

Segmentasi Pelanggan Menggunakan K-Medoids Berdasarkan Model Length, Recency, Frequency, Monetary (LRFM)

Gilang Purnama, Tacbir Hendro Pudjiantoro, Puspita Nurul Sabrina

Jurusan Informatika, Fakultas Sains dan Informatika

Universitas Jenderal Achmad Yani

Jl. Terusan Sudirman, Cimahi

gilangpurnama17@if.unjani.ac.id, thp@if.unjani.ac.id, puspitasabrina14@gmail.com

Abstrak— Penerapan strategi yang sama akan menimbulkan beberapa dampak, seperti promosi yang tidak tepat sasaran sehingga strategi promosi tersebut hanya akan bermanfaat terhadap pelanggan tertentu saja. Selain itu proses analisis data transaksi penjualan membutuhkan waktu yang cukup lama dan variabel yang tepat untuk mendapatkan informasi seperti segmentasi pelanggan dan nilai pelanggan yang diperlukan dalam menargetkan strategi pemasaran ke segmen yang paling menjanjikan. Segmentasi pelanggan dapat digunakan untuk mendukung tindakan strategis perusahaan mengenali perbedaan pelanggan dapat menjadi kunci pemasaran yang sukses dan lebih efektif. Oleh karena, penelitian ini akan berfokus pada segmentasi pelanggan dengan menggunakan metode data mining yaitu clustering menggunakan algoritma K-medoids dan model Length, Recency, Frequency, Monetary (LRFM) untuk mengidentifikasi tingkat loyalitas pelanggan. Berdasarkan hasil clustering menunjukan cluster 1 termasuk kelompok uncertain new customers dengan hasil nilai length 0.437, recency 0.436, frequency 0.003 dan monetary 0.002, Strategi yang diusulkan untuk kelompok ini yaitu “Let-go” strategy. Cluster 2 termasuk kelompok lost customer dengan hasil nilai length 0.52933, recency 0.457, frequency 0.094 dan monetary 0.049, Strategi yang diusulkan untuk kelompok ini yaitu Enforced Strategy. Cluster 3 termasuk kelompok core customer dengan hasil nilai length 0.546, recency 0.540, frequency 0.007 dan monetary 0.009, Strategi yang diusulkan untuk kelompok ini yaitu Offensive Strategy.

Kata kunci—Segmentasi pelanggan; K-Medoids; LRFM; CRISP-DM

I. PENDAHULUAN

Perusahaan menghasilkan dan menyimpan sejumlah besar data atau informasi setiap harinya, tetapi data yang disimpan biasanya tidak memberikan manfaat langsung. Nilai informasi sebenarnya terletak pada data atau informasi yang dapat kita ekstrak darinya [1]. Pengumpulan data transaksi dengan rentang waktu Januari 2020 sampai Juni 2020, data transaksi penjualan menunjukkan penurunan jumlah transaksi pada bulan April 2020 yang signifikan sebanyak 52% dari bulan sebelumnya. Adapun terjadinya penurunan penjualan disebabkan oleh berbagai macam faktor. Salah satunya adalah penerapan strategi promosi yang sama sehingga strategi tidak tepat sasaran dan hanya bermanfaat pada pelanggan tertentu saja. Loyalitas pelanggan dapat menjamin pelaku usaha untuk tetap memberikan layanan

terbaik agar pelanggan dapat bertambah dari waktu ke waktu ataupun tidak kehilangan pelanggan [2].

Segmentasi pelanggan dapat membantu pebisnis menyesuaikan rencana pemasaran, mengidentifikasi tren, merencanakan pengembangan produk, promosi iklan dan menawarkan produk yang relevan [3]. Segmentasi pelanggan dapat digunakan untuk mendukung tindakan strategis perusahaan. Mengenali perbedaan pelanggan dapat menjadi kunci pemasaran yang sukses dan lebih efektif [4].

Sampai saat ini Artch menggunakan banyak strategi untuk menjaga loyalitas pelanggan seperti memberikan diskon untuk pembelian berikutnya, paket hemat, kartu anggota dan lainnya secara menyeluruh kepada semua pelanggan. Segmentasi merupakan strategi pemasaran klasik yang dijelaskan dalam setiap buku pegangan yang dikhususkan untuk bisnis. Oleh karena itu, membagi pelanggan menjadi beberapa kelompok yang disebut segmentasi pelanggan atau segmen pasar, dan kemudian menargetkan upaya pemasaran ke segmen yang paling menjanjikan [5]. Model RFM banyak digunakan karena mudah digunakan dan cepat diimplementasikan di perusahaan, dan selain itu RFM mudah dipahami oleh manajer dan pengambil keputusan pemasaran [6]. Namun model RFM tidak dapat mengsegmentasi pelanggan mana yang memiliki hubungan jangka panjang atau jangka pendek dengan perusahaan dan loyalitas pelanggan dibangun dari hubungan yang panjang, oleh karena itu, dikembangkan model LRFM (Length, Recency, Frequency, Monetary) dengan menambahkan parameter *length* untuk mengatasi permasalahan tersebut [7].

