

ANALISIS SEGMENTASIS PELANGGAN TOKO SEPATU DI *E-COMMERCE* SHOPEE BERDASARKAN MODEL FRM (*RECENCY, FREQUENCY, MONETARY*) MENGGUNAKAN ALGORITMA *K-MEANS* KLAUSTERING

Tety Fera Febriana, M. Imron Rosadi

Teknik Informatika, Universitas Yudharta Pasuruan
Jalan Universitas Yudharta No.7 Sengonagung Purwosari Pasuruan
Tetyfera839@gmail.com

ABSTRAK

Dalam perkembangan teknologi platform e-commerce yang sangat pesat, memudahkan konsumen membeli berbagai produk melalui platform shopee. Berdasarkan permasalahan yang ditemukan yaitu masih banyak pelanggan yang hanya melakukan satu kali pembelian dan tidak kembali untuk melakukan pembelian. Ini menunjukkan rendahnya Tingkat retensi pelanggan. Penelitian ini bertujuan untuk melakukan segmentasi pelanggan toko sepatu di e-commerce shopee berdasarkan model RFM dan algoritma *k-means clustering*. Hasil penelitian menunjukkan bahwa menggunakan kombinasi analisis FRM dan *K-Means* memberikan wawasan tentang karakteristik dan nilai setiap pelanggan, sehingga dapat digunakan untuk merancang strategi pemasaran yang lebih efektif dan meningkatkan retensi pelanggan.

Kata kunci : Segmentasi Pelanggan, Perilaku Pembelian, Strategi Pemasaran, FRM, K-Means

1. PENDAHULUAN

Pelanggan e-commerce merupakan individu atau organisasi yang melakukan pembelian barang atau jasa melalui platform digital. Mereka berinteraksi dengan marketplace untuk memenuhi kebutuhan mereka. Mulai dari kebutuhan sehari-hari hingga produk khusus. Segmentasi pelanggan merupakan proses pembagian pasar atau basis potensi pelanggan ke dalam kelompok berdasarkan karakteristik. Tujuannya untuk mengidentifikasi dan memahami perbedaan kebutuhan dan preferensi pelanggan[1].

Penelitian dalam study ini adalah pelanggan dari sebuah toko sepatu yang beroperasi di platform e-commerce Shopee khususnya di Mojokerto. Toko ini memiliki berbagai jenis produk sepatu. Data yang digunakan adalah data historis pembelian pelanggan yang mencakup tiga komponen utama seperti (Recency kapan terakhir pelanggan melakukan pembelian), (Frequency seberapa sering pelanggan melakukan pembelian), dan (Monetary berapa uang yang dihabiskan pelanggan). Contoh segmentasi pelanggan sering melakukan pembelian maka strategi memberikan penghargaan program loyalitas untuk mempertahankan mereka[2].

Penelitian ini bertujuan untuk melakukan segmentasi pelanggan pada toko sepatu di e-commerce shopee menggunakan model RFM (Recency, Frequency, Monetary) yang dikombinasikan dengan algoritma K-Means klastering untuk mengevaluasi nilai pelanggan berdasarkan perilaku pembelian mereka, dan digunakan untuk mengelompokkan pelanggan ke dalam klaster yang memiliki karakteristik serupa.

Penelitian sebelumnya yang dilakukan oleh [3] menggunakan model RFM untuk segmentasi

pelanggan, namun diterapkan pada online Retail. Penelitian tersebut menunjukkan bahwa penggunaan model RFM efektif dalam mengidentifikasi pelanggan yang berpotensi memberikan nilai tinggi bagi Perusahaan. Penelitian ini memperluas metode algoritma K-Means memungkinkan segmentasi yang telah spesifik pada pelanggan toko sepatu di platform e-commerce shopee.

