

Segmentasi Pelanggan Toko EMAS Menggunakan Metode RFM dan K-MEANS Clustering

Dandi Ari Setiarno¹, Luqman Hakim²

Program Studi Teknik Industri, Fakultas Teknik, Universitas Nahdlatul Ulama Sidoarjo
Jl. Lingkar Timur KM 5,5, Rangkah Kidul, Kec. Sidoarjo, Kabupaten Sidoarjo, Jawa Timur 61234
Email: dandiari75@unusida.ac.id, hqm_ft@unusida.ac.id

ABSTRAK

Dalam dunia bisnis yang semakin kompetitif, perusahaan harus bisa menggunakan data pelanggan dengan baik agar bisa membuat keputusan yang lebih baik. Toko Emas XYZ di Sidoarjo memiliki banyak data transaksi dari pelanggan, tetapi data tersebut belum digunakan sepenuhnya untuk menganalisis strategi, terutama dalam membagi pelanggan berdasarkan segmen mereka. Penelitian ini bertujuan menganalisis sifat-sifat pelanggan dan membagi pelanggan berdasarkan pola belanjanya dengan menggunakan metode RFM yaitu Recency, Frequency, dan Monetary, serta metode K-Means Clustering. Hasil penelitian diharapkan bisa memberikan pemilahan pelanggan yang jelas dan rapi, seperti pelanggan prioritas, pelanggan yang masih potensial, pelanggan biasa, dan pelanggan yang tidak aktif lagi. Segmentasi ini bisa jadi dasar dalam membuat strategi pemasaran, meningkatkan kualitas layanan, serta mengelola hubungan dengan pelanggan secara lebih efektif dan efisien.

Kata kunci: Segmentasi Pelanggan, Analisa RFM Pelanggan Loyal, K-Means Clustering.

ABSTRACT

In an increasingly competitive business world, companies must be able to utilize customer data effectively to make better decisions. XYZ Gold Store in Sidoarjo has a wealth of customer transaction data, but this data has not been fully utilized for strategic analysis, particularly in segmenting customers. This study aims to analyze customer characteristics and segment customers based on their shopping patterns using the RFM method (Recency, Frequency, and Monetary) and the K-Means Clustering method. The research results are expected to provide a clear, precise classification of customers, including priority, potential, regular, and inactive customers. This segmentation can serve as a basis for developing marketing strategies, improving service quality, and managing customer relationships more effectively and efficiently.

Keywords: Customer Segmentation, Loyal Customer RFM Analysis, K-Means Clustering.

Pendahuluan

Perkembangan dunia bisnis yang semakin kompetitif menuntut perusahaan untuk mampu mengelola sumber daya secara efektif dan efisien, termasuk dalam memanfaatkan data pelanggan sebagai dasar pengambilan keputusan. Dalam konteks Teknik Industri, pengelolaan data pelanggan menjadi bagian penting dalam upaya meningkatkan kinerja sistem bisnis secara keseluruhan, khususnya dalam bidang pemasaran dan manajemen hubungan pelanggan. Data transaksi yang dimiliki perusahaan dapat memberikan informasi berharga mengenai perilaku pelanggan apabila diolah dengan metode yang tepat. [1], [2], [3]

Industri perdagangan emas merupakan salah satu sektor yang memiliki karakteristik unik, di mana produk tidak hanya berfungsi sebagai barang konsumsi tetapi juga sebagai instrumen investasi. Perilaku pembelian pelanggan pada sektor ini dipengaruhi oleh berbagai faktor seperti fluktuasi harga emas, tren pasar, serta preferensi individu. Oleh karena itu, pemahaman terhadap karakteristik pelanggan menjadi sangat penting bagi perusahaan agar mampu menyusun strategi pemasaran yang tepat sasaran dan meningkatkan loyalitas pelanggan[4].

Toko Emas XYZ Sidoarjo sebagai objek penelitian memiliki data transaksi pelanggan dalam jumlah yang cukup besar, yang mencakup informasi mengenai waktu transaksi, frekuensi pembelian, serta nilai transaksi. Namun, data tersebut belum dimanfaatkan secara optimal dan masih terbatas pada fungsi administratif. Akibatnya, perusahaan belum memiliki segmentasi pelanggan yang jelas, sehingga strategi pemasaran yang diterapkan cenderung bersifat umum dan kurang efektif. Hal ini berpotensi menyebabkan inefisiensi dalam alokasi sumber daya serta kurang optimalnya pelayanan kepada pelanggan.