Adapun penelitian sebelumnya yang berjudul *Customers Segmentation in the Insurance Company (TIC) Dataset* [8]. Penelitian ini menyelidiki bagaimana algoritma analitik data yang berbeda yang dapat berguna dalam mempelajari segmentasi pelanggan. Penelitian lain sebelumnya yang berjudul *Customer Segmentation Based on RFM Model Using K-Means, K-Medoids, and DBSCAN Methods*, penelitian ini menganalisis 334.641 data transaksi untuk memperoleh kelompok dan mengidentifikasi calon pelanggan. Namun hasil menunjukkan bahwa waktu yang dibutuhkan dalam pemilihan kepala klaster dan kompleksitas ruang dari tumpukan tindh klaster jauh lebih baik pada K-Medoids daripada K-Means. Dan juga K-Medoids lebih baik dalam hal waktu eksekusi, tidak sensitif terhadap

outlier dibandingkan dengan K-Means karena meminimalkan jumlah ketidaksamaan objek data [9].

Penerapan strategi yang sama terhadap seluruh pelanggan dapat menimbulkan beberapa dampak, seperti strategi promosi yang tidak tepat sasaran sehingga strategi tersebut hanya akan bermanfaat terhadap pelanggan tertentu saja. Selain itu proses analisis data transaksi penjualan membutuhkan waktu yang cukup lama dan variabel yang tepat untuk mendapatkan informasi seperti segmen pelanggan dan nilai pelanggan yang diperlukan dalam menargetkan strategi pemasaran ke segmen yang paling menjanjikan. Banyak metode yang dapat digunakan dalam melakukan proses segmentasi pelanggan, salah satunya adalah penggalian data terhadap riwayat transaksi penjualan pada periode tertentu. Data transaksi tersebut kemudian digunakan untuk menentukan variabel LRFM. Data transaksi yang digunakan adalah periode 1 Januari 2020 sampai 31 Maret 2020. Algoritma *clustering* yang digunakan adalah *k-medoids clustering* berdasarkan variabel LRFM dan CLV.

II. METODE

Metode penelitian berisi langkah - langkah yang akan dilakukan dalam melakukan segmentasi pelanggan pada Artch Indonesia. Tujuan praktis keseluruhan dari tugas *data mining* adalah untuk mengekstrak atau menggali informasi dari sejumlah besar data dan mengubahnya menjadi lebih terstruktur yang dapat dimengerti dan digunakan lebih lanjut [10]. *Cross Industry Standard Process for Data Mining* (CRISP-DM) merupakan standar proses *data mining* yang dikhususkan dalam melakukan proses analisis dari suatu industri atau organisasi sebagai dasar acuan pemecahan masalah dari bisnis [11]. Tahap CRISP-DM diantaranya :

A. Business Understanding

Pada tahap ini akan dilakukan pemahaman tujuan dan kebutuhan dari sudut pandang bisnis. Studi kasus pada penelitian ini dilakukan salah satu perusahaan *fashion* di kota Bandung yaitu Artch yang memproduksi berbagai macam kebutuhan *fashion*.

B. Data Understanding

Dalam tahap ini dilakukan pengumpulan data – data yang diperlukan untuk penelitian yaitu data transaksi penjualan produk ke pelanggan. Data yang telah dikumpulkan yaitu data transaksi penjualan produk ke pelanggan dari 1 Januari 2020 hingga 31 maret 2020. Data transaksi penjualan yang diperoleh memiliki 44 atribut dan 3736 *record data*.

C. Data Preparation

Tahap selanjutnya dilakukan persiapan data sebelum masuk ke *clustering*. persiapan data dilakukan agar data mentah yang telah dikumpulkan dapat diolah sesuai dengan kriteria yang dibutuhkan. Persiapan data meliputi tahap :

1) Data Selection

Dari 44 atribut akan dilakukan *data selection* untuk menseleksi data yang dibutuhkan untuk mengidentifikasi nilai LRFM. Berikut adalah cuplikan data transaksi setelah dilakukan *data selection* ditunjukkan pada Tabel 1.

