequency* dan *monetary*, perlu dikali dengan bobot masing-masing *layer* sehingga diperoleh skor R, total skor F dan M.
- Skor akhir MLRFM dihitung menggunakan persamaan (2) dengan R_i merupakan skor *recency* pelanggan ke- i yang dikalikan dengan bobot *recency* (WR_i). Sedangkan, $\sum_{i=1}^n F_i \times WF_{ij}$ adalah jumlah dari total skor *frequency*

pelanggan ke- i yang dikali dengan bobot *frequency* pada periode j . Kemudian, $\sum_{j=1}^n M_i \times WM_{ij}$ merupakan jumlah dari total skor *monetary* pelanggan ke- i yang dikali dengan bobot *monetary* dalam periode j .

$$MLRFM \text{ Poin} = (R_i \times WR_i) + (\sum_{j=1}^n F_i \times WF_{ij}) + (\sum_{j=1}^n M_i \times WM_{ij}) \quad (2)$$

2.4. Repetitive Median K-Means (RM K-Means)

RM K-Means adalah pengembangan dari algoritma K-Means tradisional dengan kelebihan mampu mengurangi waktu komputasi dan jumlah iterasi yang disebabkan oleh penentuan *centroid* awal secara acak. Pada metode ini, *centroid* awal ditetapkan berdasarkan nilai median setiap *array* dari variabel LRFM dan MLRFM [15]. Tahapan RM K-Means sebagai berikut:

- Urutkan skor R, F, M dan/atau L, R, F, M sebagai R', F', M' dan/atau L', R', F', M' dan simpan dalam *array*.
- Hitung nilai S. Caranya adalah membagi jumlah baris data dengan jumlah *cluster* (k).
- Bagi *array* R', F', M' dan/atau L', R', F', M' sebanyak k , dimana setiap *cluster* berjumlah S baris.
- Hitung median untuk setiap *cluster* ke- i sebagai *centroid* awal untuk K-Means.
- Hitung jarak RFM dan/atau LRFM setiap objek terhadap *centroid* awal menggunakan *Euclidean Distance* seperti pada persamaan (3)

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Pada persamaan (3), $d(x, y)$ merupakan jarak data ke- i dengan *centroid cluster* ke- j . Sedangkan, x_i adalah data ke- i dari variabel yang ada dan y_i adalah *centroid* dari *cluster* ke- j [17].

- Kelompokkan objek berdasarkan jarak minimum.
- Hitung ulang *centroid cluster* dengan menghitung nilai rata-rata setiap *cluster*.
- Ulangi langkah e hingga g sampai *centroid* tidak berubah pada iterasi yang dilakukan.

2.5. Metode Elbow

Jumlah *cluster* (k) optimal ditentukan menggunakan metode *elbow* dengan melihat selisih nilai *Sum of Squared Errors* (SSE) antara seluruh jumlah *cluster* yang nantinya membentuk sebuah siku pada grafik [18]. *Elbow* ditentukan berdasarkan turunan terbesar atas hasil perhitungan SSE. Persamaan (4) merupakan penulisan formula SSE. Dengan X_i adalah data atau objek ke- i . Kemudian, untuk C_k adalah *centroid* dari *cluster* ke- k [19].

$$SSE = \sum_{k=1}^k \sum_{x_i} \|X_i - C_k\|_2^2 \quad (4)$$

2.6. Silhouette Coefficient

Metode ini berdasarkan perhitungan kohesi (y_i) yang merupakan jarak rata-rata antara suatu data dengan data lainnya di kelompok yang sama dan separasi (x_i) atau jarak rata-rata antara suatu data dengan seluruh data lain di *cluster* berbeda [20] [21]. Formula untuk *silhouette coefficient* ditunjukkan oleh persamaan (5).

$$S = \frac{x-y}{\max(y, x)} \quad (5)$$

2.7. Tipe Segmen Pelanggan

Ha dan Park membandingkan rata-rata nilai variabel RFM untuk tiap-tiap *cluster* dengan nilai rata-rata RFM dari seluruh *cluster* untuk mengidentifikasi tipe segmen pelanggan [22]. Nilai rata-rata RFM suatu *cluster* yang lebih besar dari rata-rata keseluruhan ditandai dengan panah ke atas (\uparrow). Sebaliknya, panah ke bawah (\downarrow) menandakan nilai rata-rata *cluster* lebih kecil dari keseluruhan. Terdapat 4 (empat) tipe pelanggan dalam model RFM [23] seperti pada Tabel 2.

Tabel 2. Tipe Pelanggan Berdasarkan Model RFM

Simbol RFM	Tipe Pelanggan
R \uparrow F \uparrow M \uparrow	Loyal customers
R \uparrow F \downarrow M \uparrow	Promising customers
R \uparrow F \downarrow M \downarrow	New customers
R \downarrow F \downarrow M \downarrow	Lost customers
R \downarrow F \uparrow M \uparrow	

Dari Tabel 2 dapat diketahui bahwa tipe pelanggan berdasarkan model RFM terdiri dari *loyal customer*, *promising customer*, *new customer*, dan *lost customer*.