Untuk mengatasi permasalahan tersebut, diperlukan suatu pendekatan analitis yang mampu mengelompokkan pelanggan berdasarkan perilaku pembelian. Salah satu metode yang dapat digunakan adalah model Recency, Frequency, Monetary (RFM), yang mampu menggambarkan tingkat aktivitas dan nilai pelanggan berdasarkan data historis transaksi. [5],

[6]Metode ini menilai pelanggan berdasarkan tiga dimensi utama, yaitu waktu terakhir transaksi, frekuensi pembelian, dan nilai transaksi.

Selanjutnya, untuk melakukan pengelompokan pelanggan secara objektif, digunakan metode K-Means Clustering, yaitu salah satu teknik dalam data mining yang bertujuan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik. Kombinasi metode RFM dan K-Means Clustering diharapkan mampu menghasilkan segmentasi pelanggan yang lebih akurat dan mudah diinterpretasikan[7], [8], [9].

Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining untuk melakukan segmentasi pelanggan pada Toko Emas XYZ Sidoarjo. Data yang digunakan berupa data transaksi pelanggan dalam periode tertentu yang meliputi informasi tanggal transaksi, frekuensi pembelian, dan total nilai transaksi.[10]

Tahap pertama penelitian dilakukan dengan pengumpulan dan preprocessing data, meliputi pembersihan data dan pengelompokan transaksi pelanggan. Selanjutnya dilakukan perhitungan nilai RFM (Recency, Frequency, Monetary), di mana Recency dihitung berdasarkan selisih waktu sejak transaksi terakhir, Frequency berdasarkan jumlah transaksi, dan Monetary berdasarkan total nilai pembelian pelanggan[11], [12].

Data RFM yang diperoleh kemudian dinormalisasi untuk menghindari perbedaan skala antar variabel. Proses berikutnya adalah pengelompokan data menggunakan algoritma K-Means Clustering dengan penentuan jumlah cluster menggunakan metode Elbow[13].

Tahapan terakhir berupa evaluasi dan interpretasi hasil clustering untuk mengidentifikasi karakteristik masing-masing segmen pelanggan. Hasil segmentasi ini digunakan sebagai dasar dalam penyusunan strategi pemasaran yang lebih efektif dan tepat sasaran. [14], [15], [16]

Hasil Dan Pembahasan

Analisis Recency, Frequency, Monetary (RFM).

Penelitian ini menggunakan data transaksi pelanggan Toko Emas XYZ Sidoarjo selama periode satu tahun, yaitu Januari 2025 hingga Januari 2026. Data yang diolah meliputi tanggal transaksi terakhir, jumlah transaksi, serta total nilai pembelian pelanggan. Data tersebut kemudian diolah menggunakan metode *Recency, Frequency, Monetary* (RFM) untuk mengetahui karakteristik perilaku pelanggan. [17]

Table 1. Analisis recency, frequency, monetary

Pembeli	Recency (Hari)	Frequency	Monetary (RP)
AG	10	5	13,500,000
AH	3	14	51,000,000
AM	3	24	102,000,000
AN	40	16	67,500,000
AS	44	4	10,500,000
BM	50	7	28,500,000
CC	44	8	22,500,000
ET	50	12	37,500,000
FG	86	2	6,000,000
FZ	227	2	7,500,000
HA	40	10	33,000,000
HK	58	15	43,500,000
HY	3	13	40,500,000
IN	33	12	36,000,000
KY	241	3	6,000,000
MF	3	24	79,500,000
MJ	73	3	12,000,000
NG	148	5	21,000,000
NK	0	3	12,000,000
NR	17	17	54,000,000
SF	36	16	52,500,000
SR	40	17	48,000,000
TY	163	2	7,500,000
WT	0	15	169,500,000

XR	230	2	4,500,000
XY	150	2	3,000,000
YZ	318	2	4,500,000
ZA	17	17	55,500,000

Data pada tabel menunjukkan karakteristik pelanggan berdasarkan tiga variabel utama dalam metode *Recency, Frequency, Monetary (RFM)*, yaitu *recency* (jarak waktu sejak transaksi terakhir dalam hari), *frequency* (jumlah transaksi), dan *monetary* (total nilai pembelian dalam rupiah). Setiap pelanggan diberi kode unik, seperti AG, AH, AM, dan seterusnya, untuk memudahkan identifikasi dalam proses analisis.