Metode analisis yang digunakan adalah klasterisasi dengan algoritma K-Means berdasarkan RFM. Data akan diolah dan dikelompokkan ke dalam segmen-segmen pelanggan sesuai karakteristik RFM. Hasil penelitian akan difokuskan pada pengelompokkan pelanggan dan pengembangan strategi pemasaran yang lebih efektif untuk meningkatkan retensi pelanggan[4]

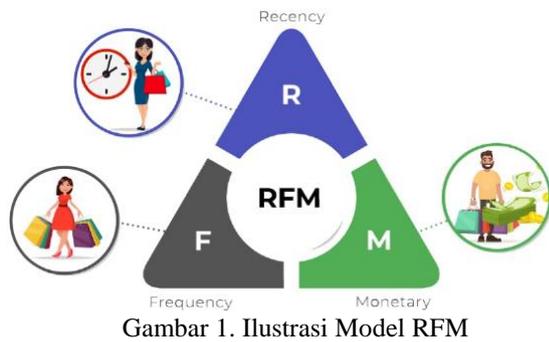
2. TINJAUAN PUSTAKA

2.1. Segmentasi Pelanggan

Segmentasi pelanggan adalah proses membagi basis pelanggan ke dalam kelompok yang lebih kecil, lebih homogen. Berdasarkan berbagai karakteristik yang relevan. Tujuan utama segmentasi pelanggan untuk memungkinkan perusahaan mengembangkan dan menerapkan strategi pemasaran yang lebih efektif dan efisien dengan menargetkan setiap segmen secara spesifik.

Pentingnya segmentasi pelanggan meliputi Personalisasi pemasaran yang lebih relevan bagi setiap segmen pelanggan. Efisiensi pemasaran menggunakan sumber daya pemasaran dengan menargetkan segmen yang paling potensi. Meningkatkan kepuasan pelanggan dengan menawarkan produk atau layanan yang lebih sesuai dengan kebutuhan dan preferensi.

2.2. Definisi Model RFM

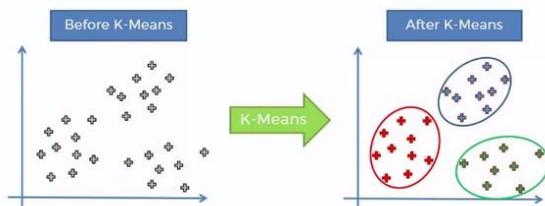


Gambar 1. Ilustrasi Model RFM

Model analisis yang digunakan untuk mengidentifikasi nilai pelanggan berdasarkan 3 dimensi [3] :

- a. *Regency (R)*: mengukur seberapa baru pelanggan melakukan pembelian terakhir. Pelanggan yang baru saja melakukan pembelian dianggap lebih mungkin melakukan pembelian lagi.
- b. *Frequency (F)*: mengukur seberapa sering pelanggan melakukan pembelian dalam periode waktu tertentu. Pelanggan yang lebih sering berbelanja cenderung lebih setia.
- c. *Monetary (M)*: mengukur jumlah uang yang dihabiskan oleh pelanggan. Yang mengeluarkan lebih banyak uang dianggap lebih bernilai.

2.3. Algoritma K-Means



Gambar 2. Diagram Alur Proses Pengolahan Data

Algoritma *clustering* digunakan untuk mengelompokkan data kedalam *K* kluster berdasarkan kesamaan antar data. Algoritma ini bekerja dengan memaksimalkan variasi antar kluster.

2.4. Evaluasi Model

Setelah berhasil menerapkan algoritma k-means untuk melakukan segmentasi pelanggan berdasarkan RFM hasil klustering dianalisis melalui *scatter plot* yang memetakan hubungan antara *recency*, *frequency*, *moentary*, dengan setiap titik data diwarnai menurut kluster masing-masing

3. METODE PENELITIAN

Tahapan untuk melakukan sebuah penelitian agar penelitian dilakukan dengan baik. Tahapan meliputi :

3.1. Study Literatur

Proses sistematis dalam mengumpulkan, mengevaluasi, menganalisa informasi yang telah di terbitkan sebelumnya mengenai segmentasi pelanggan guna mendukung peneliti dan meningkatkan pemahaman peneliti terhadap permasalahan yang ada. Pengetahuan ini diambil dari berbagai jurnal dan *library research*.