TABEL 1. DATA SELECTION

No.	Username	Waktu Pesanan	Total Harga
1.	anandaaryatiana	2020-01-01 17:30	Rp 45.000
2.	fachrunnisaaazhar	2020-01-01 18:34	Rp 45.000
3.	navisasyawari25	2020-01-01 14:35	Rp 45.000
4.	yulianiaisyh	2020-01-01 15:25	Rp 50.000
5.	eleneew	2020-01-02 07:37	Rp 160.000
...
3736	gyyoktaviani	2020-03-27 16:37	Rp 99.000

2) Data Cleaning

Selanjutnya data akan dilakukan tahap *data cleaning* seperti ditunjukkan pada Tabel 2.

TABEL 2. DATA CLEANING

No.	Username	Waktu Pesanan	Total Harga
1.	anandaaryatiana	2020-01-01	45000
2.	fachrunnisaaazhar	2020-01-01	45000
3.	navisasyawari25	2020-01-01	45000
4.	yulianiaisyh	2020-01-01	50000
5.	eleneew	2020-01-02	160000
...
3736	gyyoktaviani	2020-03-27	99000

Selanjutnya data akan ditransformasikan menjadi variabel LRFM. Data transaksi pelanggan tidak dapat langsung dijadikan masukan dalam melakukan segmentasi.

3) Data transformation

Pada data *transformation* dilakukan konversi variabel dari data transaksi menjadi variabel LRFM, Model LRFM adalah pengembangan dari RFM yang dikembangkan oleh Arthur Hughes [12]. Penambahan variabel *length* (L) untuk periode kunjungan pertama dan kunjungan terakhir pelanggan tertentu. lalu menjadi model LRFM [13]. Maka setiap pelanggan harus didefinisikan terlebih dahulu nilai LRFM sebelum dilakukan pengelompokkan berdasarkan :

- *Length* (L), Diperoleh dari selisih hari sejak transaksi pertama kali dilakukan oleh setiap pelanggan dengan hari pengerjaan, yaitu tanggal 1 Januari 2021.
- *Recency* (R), Diperoleh dari selisih hari sejak transaksi terakhir kali dilakukan oleh setiap pelanggan dengan hari pengerjaan, yaitu tanggal 1 Januari 2021.
- *Frequency* (F), Diperoleh dari jumlah transaksi yang dilakukan oleh setiap pelanggan.
- *Monetary* (M), Merupakan jumlah pemasukan yang diterima perusahaan dari tiap pelanggan.

Pada *data transformation*, data transaksi diubah menjadi atribut LRFM seperti ditunjukkan pada Tabel 3.

TABEL 3. DATA TRANSFORMATION

No.	Username	Length	Recency	Frequency	Monetary
1.	_amly_	340	340	1	31999
2.	_putriatha261	276	276	1	99000
3.	00ilyansyah	303	303	1	34999
4.	0812955794	278	278	2	69998
5.	10ilmi	343	343	1	130500
...
2866	zzhrraa	316	309	9	500494

Nilai pelanggan akan tersegmentasi menurut 5 *cluster* dengan simbol ↑ merepresentasikan nilai kelompok di atas rata-rata keseluruhan. Sementara simbol ↓ merepresentasikan nilai kelompok di bawah rata-rata keseluruhan yang ditunjukkan pada Tabel 4.

TABEL 4. CLUSTER LOYALITAS PELANGGAN [14]

Customer Clustering	Simbol LRFM	Keterangan
Core Costumers	(LRFM, ↑↑↑↑)	Kelompok pelanggan yang melakukan pembelian dengan interval waktu antara pembelian awal dan pembelian akhir tinggi. Kelompok ini baru saja melakukan pembelian akhir dengan jumlah transaksi dan total biaya yang dibelanjakan seluruhnya lebih tinggi dari rata-rata.
Lost Costumers	(LRFM, ↓↑↑↑)	Kelompok pelanggan yang melakukan pembelian dengan interval waktu antara pembelian awal dan pembelian akhir rendah. Grup ini sudah lama tidak melakukan pembelian, tetapi jumlah transaksi dan total biaya yang dikeluarkan lebih tinggi dari rata-rata.
Consuming Resource Costumer	(LRFM, ↑↓↓↓)	Kelompok pelanggan yang melakukan pembelian dengan interval waktu antara pembelian awal dan pembelian akhir tinggi. Grup ini baru saja melakukan pembelian akhir, namun jumlah total transaksi dan total biaya yang dibelanjakan seluruhnya lebih rendah dari rata-rata.
Uncertain Lost Customers	(LRFM, ↓↓↓↓)	Kelompok pelanggan yang melakukan pembelian dengan interval waktu antara pembelian awal dan pembelian akhir rendah. Grup ini telah lama tidak melakukan pembelian dan jumlah transaksi serta total biaya yang dihabiskan seluruhnya lebih rendah dari rata-rata.
Uncertain New Customers	(LRFM, ↓↓↓↓)	Kelompok pelanggan yang melakukan pembelian dengan interval waktu antara pembelian awal dan pembelian akhir rendah. Kelompok ini baru saja melakukan pembelian dengan jumlah transaksi dan total uang yang dihabiskan seluruhnya lebih rendah dari rata-rata.