Berdasarkan variabel *recency*, terlihat bahwa terdapat variasi waktu transaksi terakhir yang cukup signifikan antarpelanggan. Beberapa pelanggan seperti Tabel 1 Skor RFM WT memiliki nilai *recency* yang sangat rendah (0 hari), yang menunjukkan bahwa pelanggan tersebut baru saja melakukan transaksi. Sebaliknya, pelanggan seperti YZ dan KY memiliki nilai *recency* yang tinggi, yaitu di atas 200 hari, yang mengindikasikan bahwa pelanggan tersebut sudah lama tidak melakukan transaksi. Hal ini menunjukkan adanya perbedaan tingkat keaktifan pelanggan dalam berinteraksi dengan toko.

Dari sisi *frequency*, terlihat bahwa terdapat pelanggan dengan tingkat transaksi yang tinggi, seperti AM dan MF yang masing-masing memiliki frekuensi transaksi sebesar 24 kali. Hal ini menunjukkan loyalitas pelanggan yang tinggi. Di sisi lain, terdapat pula pelanggan dengan frekuensi rendah, seperti FG, FZ,

Skor RFM

Skor hasil perhitungan *Recency, Frequency, Monetary (RFM)* dapat dilihat pada table 2. [18]

Table 2. Skor hasil perhitungan *Recency, Frequency, Monetary (RFM)*

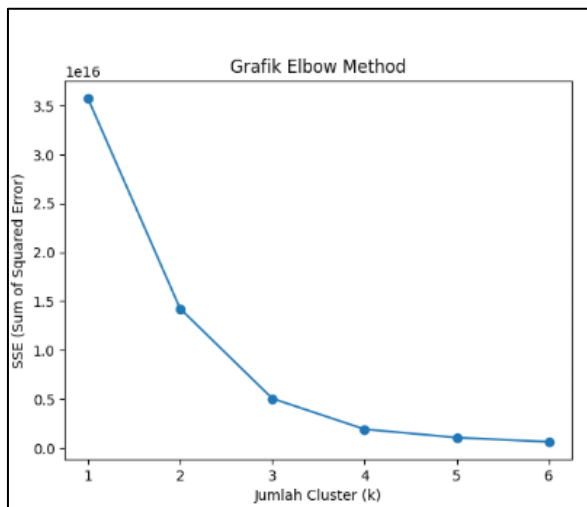
Pembeli	R_score	F_score	M_score	RFM
AG	4	2	2	422
AH	5	4	4	544
AM	5	5	5	555
AN	3	4	5	345
AS	3	2	2	322
BM	3	3	3	333
CC	3	3	3	333
ET	3	3	3	333
FG	2	1	1	211
FZ	1	1	1	111
HA	3	3	3	333
HK	2	4	4	244
HY	5	4	4	544
IN	4	3	3	433
KY	1	2	1	121
MF	5	5	5	555
MJ	2	2	2	222
NG	2	3	3	233
NK	5	2	2	522
NR	4	5	5	455
SF	4	5	4	454
SR	3	5	4	354
TY	1	1	1	111
WT	5	4	5	545
XR	1	1	1	111
XY	1	1	1	111
YZ	1	1	1	111
ZA	4	5	5	455

Keterangan:

1. Blok tabel warna merah merupakan hasil dari total penjumlahan RFM adalah nama-nama pelanggan yang mempunyai nilai tertinggi.
2. Blok tabel warna orange merupakan hasil dari total penjumlahan RFM adalah nama-nama pelanggan yang mempunyai nilai tingkat Tengah.

3. Sedangkan yang tidak diberi warna (netral) pada tabel total skor RFM merupakan nama pelanggan yang mempunyai nilai terendah.

Elbow



Gambar 1 Elbow

Metode Elbow digunakan untuk menentukan jumlah cluster optimal dengan melihat nilai Sum of Squared Error (SSE) pada setiap variasi jumlah cluster. Berdasarkan hasil perhitungan, terjadi penurunan SSE yang signifikan dari k=1 hingga k=3. Namun, setelah k=3, penurunan nilai SSE cenderung melambat dan tidak signifikan.[19]