3.2. Pengumpulan Data

Data transaksi pelanggan dari toko sepatu di shopee. Sumber data : database internal shopee atau laporan penjualan toko sepatu. Periode waktu 1 tahun terakhir. Proses pengumpulan data yaitu mengunduh atau mengakses data transaksi yang mencakup informasi tentang tanggal pembelian, jumlah transaksi, dan total pengeluaran pelanggan.

3.3. Data selection

Digunakan melalui pilihan data berdasarkan kolom yang tidak relevan. Processing Data dengan menghapus data transaksi yang tidak lengkap. Membuat kolom untuk fitur RFM seperti *CustomerID*, *InvoiceDate*, *Invoice*, *Total Price*. Transformasi data mengelompokkan hasil RFM sesuai hasil normalisasi atau standarisasi dari tiga fitur RFM.

- a. Menerapkan algoritma *k-means* dengan menentukan jumlah kluster optimal menggunakan metode *silhouette score*.
- b. Melakukan Uji Coba pada tahap ini hasil segmentasi pelanggan dihitung nilai akurasi berdasarkan *Silhouette Score* nilai yang tinggi menunjukkan bahwa kluster pembagian yang baik. Dan evaluasi segmentasi pelanggan berdasarkan RFM hasil klustering.

3.4. Analisa

Setelah mendapatkan berbagai informasi yang dibutuhkan dari study literatur selanjutnya tahap analisa yaitu :

3.5. Kebutuhan Data

Pada tahap ini, peneliti mengumpulkan berbagai data yang dibutuhkan. Setelah data terkumpul data diolah dengan model *Recency*, *Frequency*, *Monetary*, dan metode *K-Means* pada segmentasi pelanggan toko sepatu di *marketplace* Shopee.

3.6. Tahapan Alur



4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Data yang digunakan merupakan dataset dari Perusahaan, dataset ini merupakan Kumpulan data seperti ID Transaksi, tanggal pembelian, jumlah produk yang dibeli, dan harga dari tanggal 01/01/2023 hingga 30/12/2023 yang berbasis online di e-commerce. Dari data yang didapat maka penulis dapat menganalisis berdasarkan frekuensi pembelian, nilai pembelian, dan loyalitas pelanggan dan mengidentifikasi pola pembelian berdasarkan preferensi perilaku belajar.

4.2. Pengolahan Data

Data preparation tiga hal yang mendasar yang dilakukan pada tahap data preparation, antara lain :

- a. Kolom yang tidak relevan dengan metode yang akan digunakan akan dihilangkan ketika data sekunder dipilih.
- b. Setelah tahap pilihan data, data sekunder akan melalui proses preprocessing data, yang akan membuat data yang digunakan bersih dan bising dan tidak memiliki nilai yang hilang. Data akan disortir berdasarkan nilai jumlah untuk menghilangkan kolom yang memiliki nilai nol.

```
df_data.isnull().sum()
```

```
Customer ID    0
Username       0
Invoice        0
InvoiceDate    0
Quantity       0
Price          0
dtype: int64
```

- c. Membuat kolom baru untuk Fitur RFM berdasarkan *customer ID*, *InvoiceDate*, *Invoice*, dan *totalprice*.

	Customer ID	InvoiceDate	Invoice	TotalPrice
0	1	535	1	700.296
1	2	235	5	973.560
2	3	549	1	263.892
3	4	468	1	103.620
4	5	523	1	241.120

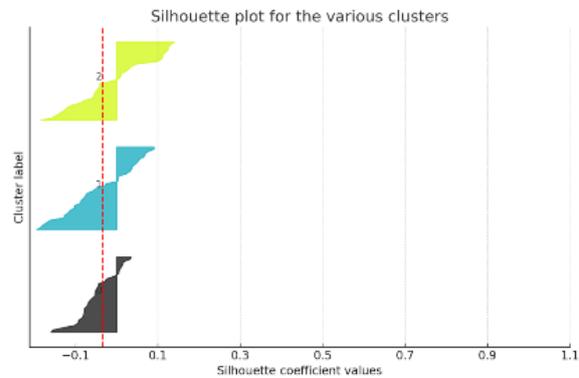
- d. Setelah data yang akan digunakan, nilai diubah dengan mengelompokkan dan memberi bobot pada masing-masing kolom frekuensi, frekuensi, dan uang. Hasil penyesuaian atribut RFM untuk frekuensi, frekuensi, dan uang ditunjukkan dalam tabel 1.