Berdasarkan pada klasifikasi nilai LRFM, maka hasil *cluster* memberikan informasi tentang perilaku pembelian pelanggan. Perbedaan pola perilaku pelanggan akan memberikan hasil yang berbeda terhadap keuntungan bagi perusahaan, untuk mengoptimalkan alokasi promosi yang ada maka perusahaan perlu merancang strategi *marketing* yang berbeda pada setiap kelompok pelanggan [15].

Kelompok pelanggan yang terbentuk dari 2 kategori yakni *frequency* dan *monetary* dapat dibagi menjadi 4 segmen yang ditunjukkan pada Tabel 5.

TABEL 5. KATEGORI STRATEGI

Customer Value	High Monetary	Low Monetary
High Frequency	Enforced Strategy	Offensive Strategy
Low Frequency	Defensive Strategy	“Let-go” Strategy

Berikut 4 usulan strategi yang dapat mengelola kelompok yang dihasilkan ditunjukkan pada Tabel 6.

TABEL 6. USULAN STRATEGI MARKETING

Strategi Marketing	Program strategi	Aktivitas Pemasaran
Enforced strategy	Menjaga komunikasi dengan Menjaga interaktif jangka panjang. Merancang program loyalitas pelanggan.	Memberikan informasi promosi melalui telepon, email atau yang lainnya. Memberikan diskon pada event tertentu. Melakukan wawancara untuk ulasan pelayanan.
Defensive strategy	Mengembangkan kegiatan promosi untuk meningkatkan frekuensi. Mengirim informasi produk dan layanan secara berkala.	Menawarkan produk up selling dengan harga khusus. Menawarkan voucher potongan harga.
Offensive strategy	Mempertahankan loyalitas pelanggan dengan <i>cross selling</i> atau <i>up selling</i> . Mengembangkan kegiatan promosi untuk meningkatkan frekuensi.	Mempromosikan produk baru atau produk pelengkap. Menawarkan produk yang paling banyak dibeli.
“Let-go” strategy	Tidak ada keharusan perusahaan untuk memperhatikan segmen ini. Memilih produk utama yang dibutuhkan pelanggan.	Memisahkan pelanggan baru dan pelanggan lama. Melakukan komunikasi dengan pelanggan baru.

4) Data normalization

Selanjutnya data LRFM akan dinormalisasi dengan metode *min – max*. Normalisasi dilakukan bertujuan agar data L, R, F, dan M tidak terlalu jauh sehingga digunakan *range* antara 0 sampai 1 dengan menggunakan metode *min - max*. *Range* 0 sampai 1 digunakan agar nilai normalisasi pada penelitian ini memiliki nilai antara 0 dan 1 [16]. Normalisasi *min - max* dihitung dengan menggunakan persamaan (1).

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Dimana :

- x'_i : Nilai baru hasil normalisasi
- x_i : Nilai awal sebelum dinormalisasi
- x_{min} : Nilai terkecil atribut
- x_{max} : Nilai terbesar

Berikut adalah hasil dari normalisasi dengan metode *min – max* ditunjukkan pada Tabel 7.

TABEL 7. DATA NORMALIZATION

No.	Username	NLength	NRecency	NFreq	NMonetary
1.	_amly_	0.71111	0.71111	0.00000	0.00231
2.	_putriatha	0.00000	0.00000	0.00000	0.00804
3.	00ilyansyah	0.30000	0.30000	0.00000	0.00257
4.	081295579	0.02222	0.02222	0.02857	0.00556
5.	10ilmi	0.74444	0.74444	0.00000	0.01073
...
2866	5875j1ucf8	0.44444	0.36667	0.22857	0.04237

NLength mewakili nilai normalisasi dari variabel *length*, NRecency mewakili nilai normalisasi dari *recency*, NFreq mewakili nilai normalisasi dari *frequency*, dan NMonetary mewakili nilai normalisasi dari *monetary*.