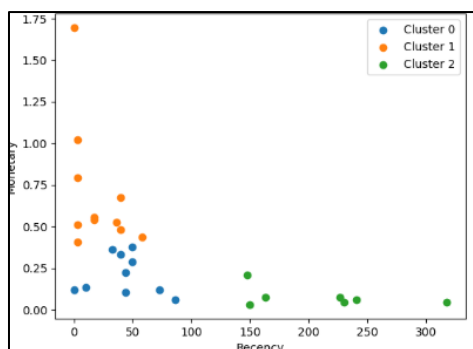
Hal ini menunjukkan bahwa penambahan jumlah cluster setelah k=3 tidak memberikan peningkatan yang berarti dalam kualitas pengelompokan data. Oleh karena itu, dapat disimpulkan bahwa jumlah cluster optimal dalam penelitian ini adalah sebanyak tiga cluster (k=3), yang mampu merepresentasikan segmentasi pelanggan secara efektif berdasarkan nilai Recency, Frequency, dan Monetary (RFM).

K-Means Clustering

Berdasarkan analisis menggunakan metode *K-Means* pada RFM, pelanggan Toko Emas XYZ terbagi menjadi tiga kelompok. [20]

Tabel 2 Clustering

Pembeli	RFM	Cluster	Pembeli	RFM	Cluster
AH	544	1	SF	454	2
AM	555	1	SR	354	2
HY	544	1	ZA	455	2
MF	555	1	AS	322	3
NK	522	1	FG	211	3
WT	545	1	FZ	111	3
AG	422	2	HK	244	3
AN	345	2	KY	121	3
BM	333	2	MJ	222	3
CC	333	2	NG	233	3
ET	333	2	TY	111	3
HA	333	2	XR	111	3
IN	433	2	XY	111	3
NR	455	2	YZ	111	3



Gambar 2 Clustering

Berdasarkan hasil pengolahan data menggunakan metode *Recency, Frequency, Monetary (RFM)* yang dilanjutkan dengan proses *clustering*, diperoleh tiga kelompok pelanggan (*cluster*) dengan karakteristik yang berbeda-beda. Pengelompokan ini bertujuan untuk mengidentifikasi nilai dan perilaku pelanggan sehingga dapat menjadi dasar dalam pengambilan keputusan strategis.

Cluster 1 merupakan kelompok pelanggan dengan nilai RFM tertinggi, yang terdiri dari pelanggan seperti AH, AM, HY, MF, NK, dan WT. Nilai RFM pada cluster ini berada pada rentang tinggi (di atas 500), yang menunjukkan bahwa pelanggan dalam kelompok ini memiliki tingkat aktivitas transaksi yang tinggi, frekuensi pembelian yang sering, serta nilai pembelian yang besar. Dengan demikian, pelanggan pada cluster ini dapat dikategorikan sebagai pelanggan loyal atau *high-value customers*. Kelompok ini memiliki kontribusi terbesar terhadap pendapatan perusahaan, sehingga perlu dipertahankan melalui strategi seperti pemberian loyalitas program, diskon eksklusif, atau layanan prioritas.

Cluster 2 menunjukkan pelanggan dengan nilai RFM menengah, seperti AG, AN, BM, CC, ET, HA, dan lainnya. Nilai RFM pada kelompok ini berada pada kisaran sedang, yang mencerminkan bahwa pelanggan masih cukup aktif namun tidak seintensif cluster 1. Kelompok ini dapat dikategorikan sebagai pelanggan potensial (*potential customers*) yang masih memiliki peluang untuk meningkatkan loyalitasnya. Strategi yang dapat diterapkan adalah promosi berkala, penawaran khusus, serta peningkatan engagement agar pelanggan dapat berpindah ke cluster dengan nilai yang lebih tinggi.

Cluster 3 merupakan kelompok pelanggan dengan nilai RFM terendah, seperti FG, FZ, HK, KY, MJ, NG, TY, XR, XY, dan YZ. Nilai RFM pada cluster ini relatif rendah, yang menunjukkan bahwa pelanggan memiliki frekuensi pembelian yang jarang, nilai transaksi yang kecil, serta waktu transaksi terakhir yang sudah lama. Kelompok ini dapat dikategorikan sebagai pelanggan tidak aktif atau *low-value customers*. Strategi yang dapat dilakukan adalah melakukan reaktivasi melalui promosi khusus, diskon besar, atau kampanye pemasaran ulang (*remarketing*). Namun, apabila tidak menunjukkan peningkatan, perusahaan dapat mempertimbangkan untuk tidak memprioritaskan kelompok ini dalam alokasi sumber daya pemasaran.