Tabel 3. Tipe Pelanggan Berdasarkan Model LRFM

Simbol LRFM	Tipe Pelanggan
L \uparrow R \uparrow F \uparrow M \uparrow	High-value loyal customers
L \downarrow R \uparrow F \uparrow M \uparrow	High-value new customers
L \uparrow R \downarrow F \uparrow M \uparrow	Potential loyal customers
L \downarrow R \downarrow F \uparrow M \uparrow	High-value lost customers
L \uparrow R \uparrow F \downarrow M \uparrow	Platinum customers
L \downarrow R \uparrow F \downarrow M \uparrow	Consuming promotional customers
L \uparrow R \downarrow F \downarrow M \uparrow	Potential consuming customers
L \downarrow R \downarrow F \downarrow M \uparrow	Consuming churn customers
L \uparrow R \uparrow F \uparrow M \downarrow	High frequency customers
L \downarrow R \uparrow F \uparrow M \downarrow	Frequency promotional customers
L \uparrow R \downarrow F \uparrow M \downarrow	Potential frequency customers
L \downarrow R \downarrow F \uparrow M \downarrow	Frequency churn customers
L \uparrow R \uparrow F \downarrow M \downarrow	Low-cost consuming customers
L \downarrow R \uparrow F \downarrow M \downarrow	Uncertain new customers
L \uparrow R \downarrow F \downarrow M \downarrow	High-cost consuming customers
L \downarrow R \downarrow F \downarrow M \downarrow	Uncertain lost customers

Prinsip yang sama juga berlaku untuk mengidentifikasi tipe pelanggan berdasarkan model penilaian LRFM. Tabel 3 menampilkan 16 (enam belas) jenis pelanggan berdasarkan kombinasi simbol dari hasil skor variabel LRFM yang dibuat oleh Chang dan Tsay [16].

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan dan Pre-processing Data

Kegiatan pengumpulan, pembersihan, dan pemilihan data menghasilkan sebanyak 7 kolom atribut x 5,826 baris data riwayat transaksi *online* pelanggan di Toko X yang siap diolah. Data tersebut diperoleh dalam bentuk excel (.xlsx) dengan periode Maret 2021 - Maret 2023. Adapun kolom atribut yang dipilih telah disesuaikan dengan kebutuhan proses segmentasi seperti pada Tabel 4.

Tabel 4. Atribut Data Setelah Dilakukan Pre-processing Data

Variabel	Tipe Data	Keterangan
Nomor Transaksi	Varchar	Kombinasi unik antara huruf dan angka membentuk sebuah nomor transaksi
Username Pelanggan	Varchar	Username untuk mengenali pelanggan adalah <i>unique value</i> .
Tanggal Transaksi	Date	Tanggal transaksi menunjukkan kapan pesanan dibuat
Nominal Transaksi	Int	Nominal transaksi merupakan total uang yang dibelanjakan dalam satu nomor transaksi

Berdasarkan Tabel 4, atribut data yang digunakan dalam proses analisis LRFM dan MLRFM pelanggan berjumlah 4 atribut. Atribut-atribut yang dimaksudkan adalah nomor transaksi, *username* pelanggan untuk mencari nilai *frequency* pelanggan, tanggal transaksi yang dibutuhkan untuk mengetahui nilai *length* dan *recency* pelanggan, serta nominal transaksi guna mendapatkan nilai *monetary* pelanggan.

3.2. Analisis LRFM

Dataset kemudian diekstraksi menggunakan model LRFM sehingga menghasilkan nilai variabel *Length* (L), *Recency* (R), *Frequency* (F), dan *Monetary* (M) setiap pelanggan. Untuk variabel *recency* dihitung interval antara tanggal terakhir kali pelanggan melakukan pemesanan sampai dengan tanggal penelitian dilakukan, yakni 28 Oktober 2023. Proses ini memberikan informasi bahwa sebanyak 4,392 pelanggan telah bertransaksi di Toko *Online X* selama Maret 2021 – Maret 2023. Cuplikan nilai awal LRFM pelanggan terdapat pada Tabel 5.

Tabel 5. Nilai LRFM Pelanggan Toko *Online X*

No.	Username Pelanggan	L (Hari)	R (Hari)	F	M (IDR)
1	Riz*** Syah***	349	247	86	4,636,600
2	Ieri***	301	581	36	4,151,400
3	Kre***	475	245	25	1,470,854
4	Iswi*****	99	233	9	963,900
5	Kisw***	106	242	7	429,700
6	alifia2***	193	555	4	170,053
7	Keyzshop***	13	893	2	145,823
8	Sheeta***	1	937	2	107,000

Nilai LRFM yang ada pada Tabel 5 menunjukkan bahwa adanya perbedaan nilai antar variabel yang cukup jauh, khususnya variabel *monetary* dengan ketiga variabel lain. Sehingga

berikutnya dilakukan normalisasi menggunakan persamaan (1). Hasil normalisasi beserta pemetaan skor LRFM yang mengacu pada Tabel 1 telah disajikan pada Tabel 6.

Tabel 6. Normalisasi Nilai LRFM Pelanggan Toko *Online X*

No	Username Pelanggan	Kategori	L	R	F	M
1	Riz***	Normalisasi	0,507	0,953	0,85	0,743
	Syah***	Skor	4	5	5	5
2	Ieri***	Normalisasi	0,438	0,513	0,35	0,665
		Skor	3	4	3	4
3	Kre***	Normalisasi	0,690	0,955	0,24	0,236
		Skor	4	5	2	2
4	Iswi*****	Normalisasi	0,144	0,971	0,08	0,154
		Skor	2	5	1	2
5	Kisw***	Normalisasi	0,154	0,960	0,06	0,069
		Skor	2	5	1	1
6	alifia2***	Normalisasi	0,281	0,548	0,03	0,027
		Skor	2	4	1	1
7	Keyzshop**	Normalisasi	0,019	0,103	0,01	0,023
		Skor	1	2	1	1
8	Sheeta***	Normalisasi	0,001	0,045	0,01	0,017
		Skor	1	1	1	1

Setelah skor LRFM pelanggan diketahui dari Tabel 6, masing-masing skor variabel dikalikan dengan bobot variabel sesuai hasil diskusi dengan Toko *Online X*. Setiap variabel memiliki bobot yang sama, yaitu 0.25. Sehingga skor akhir LRFM pelanggan diperoleh seperti pada Tabel 7.