Setelah nilai normalisasi dari LRFM diketahui, selanjutnya akan dilakukan perhitungan CLV. Nilai umur pelanggan didapat dari rentang waktu pelanggan tersebut menggunakan produk atau jasa dari perusahaan. Perhitungan CLV menggunakan persamaan (2).

$$CLV = NL * WL + NR * WR + NF * WF + NM * WM \quad (2)$$

Dimana :

CLV : nilai CLV yang dicari

NL : nilai normalisasi variabel *length*

WL : nilai bobot variabel *length*

NR : nilai normalisasi variabel *recency*

WR : nilai bobot variabel *recency*

NF : nilai normalisasi variabel *frequency*

WF : nilai bobot variabel *frequency*

NM : nilai normalisasi variabel *monetary*

WM : nilai bobot variabel *monetary*

Setelah nilai normalisasi setiap masing – masing atribut L, R, F, dan M diketahui maka dapat dicari nilai CLV. Hasil perhitungan CLV dapat dilihat pada Tabel 8.

TABEL 8. CUSTOMER LIFETIME VALUE

No.	NLength	NRecency	NFreq	NMonetary	CLV
1.	0.711111	0.711111	0.00000	0.00231	557.430
2.	0.00000	0.00000	0.00000	0.00804	795.749
3.	0.30000	0.30000	0.00000	0.00257	271.579
4.	0.02222	0.02222	0.02857	0.00556	401.457
5.	0.74444	0.74444	0.00000	0.01073	1911.13
...
2866	0.44444	0.36667	0.22857	0.04237	21461.4

Pada hasil perhitungan CLV, akan terbentuk nilai setiap pelanggan, semakin besar nilai CLV maka semakin besar pelanggan tersebut memberikan keuntungan pada perusahaan.

D. Modeling

Metode *clustering* yang akan digunakan pada pemodelan penelitian ini yaitu metode K-Medoids. Algoritma K-Medoids memiliki kemampuan untuk mengatasi *noise* dan *outlier*, dimana objek dengan nilai yang besar memiliki kemungkinan untuk menyimpang dari sebaran data [17]. Tahapan-tahapan penyelesaian K-Medoids adalah sebagai berikut [18]:

- 1) Inisialisasikan pusat cluster sebanyak jumlah cluster (k) secara acak.
- 2) Setiap data atau objek dialokasikan ke cluster terdekat menggunakan persamaan ukuran jarak Euclidian Distance dengan persamaan (3).

$$d_{xy} = \sqrt{\sum_{i=1}^n (x_{ik} - x_{ij})^2} \quad (3)$$

Dimana :

d_{xy} : Euclidean distance, Jarak antara objek i dan j

n : Dimensi data yang digunakan

x_{ij} : Koordinat data ke-i

x_{ik} : Koordinat sentroid ke-j

- 1) Pilih objek pada masing - masing cluster secara acak sebagai kandidat medoid baru.
- 2) Hitung jarak objek yang terdapat pada masing - masing cluster dengan calon medoid baru.
- 3) Hitung total simpangan S dengan menghitung jumlah total jarak baru - total jarak lama. Jika didapatkan $S < 0$, maka tukar objek dengan data cluster untuk membuat sekumpulan k objek baru sebagai medoid.
- 4) Ulangi langkah 3 sampai dengan 5 hingga tidak terjadi perubahan medoid, sehingga diperoleh cluster serta anggota cluster masing-masing.

E. Evaluation

Tahap evaluasi dilakukan setelah kelompok - kelompok terbentuk dari hasil K-medoids *clustering*. *Silhouette coefficient* merupakan suatu metode yang digunakan untuk melihat kualitas dari sebuah *cluster*, dimana *cluster* tersebut diukur seberapa baik suatu objek ditempatkan dalam suatu *cluster* [19]. Tahapan – tahapan perhitungan uji kualitas *cluster* menggunakan metode *silhouette coefficient* adalah sebagai berikut [20]:

- 1) Ambil hasil perhitungan setiap cluster dari perhitungan algoritma k-medoids.
- 2) Menghitung rata – rata jarak obyek lainnya yang berada didalam satu cluster dengan persamaan (4).

$$C_a = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (4)$$

Dimana :

C_a : Cluster dalam ke $X_0 \dots X_n$

i : Banyaknya obyek dalam $X_0 \dots X_n$

x , y : koordinat obyek luar

s , y : koordinat centroid luar

Menghitung rata – rata jarak obyek dengan semua obyek lain yang berada pada luar cluster lain menggunakan persamaan berikut, kemudian diambil nilai yang paling kecil atau minimum dengan persamaan (5).

$$C_b = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (5)$$

Dimana :