Centroid

Tabel 3 Centroid

Clustering	R	F	M
Cluster 1	5,00	4,00	4,17
Cluster 2	3,45	3,73	3,64
Cluster 3	1,55	1,73	1,64

Cluster 1 (Pelanggan Bernilai Tinggi).

Cluster 1 memiliki nilai *centroid* tertinggi dibandingkan dengan cluster lainnya, dengan skor rata-rata *Recency* sebesar 5,00, *Frequency* sebesar 4,00, dan *Monetary* sebesar 4,17 Hal ini menunjukkan bahwa pelanggan dalam cluster ini melakukan transaksi relatif baru, memiliki frekuensi pembelian yang cukup sering, serta nilai pembelian yang tinggi. Pelanggan pada cluster ini merupakan pelanggan utama perusahaan yang memberikan kontribusi besar terhadap pendapatan, sehingga perlu dipertahankan melalui pelayanan prioritas, serta program loyalitas khusus.

Cluster 2 (Pelanggan Bernilai Menengah).

Cluster 2 memiliki nilai *centroid* menengah dengan skor *Recency* sebesar 3,45, namun skor *Frequency* dan *Monetary* relatif lebih rendah dibandingkan dengan *Cluster 1*. Kondisi ini menunjukkan bahwa pelanggan dalam *cluster* ini masih tergolong aktif melakukan pembelian, tetapi jumlah transaksi dan nilai pembelannya belum maksimal. Pelanggan pada *cluster* ini memiliki potensi untuk meningkatkan nilainya melalui strategi pemasaran seperti promosi, diskon pembelian, atau penawaran bundling produk agar frekuensi dan nilai transaksinya meningkat.

Cluster 3 (Pelanggan Bernilai Rendah).

Cluster 3 memiliki nilai centroid terendah dengan skor *Recency*, *Frequency*, dan *Monetary* yang relatif kecil. Hal ini menandakan bahwa pelanggan pada cluster ini jarang melakukan transaksi, sudah cukup lama tidak melakukan pembelian, serta memiliki nilai pembelian yang rendah. Pelanggan dalam cluster ini tergolong pelanggan kurang aktif, sehingga perusahaan perlu melakukan evaluasi lebih lanjut apakah pelanggan ini masih layak dipertahankan atau cukup diberikan pelayanan standar tanpa prioritas khusus. [21]

Menghitung Jarak Setiap Data Ke Centroid Dengan Menggunakan Jarak Eucliden

Tabel 4 Eucliden

Cluster	Jarak_C1	Jarak_C2	Jarak_C3	Cluster_Final
2	3.11591078	2.446425965	2.49098374	Jarak_C2
1	0.17	1.614001239	4.7565744	Jarak_C1
1	1.29957685	2.421776208	5.82108237	Jarak_C1
2	2.1653868	1.457737974	4.3063906	Jarak_C2
3	3.56495442	2.425901894	1.51822265	Jarak_C3
2	2.52366797	1.070046728	2.35902522	Jarak_C2
2	2.52366797	1.070046728	2.35902522	Jarak_C2
2	2.52366797	1.070046728	2.35902522	Jarak_C2
3	5.29612122	4.065095325	1.07004673	Jarak_C3
3	5.92021114	4.519402615	1.11579568	Jarak_C3
2	2.52366797	1.070046728	2.35902522	Jarak_C2
3	3.00481281	1.518222645	3.30529878	Jarak_C2
1	0.17	1.614001239	4.7565744	Jarak_C1
2	1.83545635	1.11579568	3.07652401	Jarak_C2
3	5.4816877	3.995622605	0.88600226	Jarak_C3
1	1.29957685	2.421776208	5.82108237	Jarak_C1
3	4.20819439	2.790161286	0.6363961	Jarak_C3
3	3.37177995	1.744992837	1.91441897	Jarak_C2
1	2.95108455	2.843413442	3.47922405	Jarak_C2
2	1.63978657	1.940360791	5.29008507	Jarak_C1
2	1.42439461	1.430034965	4.71858029	Jarak_C1
2	2.2425209	1.394632568	4.28544047	Jarak_C2
3	5.92021114	4.519402615	1.11579568	Jarak_C3
1	0.83	2.079663434	5.32400225	Jarak_C1
3	5.92021114	4.519402615	1.11579568	Jarak_C3
3	5.92021114	4.519402615	1.11579568	Jarak_C3
3	5.92021114	4.519402615	1.11579568	Jarak_C3
2	1.63978657	1.940360791	5.29008507	Jarak_C1