Tabel 7. Skor Akhir LRFM Pelanggan Toko *Online X*

No.	Username Pelanggan	Skor L	Skor R	Skor F	Skor M
1	Riz*** Syah***	1	1,25	1,25	1,25
2	Ieri***	0,75	1	0,75	1
3	Kre***	1	1,25	0,5	0,5
4	Iswi*****	0,5	1,25	0,25	0,5
5	Kisw***	0,5	1,25	0,25	0,25
6	alifia2***	0,5	1	0,25	0,25
7	Keyzshop***	0,25	0,5	0,25	0,25
8	Sheeta***	0,25	0,25	0,25	0,25

Berdasarkan Tabel 7, terlihat bahwa pelanggan dengan gabungan skor akhir LRFM terbesar menempati nomor pertama, terbesar selanjutnya menempati posisi kedua, dan seterusnya hingga posisi kedelapan memiliki skor akhir yang paling kecil.

3.3. Analisis MLRFM

Dataset kemudian diproses dengan model MLRFM yang menghasilkan skor akhir variabel *recency*, *frequency*, dan *monetary*. Perhitungan variabel *frequency* dan *monetary* akan dibagi menjadi beberapa *layer* berdasarkan periode waktu yang ada pada Tabel 8.

Tabel 8. Pembobotan *Multi Layer*

Periode Waktu (Hari)	Bobot Layer
760	0.5
365	0.3
180	0.2

Daftar *layer* periode waktu, beserta bobotnya masing-masing yang disajikan oleh Tabel 8 didasarkan pada pertimbangan bersama dengan

pemilik dan kepala tim penjualan Toko *Online X*. Setelah *layer* dan bobot *layer* diketahui, maka selanjutnya meringkas nilai RFM masing-masing pelanggan, seperti yang dituliskan pada Tabel 9.

Ringkasan nilai RFM pelanggan pada Tabel 9 kemudian dilakukan normalisasi dengan persamaan

(1) dan diberi skor seperti yang dilakukan pada analisis LRFM. Pemetaan skor untuk setiap hasil normalisasi juga mengacu pada Tabel 1 sebelumnya. Hasil normalisasi dapat dilihat pada Tabel 10.

Tabel 9. Nilai Awal RFM Pelanggan Toko *Online X*

No.	Username Pelanggan	Recency - Hari	Frequency (180 Hari)	Frequency (365 Hari)	Frequency (760 Hari)	Monetary - IDR (180 Hari)	Monetary - IDR (365 Hari)	Monetary - IDR (760 Hari)
1	Riz*** Syah***	247	30	79	86	1,612,600	4,258,600	4,636,600
2	Ieri***	581	0	0	36	0	0	4,151,400
3	Kre***	245	2	13	25	123,160	771,752	1,470,854
4	Iswj*****	233	9	9	9	963,900	963,900	963,900
5	Kisw***	242	7	7	7	429,700	429,700	429,700
6	alifia2***	555	0	2	4	0	40,273	170,053
7	Keyzshop***	893	0	0	2	0	0	145,823
8	Sheeta***	937	0	0	2	0	0	107,000

Tabel 10. Nilai Normalisasi dan Skor MLRFM Pelanggan Toko *Online X*

No.	Username Pelanggan	Kategori	Recency	Frequency (180 Hari)	Frequency (365 Hari)	Frequency (760 Hari)	Monetary (180 Hari)	Monetary (365 Hari)	Monetary (760 Hari)
1	Riz*** Syah***	Normalisasi	0,953	0,305	0,78	0,85	0,272	0,683	0,743
		Skor	5	3	5	5	2	4	5
2	Ieri***	Normalisasi	0,513	0,0	0,0	0,35	0,0	0,0	0,665
		Skor	4	1	1	3	1	1	4
3	Kre***	Normalisasi	0,955	0,011	0,12	0,24	0,021	0,124	0,236
		Skor	5	1	2	2	1	2	2
4	Iswj*****	Normalisasi	0,971	0,084	0,08	0,08	0,162	0,154	0,154
		Skor	5	1	1	1	2	2	2
5	Kisw***	Normalisasi	0,959	0,063	0,06	0,06	0,072	0,069	0,069
		Skor	5	1	1	1	1	1	1
6	alifia2***	Normalisasi	0,547	0,0	0,01	0,03	0,0	0,006	0,027
		Skor	4	1	1	1	1	1	1
7	Keyzshop***	Normalisasi	0,103	0,0	0,0	0,01	0,0	0,0	0,023
		Skor	2	1	1	1	1	1	1
8	Sheeta***	Normalisasi	0,045	0,0	0,0	0,01	0,0	0,0	0,017
		Skor	1	1	1	1	1	1	1

Berdasarkan Tabel 10 dapat dilihat nilai normalisasi beserta skor yang diperoleh pelanggan. Khusus untuk skor *frequency* dan *monetary* pelanggan pada setiap *layer* waktu (180 – 365 – 720 hari) dikalikan dengan bobot *layer* yang telah terdefinisi pada Tabel 8 untuk mendapatkan total skor F dan M. Terakhir, total skor masing-masing variabel dikalikan dengan bobot variabel tersebut hingga didapatkan skor akhir MLRFM pelanggan menggunakan persamaan (2). Bobot RFM yang diinginkan Toko *Online X* berturut-turut adalah 0.33, 0.33, dan 0.34.

pada perhitungan model LRFM, begitu pula dengan posisi kedua hingga kedelapan tidak mengalami perubahan posisi pelanggan.