	Recency	Frequency	Monetary
0	1.372286	-0.672763	0.988090
1	-1.582746	2.903639	1.330688
2	1.465754	-0.672763	-0.028249
3	0.890565	-0.672763	-1.001895
4	1.293197	-0.672763	-0.123011

Data diatas adalah hasil normalisasi atau standarisasi dari tiga fitur dalam model RFM

- e. Modeling K-Means

Untuk membentuk model klastering menggunakan algoritma *k-means* maka, menentukan jumlah klaster optimal dengan menggunakan metode *silhouette score*.



Gambar 3. Hasil Klastering Pelanggan Berdasarkan Algoritma K-Means

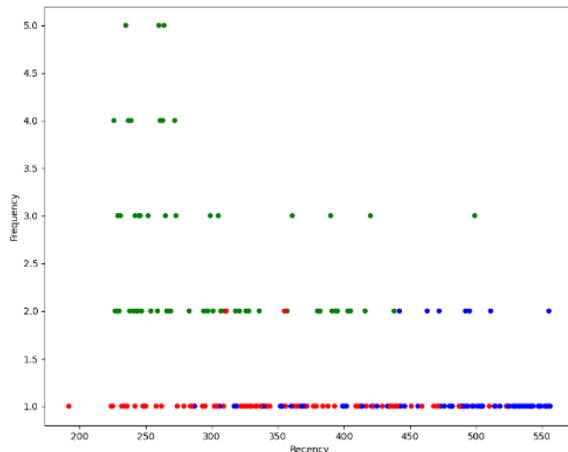
Dari diagram diatas menunjukkan klaster 2 kualitas klaster terbaik dengan nilai *silhouette score* yang paling tinggi, sedangkan klaster 1 menunjukkan klaster yang baik dengan nilai *silhouette score* positif, tetapi tidak setinggi klaster 2, dan klaster 0 memiliki beberapa titik dengan nilai *silhouette score* negative menunjukkan bahwa ada beberapa titik yang lebih dekat dengan klaster lain.

- f. Menerapkan algoritma dengan *k-means* dengan jumlah kluster yang optimal.

CustomerID	Recency	Frequency	Monetary	R	F	M	RFMGroup	RFMScore	RFM_Loyalty_Level	Cluster	
0	1	535	1	700.296	4	4	1	441	9	Gold	2
1	2	235	5	973.560	1	1	1	111	3	Platinum	1
2	3	549	1	263.892	4	4	3	443	11	Silver	2
3	4	468	1	103.620	3	4	4	344	11	Silver	0
4	5	523	1	241.120	4	4	3	443	11	Silver	2

Pelanggan platinum merupakan pelanggan berharga dengan pembelian yang tinggi, pembelian baru-baru ini, dan pengeluaran yang tinggi. Pelanggan gold merupakan pelanggan yang menghabiskan banyak uang tetapi jarang melakukan pembelian dan mungkin perlu diaktifkan Kembali. Pelanggan silver merupakan pelanggan dengan nilai yang rendah tidak terlalu sering berbelanja, dan tidak menghabiskan banyak

- uang. Dan merupakan target yang baik untuk kampanye reaktivitas atau peningkatan frekuensi pembelian.
- g. Evaluasi setelah berhasil menerapkan algoritma k-means untuk melakukan segmentasi pelanggan berdasarkan RFM hasil kluastering dianalisis melalui *scatter plot* yang memetakan hubungan antara *recency*, *frequency*, *moentary*, dengan setiap titik data diwarnai menurut kluster masing-masing



Gambar 4. Scatter Plot Hubungan Recency, Frequency, dan Monetary

Berdasarkan metrik RFM menunjukkan bahwa ada perbedaan yang jelas dalam pola pembelian d antara kluster-kluster yang terbentuk dengan warna hijau berisi pelanggan yang sering bertransaksi dan bernilai tinggi, yang merupakan target utama untuk upaya retensi dan loyalitas. Sebaliknya, kluster dengan warna biru dan merah berisi pelanggan dengan transaksi sedang atau rendah, memerlukan strategi yang tepat berdasarkan interpretasi ini untuk membantu perusahaan dalam meningkatkan nilai pelanggan.

```
rfm.Cluster.value_counts()
Cluster
0    70
2    67
1    58
Name: count, dtype: int64
```

Berdasarkan hasil tersebut memiliki 3 segmen untuk pelanggan kategori platinum sebanyak 52 pelanggan, gold sebanyak 67 pelanggan, silver sebanyak 58 pelanggan.