Setelah diperoleh nilai *centroid* akhir dari masing-masing cluster, langkah selanjutnya adalah menghitung jarak setiap data pelanggan terhadap seluruh *centroid*. Perhitungan jarak ini bertujuan untuk menentukan kedekatan masing-masing pelanggan terhadap pusat cluster sehingga dapat ditentukan kelompok yang paling sesuai. [22]

Metode yang digunakan dalam penelitian ini adalah *Euclidean Distance*, yaitu metode pengukuran jarak berdasarkan garis lurus dalam ruang multidimensi. Karena penelitian ini menggunakan tiga variabel, yaitu *Recency* (R), *Frequency* (F), dan *Monetary* (M), maka perhitungan jarak dilakukan dalam tiga dimensi.

Rumus *Euclidean Distance* adalah sebagai berikut:

$$[d(x, y) = \sqrt{(R_i - R_c)^2 + (F_i - F_c)^2 + (M_i - M_c)^2}]$$

Dimana:

R_i, F_i, M_i = nilai skor RFM pelanggan ke-i.

R_c, F_c, M_c = nilai centroid pada cluster tertentu.

Setiap pelanggan akan memiliki tiga nilai jarak:

Jarak ke Centroid 1.

Jarak ke Centroid 2.

Jarak ke Centroid 3.

Nilai jarak terkecil menunjukkan cluster yang paling dekat dengan karakteristik pelanggan tersebut.

Proses perhitungan jarak ini dilakukan pada setiap iterasi algoritma K-Means. Setelah semua data dikelompokkan berdasarkan jarak terdekat, maka centroid akan diperbarui dengan menghitung rata-rata nilai RFM

dari anggota cluster tersebut. Proses ini akan terus diulang hingga tidak terjadi perubahan signifikan pada nilai centroid atau komposisi cluster.

Dengan demikian, perhitungan jarak Euclidean merupakan inti dari proses K-Means karena menentukan bagaimana data dikelompokkan berdasarkan kedekatan karakteristik RFM masing-masing pelanggan.

Simpulan

Berdasarkan hasil penelitian yang telah dilakukan pada data transaksi pelanggan Toko Emas XYZ Sidoarjo periode Januari 2025 hingga Januari 2026, dapat disimpulkan bahwa metode Recency, Frequency, Monetary (RFM) mampu mengidentifikasi karakteristik perilaku pelanggan secara efektif berdasarkan waktu terakhir transaksi, intensitas pembelian, dan nilai transaksi. Hasil perhitungan RFM kemudian diolah menggunakan metode K-Means Clustering dan menghasilkan tiga kelompok pelanggan yang memiliki karakteristik berbeda.

Cluster 1 merupakan kelompok pelanggan bernilai tinggi dengan rata-rata nilai Recency sebesar 5,00, Frequency sebesar 4,00, dan Monetary sebesar 4,17, yang menunjukkan bahwa pelanggan dalam kelompok ini aktif bertransaksi, memiliki frekuensi pembelian tinggi, serta memberikan kontribusi pendapatan terbesar bagi perusahaan. Cluster 2 merupakan kelompok pelanggan bernilai menengah dengan tingkat aktivitas yang cukup baik namun kontribusi transaksi masih dapat ditingkatkan. Sementara itu, Cluster 3 merupakan kelompok pelanggan bernilai rendah dengan frekuensi dan nilai transaksi yang relatif kecil serta cenderung tidak aktif.

Dengan demikian, penerapan metode RFM dan K-Means Clustering terbukti mampu menghasilkan segmentasi pelanggan yang jelas dan terstruktur. Hasil segmentasi ini dapat digunakan sebagai dasar dalam pengambilan keputusan strategis, seperti pemberian layanan prioritas kepada pelanggan bernilai tinggi, pengembangan strategi promosi untuk meningkatkan pelanggan menengah, serta evaluasi terhadap pelanggan bernilai rendah. Implementasi segmentasi ini diharapkan dapat meningkatkan efektivitas strategi pemasaran dan mendukung peningkatan kinerja bisnis secara keseluruhan.