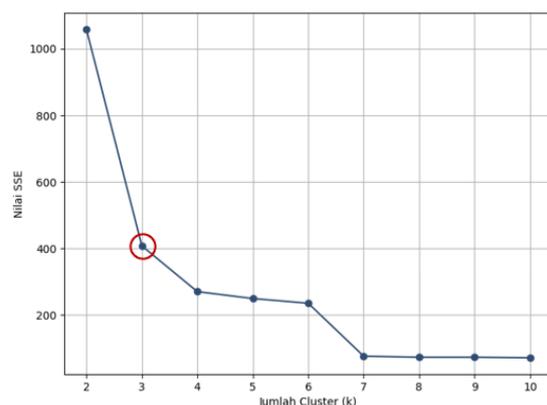
3.4. Clustering dengan RM K-Means

Tahap selanjutnya adalah melakukan pengelompokan pelanggan dengan algoritma RM K-Means. Inputan untuk menjalankan algoritma ini menggunakan hasil skor akhir dari masing-masing model analisis pelanggan (LRFM dan MLRFM) yang telah dilakukan sebelumnya.

Tabel 11. Skor Akhir MLRFM Pelanggan Toko *Online X*

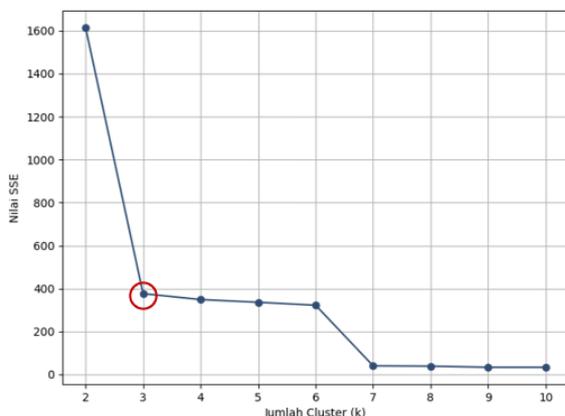
No.	Username Pelanggan	Skor R	Skor Multi Layer F	Skor Multi Layer M
1	Riz*** Syah***	1,65	1,52	1,39
2	Ieri***	1,32	0,5	0,68
3	Kre***	1,65	0,59	0,61
4	Iswj*****	1,65	0,33	0,68
5	Kisw***	1,65	0,33	0,34
6	alifia2***	1,32	0,26	0,27
7	Keyzshop***	0,66	0,17	0,17
8	Sheeta***	0,33	0,17	0,17

Berdasarkan Tabel 11, diberikan informasi bahwa pelanggan dengan skor terbesar masih sama dengan pelanggan yang menempati posisi pertama



Gambar 2. Plot *Elbow* Berdasarkan Data LRFM Pelanggan

Gambar 2 menunjukkan bahwa turunan terbesar dari perhitungan nilai SSE dengan persamaan (4) yang telah dilakukan berada pada K=3. Hal ini ditunjukkan dengan lingkaran merah pada Gambar 2. Oleh karena itu, jumlah *cluster* (*k*) optimal untuk kombinasi model LRFM dengan RM K-Means adalah 3 (tiga) *cluster*.



Gambar 3. Plot Elbow Berdasarkan Data MLRFM Pelanggan

Selanjutnya, jika dilihat pada Gambar 3 diketahui juga bahwa turunan terbesar dari perhitungan nilai SSE yang telah dilakukan berdasarkan data MLRFM pelanggan berada pada K=3. Hal ini ditunjukkan dengan lingkaran merah pada Gambar 3. Oleh karena itu, jumlah *cluster* (*k*) yang optimal adalah 3 (tiga) *cluster* untuk membentuk kelompok pelanggan dari kombinasi model MLRFM dengan RM K-Means. Selanjutnya, proses *clustering* berdasarkan masing-masing model dengan kombinasi RM K-Means memberikan ringkasan jumlah anggota untuk setiap segmen seperti pada Tabel 12.

Tabel 12. Jumlah Anggota Setiap Segmen

Model Analisis	Cluster 1	Cluster 2	Cluster 3
LRFM	1,427	912	2,053
MLRFM	1,427	912	2,053

Berdasarkan Tabel 12, jumlah anggota di setiap *cluster* dengan kombinasi model LRFM dan MLRFM tidak memiliki perbedaan. Hal ini menunjukkan bahwa model LRFM dan MLRFM tidak berpengaruh terhadap banyaknya jumlah anggota pada masing-masing *cluster*.

Tabel 13. Rata-rata Keseluruhan Skor LRFM

Average Skor L	Average Skor R	Average Skor F	Average Skor M
0,26086	1,0243	0,25052	0,25098

Pada Tabel 13 diperoleh rata-rata keseluruhan skor LRFM dari 4,392 pelanggan Toko *Online X*. Rata-rata ini akan menjadi nilai acuan bagi tahap analisis tipe segmen pelanggan.

Tabel 14. Rata-rata Keseluruhan Skor MLRFM

Average Skor R	Average Skor F	Average Skor M
1,35208	0,26078	0,26915

Selanjutnya, Tabel 14 menampilkan hasil perhitungan rata-rata keseluruhan untuk skor MLRFM pelanggan dengan jumlah pelanggan yang sama. Hasil rata-rata keseluruhan pada Tabel 14 juga digunakan untuk mengidentifikasi tipe pelanggan tiap-tiap segmen.