C_b : Cluster luar ke $X_0 \dots X_n$

i : Banyaknya obyek luar $X_0 \dots X_n$

x , y : koordinat obyek luar

s , y : koordinat centroid luar

Menghitung nilai *silhouette coefficient* dengan persamaan (6).

$$C_b = \frac{(C_b - C_a)}{\max(C_b, C_a)} \quad (6)$$

Dimana :

b : Jarak rata-rata medoid dengan obyek di luar *cluster*

a : Jarak rata-rata antara medoid dengan objek di dalam *cluster*

Bandungkan hasil dari perhitungan silhouette coefficient menggunakan tabel interpretasi Kauffman untuk menunjukkan hasil ukuran cluster yang memiliki rentang dari -1 sampai 1 ditunjukkan pada Tabel 9.

TABEL 9. INTERPRETASI KAUFFMAN

Nilai Silhouette Coefficiente	Interpretasi oleh Kauffman
0,71 – 1,00	Strong Classification
0,51 – 0,70	Good Classification
0,26 – 0,50	Weak Classification
0 – 0,25	Bad Classification

Nilai 1 menunjukkan bahwa nilai *cluster* yang terbentuk merupakan *cluster* yang baik, sedangkan untuk nilai -1 menunjukkan bahwa hasil *cluster* yang terbentuk tidak baik atau buruk.

F. Deployment

Pada tahap ini pengetahuan atau informasi yang telah diperoleh dari tahap sebelumnya akan dilakukan pelaporan.

III. HASIL DAN DISKUSI

Teknik data mining yang digunakan adalah teknik *clustering* dengan mengimplementasikan algoritma k-medoids terhadap 4870 data transaksi pelanggan setelah data melewati proses *data preparation*. Berikut adalah hasil *clustering* menggunakan k-medoids ditunjukkan pada Tabel 10.

TABEL 10. MODELING

No.	C1	C2	C3	Kedekatan	Cluster
1.	1.265490	9815.894	319.2766	1.265	1
2.	239.4264	9577.575	80.96612	80.966	3
3.	284.7445	10101.74	605.1284	284.745	1
4.	154.8670	9971.867	475.2515	154.867	1
5.	1354.815	8462.185	1034.432	1034.432	3
...
2866	20905.088	11088.087	20584.705	11088.08	2

Langkah selanjutnya yaitu mencari serta mendapatkan selisih total *cost medoid* baru dengan total *cost medoid* lama. Jika hasil selisih yang didapatkan kurang dari 0, maka cari *non-medoid* atau medoid baru lagi hingga mendapatkan selisih *cost* antara medoid lama dan medoid baru lebih dari 0. Jika mendapatkan lebih dari 0, maka *clustering* berhenti. Hasil pengurangan *cost* pada percobaan diatas sudah lebih dari 0, maka proses *clustering* dapat berhenti. Berikut adalah hasil selisih yang didapat ditunjukkan pada Tabel 11.

TABEL 11. TOTAL SIMPANGAN

Iterasi	Cost
Medoid lama	15528746.743
Medoid baru	15697391.705
Cost medoid baru – Cost Medoid lama	168644.961

Tahap *evaluation* bertujuan untuk menguji hasil *cluster* sudah optimal atau belum. Pengujian *cluster* dilakukan menggunakan metode *silhouette coefficient* yang telah dibahas sebelumnya. Implementasi terhadap perhitungan *silhouette* pada

beberapa data transaksi penjualan dengan menghitung jarak dalam dan luar *cluster*. Berikut adalah rata – rata jarak obyek berada didalam satu *cluster* dan rata – rata jarak obyek dengan semua obyek lain yang berada pada luar *cluster* ditunjukkan pada Tabel 12.

TABEL 12. SILHOUETTE COEFFICIENT

Cluster	rata – rata jarak dalam cluster	rata – rata jarak luar cluster	Silhouette
1	264.2075535	45594.02806	0.994205216
2	8430.980372	67625.34376	0.875328096
3	2974.522634	30954.45609	0.90390648

Pada hasil perhitungan *silhouette coefficient* didapatkan nilai *silhouette* untuk *cluster* 1 yaitu 0.994, *cluster* 2 yaitu 0.875 dan *cluster* 3 yaitu 0.903 termasuk kategori *strong classification* menurut tabel interpretasi Kauffman.

Selanjutnya hasil *cluster* yang telah diperoleh akan diimplementasikan dalam sebuah laporan agar lebih mudah dipahami oleh bagian *marketing* arch. Hasil perhitungan pada tahap *modeling* menggunakan algoritma k-medoids. Hasil *cluster* ditunjukkan pada Tabel 13.