Berdasarkan hasil klastering terlihat bahwa platinum memberikan kontribusi terbesar terhadap pendapatan toko, meskipun jumlahnya paling sedikit sehingga bisa menawarkan layanan pelanggan lebih personal, pengiriman gratis, atau diskon, eksklusif bisa menjadi langkah penting untuk mempertahankan loyalitas mereka. Kluster gold meskipun pengeluaran tinggi, perlu ditargetkan dengan strategi yang mendorong pelanggan untuk lebih sering belanja dan

memotivasi mereka untuk lebih aktif. Kluster Silver memerlukan perhatian lebih walaupun nilai transaksi rendah mungkin masih memiliki potensi untuk Kembali pelanggan aktif. Program pemasaran berbasis E-mail yang menawarkan promosi atau produk baru dengan harga yang lebih terjangkau untuk meningkatkan daya minat pelanggan,

5. KESIMPULAN DAN SARAN

Analisis pengelompokan dari data transaksi pelanggan *e-commerce* menghasilkan 3 cluster dari 297 data pelanggan berdasarkan segmentasi pelanggan kategori gold sebanyak 67 pelanggan, platinum sebanyak 52 pelanggan dan kategori silver sebanyak 58 pelanggan. Pelanggan kategori emas memiliki 67 pelanggan, menurut data yang dikumpulkan. Hasil penelitian menunjukkan bahwa metode RFM dan algoritma K-Means digunakan untuk segmentasi pelanggan. Dengan kata lain, data transaksi penjualan digunakan. Pelanggan kategori gold memiliki potensi besar, namun memerlukan strategi yang dapat meningkatkan frekuensi belanja pada pelanggan, sedangkan kategori silver terdiri dari pelanggan dengan transaksi rendah yang memerlukan kampanye reaktivasi. Oleh karena itu, manfaat dari penggunaan k-means untuk menentukan kluster adalah bahwa silhouette score membantu dalam menentukan kluster yang paling ideal. Dengan cara ini, perusahaan dapat lebih mudah mengidentifikasi demografi pelanggan berdasarkan kelompoknya dan dapat menerapkan strategi pemasaran yang berbeda untuk setiap kluster.

DAFTAR PUSTAKA

[1] P. Indra Pangestu, T. Iman Hermanto, and D. Irmayanti, "Analisis Segmentasi Pelanggan Berbasis Recency Frequency Monetary (Rfm) Menggunakan Algoritma K-Means," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1486–1492, 2023, doi: 10.36040/jati.v7i3.7171.

[2] A. Febriani and S. A. Putri, "Segmentasi Konsumen Berdasarkan Model Recency, Frequency, Monetary dengan Metode K-Means," *Jiems (Journal Ind. Eng. Manag. Syst.*, vol. 13, no. 2, pp. 52–57, 2020, doi: 10.30813/jiems.v13i2.2274.

[3] A. Rachman Hakim and H. Soetanto, "Penerapan Algoritma K-Means Menggunakan Model Recency Frequency Monetary Dalam Klasterisasi Penjualan Produk Sepatu," *Semin. Nas. Mhs. Fak. Teknol. Inf. Jakarta-Indonesia*, no. September, pp. 699–706, 2022, [Online]. Available: <https://senafiti.budiluhur.ac.id/index.php/>

[4] D. Wiga Maulana Baihaqi, "Penerapan Teknik Clustering Sebagai Strategi Pemasaran pada Penjualan," vol. XXI, no. 2, pp. 243–248, 2019, doi: 10.31294/p.v20i2.