3.5. Pengujian Validitas Cluster

Silhouette coefficient digunakan untuk menguji validitas *cluster* yang terbentuk oleh RM K-Means dalam penelitian ini. Kauffman dan Rousseeuw mengelompokkan hasil *silhouette coefficient* ke dalam 4 golongan seperti yang ditunjukkan oleh Tabel 15 [24].

Tabel 15. Interpretasi Kauffman

Rentang Nilai <i>Silhouette Coefficient</i>	Interpretasi Kauffman
0,71 – 1,00	<i>Strong Classification</i>
0,51 – 0,70	<i>Good Classification</i>
0,26 – 0,50	<i>Weak Classification</i>
≤ 0,25	<i>Bad Classification</i>

Dengan persamaan (5), didapatkan rata-rata nilai *silhouette coefficient* untuk kombinasi LRFM dan RM K-Means dengan 3 *cluster* sebesar 0,8485. Sedangkan, kombinasi MLRFM dengan RM K-Means dengan 3 *cluster* sebesar 0,8609. Berdasarkan hasil tersebut, kedua kombinasi perhitungan yang dilakukan untuk mensegmentasi pelanggan Toko *Online X* sama-sama termasuk kategori *Strong Classification*. Diketahui pula bahwa kombinasi MLRFM dengan RM K-Means lebih unggul dengan selisih nilai 0,0124.

3.6. Analisis Tipe Segmen Pelanggan

Proses analisis ini membandingkan nilai rata-rata variabel LRFM dan MLRFM dari masing-masing *cluster* dengan rata-rata keseluruhan. Simbol tanda panah ke atas (↑) dan tanda panah ke bawah (↓) digunakan untuk menunjukkan apakah nilai rata-rata skor variabel pada suatu *cluster* lebih tinggi atau rendah dari total rata-rata skor yang ada [19].

Tabel 16. Pemetaan Tipe Segmen Pelanggan dengan LRFM

k	Average Skor L	Average Skor R	Average Skor F	Average Skor M	Tipe Segmen
1	0,253 (↓)	0,663 (↓)	0,250 (↓)	0,250 (↓)	<i>Uncertain lost customers</i>
2	0,259 (↓)	1,0 (↓)	0,250 (↑)	0,250 (↓)	<i>Frequency churn customers</i>
3	0,266 (↑)	1,25 (↑)	0,250 (↑)	0,251 (↑)	<i>High-value loyal customers</i>

Dari Tabel 16 dapat diketahui bahwa *cluster* 1 dan 2 termasuk ke dalam tipe *lost customers* dengan sub-tipe yang berbeda. Pada *cluster* 1, rata-rata skor L, R, F, dan M rendah jika dibandingkan dengan rata-rata keseluruhan. Hal ini menunjukkan bahwa

pelanggan pada *cluster* 1 jarang melakukan pembelian dalam interval waktu sejak pertama kali hingga terakhir kali dilakukannya transaksi oleh pelanggan di Toko *Online X*. Adapun jumlah uang yang dikeluarkan pada masa lampau juga rendah, serta pelanggan ini sudah lama tidak melakukan transaksi. Kemudian, *cluster* 2 memiliki rata-rata skor L, R, dan M rendah daripada rata-rata keseluruhan. Namun, rata-rata skor F pada *cluster* ini lebih tinggi daripada rata-rata F keseluruhan. Hasil ini menunjukkan bahwa *cluster* 2 merupakan golongan pelanggan yang sering melakukan transaksi di Toko *Online X* di masa lampau, namun jumlah uang yang dikeluarkan tidak besar dan baru-baru ini tidak melakukan transaksi kembali. Lanjut, pelanggan pada *cluster* 3 adalah tipe pelanggan yang loyal karena rata-rata skor L, R, F, M berada di atas rata-rata skor keseluruhan. Pelanggan yang berada pada kelompok ini adalah pelanggan yang masih setia bertransaksi dengan Toko *Online X* sejak transaksi pertama dibuat hingga saat ini. Pelanggan ini bertransaksi dalam nominal yang besar dan baru-baru ini masih aktif melakukan transaksi.

Tabel 17. Pemetaan Tipe Segmen Pelanggan dengan MLRFM

k	Average Skor R	Average Skor F	Average Skor M	Tipe Segmen
1	0,87519 (↓)	0,16521 (↓)	0,17047 (↓)	Lost customers
2	1,31997 (↓)	0,25934 (↓)	0,26755 (↓)	Lost customers
3	1,64996 (↑)	0,31674 (↑)	0,32701 (↑)	Loyal customers

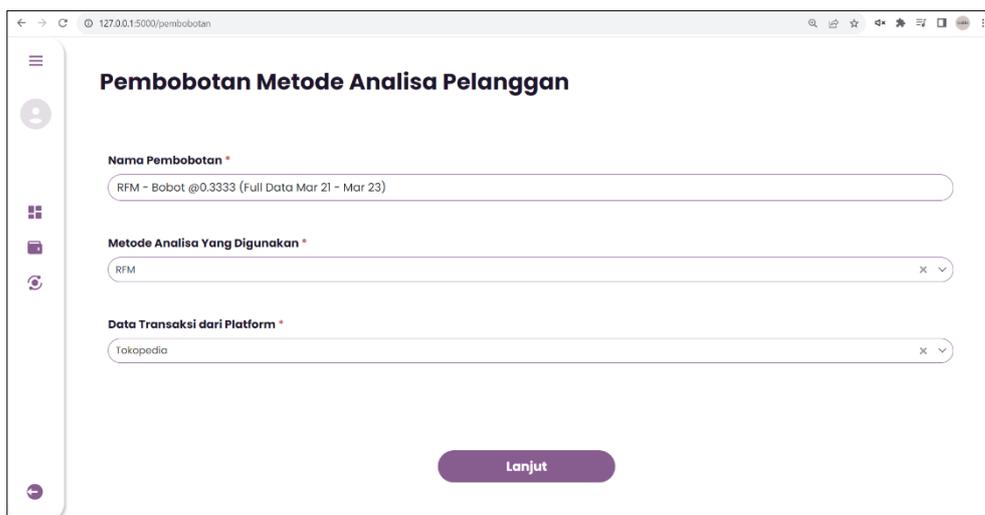
Kemudian, Tabel 17 menunjukkan bahwa tipe pelanggan di *cluster* 1 dan 2 juga termasuk sebagai *lost customers* karena memiliki rata-rata R, F, dan M yang rendah dari rata-rata keseluruhan. Hal ini berarti bahwa pelanggan tersebut sudah lama tidak kembali