TABEL 13. DEPLOYMENT

Cluster	Rata – rata Normalisasi				
	Anggota	length	recency	frequency	monetary
1	1162	0.437	0.436	0.003	0.002
2	86	0.529	0.457	0.094	0.049
3	1618	0.546	0.540	0.007	0.009

Berdasarkan tabel loyalitas pelanggan yang telah dibahas sebelumnya *cluster* 1 termasuk kelompok *Uncertain New Customers* dengan simbol $L \downarrow R \downarrow F \downarrow M \downarrow$ yang melakukan pembelian dengan interval waktu antara pembelian awal yaitu 0.437 dan pembelian akhir rendah yaitu 0.436. kelompok ini baru saja melakukan pembelian akhir, namun jumlah total transaksi yaitu 0.003 dan total biaya yang dibelanjakan seluruhnya lebih rendah dari rata - rata yaitu 0.002.

Cluster 2 termasuk kelompok *Lost customer* dengan simbol $L \downarrow R \uparrow F \uparrow M \uparrow$ yang melakukan pembelian dengan interval waktu antara pembelian awal yaitu 0.529 dan pembelian akhir rendah dengan nilai 0.457. Kelompok ini baru saja melakukan pembelian akhir dengan jumlah transaksi yaitu 0.094 dan total biaya yang dibelanjakan seluruhnya lebih tinggi dari rata - rata yaitu 0.049.

Cluster 3 termasuk kelompok *Core customer* dengan simbol $L \uparrow R \uparrow F \downarrow M \uparrow$ yang melakukan pembelian dengan interval waktu antara pembelian awal yaitu 0.546 dan pembelian akhir tinggi yaitu 0.540. kelompok ini sudah lama tidak melakukan pembelian, tetapi jumlah transaksi yaitu 0.007 dan total biaya yang dikeluarkan lebih tinggi dari rata-rata yaitu 0.009.

CLV berguna dalam menentukan nilai profitabilitas dan loyalitas pelanggan dengan model LRFM. Berdasarkan nilai CLV maka rangking tertinggi diantara 3 *cluster* tersebut berada pada *cluster* 2 dengan nilai CLV berada jauh tertinggi dari rata - rata *cluster* lainnya yaitu 172189.650. Rangking kedua yaitu

berada pada *cluster* 3 dengan nilai CLV 1647.070. Ranging ke tiga yaitu pada *cluster* 1 dengan nilai CLV 419.226.

IV. KESIMPULAN

Berdasarkan tabel loyalitas pelanggan dan tabel kategori strategi yang telah dibahas sebelumnya *cluster* 1 termasuk kelompok *Uncertain New Customers* dengan simbol $L \downarrow R \downarrow F \downarrow M \downarrow$ strategi untuk kelompok ini yaitu "*Let-go*" strategy, *Cluster* 2 termasuk kelompok *Lost customer* dengan simbol $L \downarrow R \uparrow F \uparrow M \uparrow$ strategi untuk kelompok ini yaitu *Enforced Strategy*. *Cluster* 3 termasuk kelompok *Core customer* dengan simbol $L \uparrow R \uparrow F \downarrow M \uparrow$ strategi untuk kelompok ini yaitu *Offensive Strategy*. Untuk penelitian selanjutnya, dapat menggunakan metode tertentu untuk menentukan jumlah *cluster* yang optimal.