melakukan transaksi di Toko *Online X*. Kemudian, pelanggan-pelanggan ini jarang melakukan transaksi di masa lampau dan hanya menghabiskan uangnya dengan nominal di bawah rata-rata. Namun jika *cluster* 1 dan *cluster* 2 dibandingkan, maka pelanggan pada *cluster* 2 masih lebih baik karena rata-rata skor yang lebih tinggi dari *cluster* 1. Beralih ke *cluster* 3, di mana pelanggan pada kelompok ini merupakan pelanggan yang loyal. Pelanggan-pelanggan ini masih aktif dan sering melakukan pembelian di Toko *Online X* hingga saat ini. Nominal yang dikeluarkan saat bertransaksi juga besar dan lebih tinggi daripada rata-rata skor keseluruhan.

3.7. Implementasi Sistem

Hasil penelitian ini juga diimplementasikan dalam bentuk *website* sehingga memudahkan pemilik dan tim penjualan Toko *Online X* untuk melihat hasil segmentasi yang dilakukan. Bahasa pemrograman yang digunakan adalah *Python*, HTML dan CSS, dengan MySQL sebagai basis data. Aktor yang berperan dalam sistem adalah admin.

Halaman pembobotan yang ditunjukkan oleh Gambar 4 merupakan menu yang disediakan agar admin dapat melakukan penginputan bobot variabel LRFM atau MLRFM sebelum sistem menjalankan model analisis. Selanjutnya, data pembobotan yang telah diinputkan ke dalam *database* oleh admin akan ditampilkan pada halaman data simulasi atau pembobotan perhitungan seperti yang ditunjukkan oleh Gambar 5. Jika ingin melihat hasil perhitungan dari model analisis pelanggan dan segmentasi, admin dapat menekan salah satu tombol yang tersedia pada kolom hasil sesuai dengan metode yang diinginkan. Gambar 6 merupakan contoh tampilan hasil perhitungan segmentasi berdasarkan suatu data bobot yang dipilih admin.



Gambar 4. Halaman Pembobotan

No.	Nama Bobot	Platform	Metode	Bobot	Pembobotan Multi Layer	Hasil	Aksi
1	Multi Layer RFM - Bobot @0.3333 (Full Data Mar 21 - Mar 23)	Tokopedia	MLRFM	R : 0.33 F : 0.33 M : 0.34	Periode (days) : 180 - 760 - 365 Bobot Periode : 0.2 - 0.5 - 0.3	MLRFM RM Kmeans	
2	RFM - Bobot @0.3333 (Full Data Mar 21 - Mar 23)	Tokopedia	RFM	R : 0.33 F : 0.33	-	RFM	

Gambar 5. Halaman Daftar Bobot Simulasi Perhitungan

Recency (R)	Frequency (F)	Monetary (M)
1.30956	0.25593	0.26417

Cluster	Average Skor R	Average Skor F	Average Skor M	Jumlah Pelanggan	Kategori Pelanggan
1	0.81309	0.16558	0.17107	1427	Lost Customers
2	1.32	0.25966	0.26801	912	Loyal Customers
3	1.65	0.31708	0.32717	2053	Loyal Customers

Gambar 6. Halaman Hasil Segmentasi Pelanggan

4. DISKUSI

Penelitian ini menghasilkan segmen-segmen pelanggan Toko *Online X*, beserta tipe atau karakteristiknya berdasarkan kombinasi LRFM dengan RM K-Means dan MLRFM dengan RM K-Means. Sehingga hasil penelitian ini dapat membantu Toko *Online X* dalam mengenali pelanggan loyal dan potensial berdasarkan kombinasi model analisis pelanggan dan *cluster* terbaik. Berdasarkan hasil perhitungan metode *silhouette coefficient*, kombinasi MLRFM dan RM K-Means memiliki nilai *silhouette* yang lebih baik daripada kombinasi LRFM dengan RM K-Means, yaitu sebesar 0,8609. Nilai *silhouette* dari kombinasi MLRFM dan RM K-Means menunjukkan bahwa *cluster* yang dihasilkan dapat dikategorikan sebagai *strong classification*.

Kombinasi MLRFM dan RM K-Means membagi pelanggan Toko *Online X* menjadi tiga *cluster*, di mana pelanggan pada *cluster* 1 dan 2 tergolong sebagai *lost customer* karena simbol skor RFM untuk kedua *cluster* adalah R↓ F↓ M↓. Namun, *cluster* 2 dapat dikatakan masih lebih unggul dengan rata-rata skor RFM yang lebih tinggi dari *cluster* 1.

Sedangkan, pelanggan yang berada di *cluster* 3 termasuk sebagai *loyal customer* dengan simbol skor RFM ditandai dengan R↑ F↑ M↑.

Selanjutnya, penelitian terdahulu yang juga melakukan segmentasi pelanggan telah ditemukan. Beberapa diantaranya:

- a. Penelitian yang dilakukan oleh Widiyanto dan Watanti pada 2021 berjudul "Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran" [10]. Penelitian tersebut menyatakan bahwa tingkat akurasi *cluster* yang dihasilkan menggunakan algoritma K-Means sudah baik. Namun, pengujian hanya dilakukan berdasarkan wawancara atau *User Acceptance Test (UAT)*.
- b. Penelitian dengan judul "E-Commerce Customer Segmentation Using K-Means Algorithm and Length, Recency, Frequency, Monetary Model" oleh Siagian et al. pada 2021. Hasil penelitian tersebut menggolongkan tipe pelanggan untuk setiap *cluster* yang terbentuk berdasarkan algoritma K-Means tradisional dengan menggunakan LRFM *Index* [11].

Namun, akurasi *cluster* yang dihasilkan pada penelitian ini tidak diketahui.

- c. Penelitian selanjutnya dilakukan oleh Adilla *et al.* pada 2022 dengan judul "Implementasi K-Means Clustering Analysis untuk Menentukan Hambatan Pembelajaran Daring pada Siswa Studi Kasus: SMP Swasta Yapendak Tinjowan". Penelitian tersebut membantu objek penelitian dalam menentukan siswa-siswa yang memiliki hambatan selama pembelajaran daring sehingga dapat diberi perlakuan khusus [12]. Namun, jumlah *cluster* masih ditentukan secara manual dan tingkat akurasi *cluster* yang dihasilkan juga tidak diketahui.

Berdasarkan penjabaran di atas, dapat dikatakan bahwa penelitian ini mampu menjawab kekurangan pada penelitian sebelumnya dengan mengetahui jumlah *cluster* optimal melalui metode *elbow* dan tingkat akurasi *cluster* yang dihasilkan secara objektif berdasarkan nilai *silhouette coefficient*. Pengembangan baru dari model analisis pelanggan dan metode *clustering* yang digunakan pada penelitian ini telah disesuaikan dengan kebutuhan Toko *Online X*, di mana kombinasi MLRFM dan RM K-Means lebih cocok untuk mengidentifikasi loyalitas pelanggan berdasarkan pembagian periode waktu transaksi Toko *Online X* karena tingkat akurasi yang dihasilkan lebih tinggi.

5. KESIMPULAN

Berdasarkan penelitian ini, kesimpulan yang dapat ditarik adalah hasil segmentasi kombinasi MLRFM dengan RM K-Means lebih baik dalam membantu Toko *Online X* mengidentifikasi pelanggan loyal dan potensial. Kombinasi tersebut menghasilkan tiga *cluster* dengan *cluster* 1 dengan 1,427 pelanggan dan *cluster* 2 yang memiliki 912 pelanggan merupakan kelompok *lost customer* (pelanggan yang hilang) karena simbol skor RFM yang dimiliki adalah $R \downarrow F \downarrow M \downarrow$. Sedangkan, *cluster* 3 terdiri dari 2,053 pelanggan yang memiliki loyalitas tinggi dengan ditunjukkan oleh simbol $R \uparrow F \uparrow M \uparrow$. Selain itu, *website* segmentasi pelanggan dibangun secara khusus untuk Toko *Online X* agar memudahkan perusahaan dalam memasukkan, melihat visual, dan menyimpan data, serta hasil segmentasi.

DAFTAR PUSTAKA

- [1] A. Premana, G. Fitralisma, A. Yulianto, M. B. Zaman, and M. A. Wiryo, "Pemanfaatan Teknologi Informasi Pada Pertumbuhan Ekonomi Dalam Era Disrupsi 4.0", *J. of Eco. and Manag. (JECMA)*, vol. 2, no. 2, pp. 1–6, Aug. 2020, doi: 10.46772/jecma.v1i01.219.
- [2] N. Nofiyanti and G. Wiwoho, "Pengaruh Perceived Usefulness dan Perceived Value Terhadap Online Repurchase Intention pada Marketplace Shopee di Kecamatan Kebumen dengan Customer Satisfaction sebagai Variabel Intervening", *J. Ilmiah Mahasiswa Manaj. Bis. dan Akun. (JIMMBA)*, vol. 2, no. 2, pp. 281-290, Apr. 2020, doi: 10.32639/jimmba.v2i2.477.
- [3] M. A. Rizaty, "Transaksi e-Commerce RI Tak Capai Target pada 2022," *DataIndonesia*, Jan. 24, 2023. Accessed: Oct. 5, 2023. [Online]. Available: <https://dataindonesia.id/digital/detail/transaksi-ecommerce-ri-tak-capai-target-pada-2022>.
- [4] I. Tzavlopoulos, K. Gotzamani, A. Andronikidis, and C. Vassiliadis, "Determining The Impact of E-Commerce Quality on Customers Perceived Risk, Satisfaction, Value and Loyalty," *Inter J. of Quality and Service Sci.*, vol. 11, no. 4, pp. 576-587, Dec. 2019, doi: 10.1108/ijqss-03-2019-0047.
- [5] P. Rita, T. Oliveira, and A. Farisa, "The Impact of E-Service Quality and Customer Satisfaction on Customer Behavior in Online Shopping", *Heliyon*, vol. 5, no. 10, pp. 1-14, Oct. 2019, doi: 10.1016/j.heliyon.2019.e02690.
- [6] Y. F. Rahardian and B. Syairudin, "Segmentation Analysis of Students in X Course with RFM Model and Clustering," *J. Sos. Hum. (JSH)*, vol. 0, no. 1, pp. 59-79, Aug. 2020, doi: 10.12962/j24433527.v0i1.6776.
- [7] S. Monalisa, P. Nadya, and R. Novita, "Analysis for Customer Lifetime Value Categorization with RFM Model," in *Procs. Comp. Sci.*, 2019, pp. 834-840, doi: 10.1016/j.procs.2019.11.190.
- [8] R. Pradhan, "Customer Segmentation Using Clustering Approach Based on RFM Analysis," in *Inter. Conference on Infor. Sys. and Comp. Networks (ISCON)*, 2021, pp. 1-5, doi: 10.1109/ISCON52037.2021.9702482.
- [9] A. A. D. Sulistyawati and M. Sadikin, "Penerapan Algoritma K-Medoids Untuk Menentukan Segmentasi Pelanggan," *J. Sis. Infor. (SISTEMASI)*, vol. 10, no. 3, pp. 516-526, Sept. 2021, doi: 10.32520/stmsi.v10i3.1332.
- [10] A. Widiyanto and A. Witanti, "Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global)," *Konvergensi Tek. dan Sis. Infor (Konstelasi)*, vol. 1, no. 1, pp. 204-215, Apr. 2021, doi: 10.24002/konstelasi.v1i1.4293.
- [11] R. Siagian, P. Sirait, and A. Halima, "E-Commerce Customer Segmentation Using K-Means Algorithm And Length, Recency,