DAFTAR PUSTAKA

- [1] [1] J. Silva, N. Varela, L. A. B. López, and R. H. R. Millán, "Association rules extraction for customer segmentation in the SMES sector using the apriori algorithm," *Procedia Comput. Sci.*, vol. 151, no. 2018, pp. 1207–1212, 2019, doi: 10.1016/j.procs.2019.04.173.
- [2] [2] B. P. Dewa and B. Setyohadi, "Analisis Dampak Faktor Customer Relationship Management dalam Melihat Tingkat Kepuasan Dan Loyalitas Pada Pelanggan Marketplace Di Indonesia," *Telematika*, vol. 14, no. 01, pp. 33–38, 2017, doi: 10.31315/telematika.v14i01.1964.
- [3] [3] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, "RFM ranking – An effective approach to customer segmentation," *J. King Saud Univ. - Comput. Inf. Sci.*, 2018, doi: 10.1016/j.jksuci.2018.09.004.
- [4] [4] V. L. Miguéis, A. S. Camanho, and J. Falcão E Cunha, "Customer data mining for lifestyle segmentation," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9359–9366, 2012, doi: 10.1016/j.eswa.2012.02.133.
- [5] [5] P. Q. Brito, C. Soares, S. Almeida, A. Monte, and M. Byvoet, "Customer segmentation in a large database of an online customized fashion business," *Robot. Comput. Integr. Manuf.*, vol. 36, pp. 93–100, 2015, doi: 10.1016/j.rcim.2014.12.014.
- [6] [6] I. Maryani, D. Riana, R. D. Astuti, A. Ishaq, Sutrisno, and E. A. Pratama, "Customer segmentation based on RFM model and clustering techniques with K-means algorithm," *Proc. 3rd Int. Conf. Informatics Comput. ICIC 2018*, pp. 1–6, 2018, doi: 10.1109/IAC.2018.8780570.
- [7] [7] J. T. Wei, S. Y. Lin, C. C. Weng, and H. H. Wu, "A case study of applying LRFM model in market segmentation of a children's dental clinic," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 5529–5533, 2012, doi: 10.1016/j.eswa.2011.11.066.
- [8] [8] W. Qadadeh and S. Abdallah, "Customers Segmentation in the Insurance Company (TIC) Dataset," *Procedia Comput. Sci.*, vol. 144, pp. 277–290, 2018, doi: 10.1016/j.procs.2018.10.529.
- [9] [9] P. Arora, Deepali, and S. Varshney, "Analysis of K-Means and K-Medoids Algorithm for Big Data," *Phys. Procedia*, vol. 78, no. December 2015, pp. 507–512, 2016, doi: 10.1016/j.procs.2016.02.095.
- [10] [10] C. Langin, "Introduction to Data Mining," *Scalable Comput. Pract. Exp.*, vol. 9, no. 4, pp. 1–25, 2001, doi: 10.12694/scpe.v9i4.542.
- [11] [11] B. E. Adiana, I. Soesanti, and A. E. Permanasari, "Analisis Segmentasi Pelanggan Menggunakan Kombinasi Rfm Model Dan Teknik Clustering," *J. Terap. Teknol. Inf.*, vol. 2, no. 1, pp. 23–32, 2018, doi: 10.21460/jutei.2018.21.76.
- [12] [12] D. A. Kandeil, A. A. Saad, and S. M. Youssef, "A two-phase clustering analysis for B2B customer segmentation," *Proc. - 2014 Int. Conf. Intell. Netw. Collab. Syst. IEEE INCoS 2014*, pp. 221–228, 2014, doi: 10.1109/INCoS.2014.49.
- [13] [13] R. Ait, A. Amine, B. Bouikhalene, and R. Lbibb, "Customer segmentation model in e-commerce using clustering techniques and LRFM model: The case of online stores in Morocco," *Int. J. Comput. Inf. Eng. World Acad. Sci. Eng. Technol.*, vol. 9, no. 8, pp. 1976–1986, 2015.
- [14] [14] S. Monalisa, "Analysis Outlier Data on RFM and LRFM Models to Determining Customer Loyalty with DBSCAN Algorithm," *Proceeding - 2018 Int. Symp. Adv. Intell. Informatics Revolutionize Intell. Informatics Spectr. Humanit. SAIN 2018*, pp. 1–5, 2019, doi: 10.1109/SAIN.2018.8673380.
- [15] [15] R. H. Sampieri, "Penerapan Customer Relationship Management (Crm) Dengan Menggunakan Metode Lrfm Analisis Di Pt Phapros," p. 634.
- [16] [16] M. Irwan Padli, "Segmentasi Perilaku Pembelian Pelanggan Berdasarkan Model RFM dengan Metode K-Means," vol. 5341, no. April, pp. 9–15, 2018.
- [17] [17] N. R. Maulina, I. Surjandari, and A. M. M. Rus, "Data mining approach for customer segmentation in b2b settings using centroid-based clustering," *2019 16th Int. Conf. Serv. Syst. Manag. ICSSSM 2019*, no. March, pp. 1–6, 2019, doi: 10.1109/ICSSSM.2019.8887739.
- [18] [18] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [19] [19] A. Febri, N. Ningsih, and J. Lemantara, "Aplikasi Analisis Segmentasi Pelanggan untuk Menentukan Strategi Pemasaran Menggunakan Kombinasi Metode K-Means dan Model RFM," *Sistemasi*, vol. 10, no. 1, p. 152, 2021, doi: 10.32520/stmsi.v10i1.1123.
- [20] [20] P. E. Prakasawati, Y. H. Chrisnanto, and A. I. Hadiana, "Segmentasi Pelanggan Berdasarkan Produk Menggunakan Metode K-Medoids," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 335–339, 2019, doi: 10.30865/komik.v3i1.1610.