- Frequency, Monetary Model," *J. of Inform. and Telecomm. Engineering (JITE)*, vol. 5, no. 1, pp. 21-39, Jul. 2021, doi: 10.31289/jite.v5i1.5182.
- [12] D. Adillah, N. Manurung, and A. Dermawan. "Implementation of K-Means Clustering Analysis to Determine Barriers to Online Learning Case Study: Swasta Yappendak Tinjowan Junior High School," *J. Tek. Inform. (JUTIF)*, vol. 3, no. 3, pp. 519-525, Jun. 2022, doi: 10.20884/1.jutif.2022.3.3.188.
- [13] R. Rahmadiani, A. Dhini, and E. Laoh, "Estimating Customer Lifetime Value using LRFM Model in Pharmaceutical and Medical Device Distribution Company," in *Inter. Con. on ICT for Smart Society (ICISS)*, 2020, doi: 10.1109/iciss50791.2020.9307592.
- [14] A. Handojo, N. Pujawan, B. Santosa, and M. L. Singgih, "A Multi Layer Recency Frequency Monetary Method For Customer Priority Segmentation in Online Transaction," *Cogent Engineering*, vol. 10, no. 1, pp. 1-19, Jan. 2023, doi: 10.1080/23311916.2022.2162679.
- [15] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, "RFM Ranking – An Effective Approach to Customer Segmentation," *J. of King Saud Univ.*, vol. 33, no. 10, pp. 1251-1257, Sept. 2018, doi: 10.1016/j.jksuci.2018.09.004.
- [16] S. H. Chao, M. K. Chen, and H. H. Wu, "An LRFM Model to Analyze Outpatient Loyalty From a Medical Center in Taiwan," *SAGE Open*, vol. 11, no. 3, Jul. 2021, doi: 10.1177/21582440211031899.
- [17] M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance Pada Algoritma K-Means Clustering Berbasis Chi-Square," *J. Infor.: J. Pengembangan IT (JPIT)*, vol. 4, no. 1, pp. 20-24, Jan. 2019, doi: 10.30591/jpit.v4i1.1253.
- [18] P. M. Hasugian, B. Sinaga, J. Manurung, and S. A. Al Hashim, "Best Cluster Optimization with Combination of K-Means Algorithm And Elbow Method Towards Rice Production Status Determination," *Inter. J. Of Artif. Intel. (IJAIR)*, vol. 5, no. 1, pp. 102-110, Jun. 2021, doi: 10.29099/ijair.v6i1.232.
- [19] E. Umargono, J. E. Suseno, and V. Gunawan, "K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula," in *Proceedings of the 2nd Inter. Seminar on Sci. and Tech.*, 2019, vol. 474, pp. 121-129, doi: 10.2991/assehr.k.201010.019.
- [20] H.B. Tambunan, D. H. Barus, J. Hartono, A. S. Alam, D. A. Nugraha, and H. H. H. Usman, "Electrical Peak Load Clustering Analysis Using K-Means Algorithm and Silhouette Coefficient," in *Inter. Conference on Tech. and Policy in Energy and Electric Power (ICT-PEP)*, 2020, doi: 10.1109/ict-pep50916.2020.9249773.
- [21] I. Pauletic, L. N. Prskalo, and M. B. Bakaric, "An Overview of Clustering Models with an Application to Document Clustering," in *Inter. Convention on Infor. and Communication Tech., Electronics and Microelectronic (MIPRO)*, (Croatia), 2019, doi: 10.23919/mipro.2019.8756868.
- [22] S. H. Ha and S. C. Park, "Application of Data Mining Tools To Hotel Data Mart on The Intranet Fordatabase Marketing," *Expert Syst. Appl.*, vol. 15, pp. 1–13, 1998.
- [23] T. Juhari and A. Juarna, "Implementation Rfm Analysis Model For Customer Segmentation Using The K-Means Algorithm Case Study XYZ Online Bookstore," *Explore*, vol. 12, no. 1, pp. 107-118, 2022, doi: 10.35200/explore.v12i1.548.
- [24] L. Kaufman and P. J. Rousseeuw, in *Finding Groups in Data an Introduction to Cluster Analysis*, 1st ed. Hoboken, New Jersey, U.S.: John Wiley & Sons Inc. Publication, 1990. . Accessed: Sept. 28, 2023. [Online]. Available: 10.1002/9780470316801.