Daftar Pustaka

- [1] T. W. Ningsih *Et Al.*, "Metode Clustering Dengan Model Rfm Dan Algoritma K-Means Untuk Segmentasi Pelanggan Transjakarta," Vol. 9, No. 5, Pp. 7898–7905, 2025.
- [2] S. Ben Salem, S. Naouali, And Z. Chtourou, "The Drk-M For Clustering Categorical Datasets With Uncertainty," *IEEE Intell. Syst.*, Vol. 36, No. 5, Pp. 113–121, 2021, Doi: 10.1109/MIS.2020.3038837.
- [3] F. Y. Cao, L. Q. Yu, J. Z. X. Huang, And J. Y. Liang, "K-Mw-Modes: An Algorithm For Clustering Categorical Matrix-Object Data," *Appl. Soft Comput.*, Vol. 57, Pp. 605–614, 2017, Doi: 10.1016/J.Asoc.2017.04.019.
- [4] Y. Q. Zhang And Y. M. Cheung, "A New Distance Metric Exploiting Heterogeneous Interattribute Relationship For Ordinal-And-Nominal-Attribute Data Clustering," *IEEE Trans. Cybern.*, Vol. 52, No. 2, Pp. 758–771, 2022, Doi: 10.1109/TCYB.2020.2983073.
- [5] K. Kim, "A Weighted K-Modes Clustering Using New Weighting Method Based On Within-Cluster And Between-Cluster Impurity Measures," *Journal Of Intelligent & Fuzzy Systems*, Vol. 32, No. 1, Pp. 979–990, 2017, Doi: 10.3233/JIFS-16157.
- [6] A. G. Oskouei, M. A. Balafar, And C. Motamed, "FKMAWCW: Categorical Fuzzy K-Modes Clustering With Automated Attribute-Weight And Cluster-Weight Learning," *Chaos Solitons Fractals*, Vol. 153, 2021, Doi: 10.1016/J.Chaos.2021.111494.
- [7] Z. N. Jiang And X. Y. Liu, "A Novel Consensus Fuzzy K-Modes Clustering Using Coupling DNA-Chain-Hypergraph P System For Categorical Data," *PROCESSES*, Vol. 8, No. 10, 2020, Doi: 10.3390/Pr8101326.
- [8] X. N. Gao And S. Wu, "CUBOS: An Internal Cluster Validity Index For Categorical Data," *Tehnicki Vjesnik-Technical Gazette*, Vol. 26, No. 2, Pp. 486–494, 2019, Doi: 10.17559/TV-20190109015453.
- [9] Z. G. Chen, H. S. Kang, S. N. Yin, And S. R. Kim, "An Efficient Privacy Protection In Mobility Social Network Services With Novel Clustering-Based Anonymization," *EURASIP J. Wirel. Commun. Netw.*, 2016, Doi: 10.1186/S13638-016-0767-1.
- [10] L. W. Peng And Y. G. Liu, "Attribute Weights-Based Clustering Centres Algorithm For Initialising K-Modes Clustering," *Cluster Computing-The Journal Of Networks Software Tools And ApplicationS*, Vol. 22, Pp. S6171–S6179, 2019, Doi: 10.1007/S10586-018-1889-5.
- [11] K. S. Dorman And R. Maitra, "An Efficient K-Modes Algorithm For Clustering Categorical Datasets," *Stat. Anal. Data Min.*, Vol. 15, No. 1, Pp. 83–97, 2022, Doi: 10.1002/Sam.11546.

- [12] F. Yuan, Y. L. Yang, And T. T. Yuan, “A Dissimilarity Measure For Mixed Nominal And Ordinal Attribute Data In K-Modes Algorithm,” *Applied Intelligence*, Vol. 50, No. 5, Pp. 1498–1509, 2020, Doi: 10.1007/S10489-019-01583-5.
- [13] A. Saha And S. Das, “Categorical Fuzzy K-Modes Clustering With Automated Feature Weight Learning,” *Neurocomputing*, Vol. 166, Pp. 422–435, 2015, Doi: 10.1016/J.Neucom.2015.03.037.
- [14] J. Teknologi, E. Febrianty, L. Awalina, And W. I. Rahayu, “Optimalisasi Strategi Pemasaran Dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering Pada Transaksi Online Retail Optimizing Marketing Strategies With Customer Segmentation Using K-Means Clustering On Online Retail Transactions,” Vol. 13, No. September, Pp. 122–137, 2023, Doi: 10.34010/Jati.V13i2.
- [15] T. Faouzi, L. Firinguetti-Limone, J. M. Avilez-Bozo, And R. Carvajal-Schiaffino, “The A -Groups Under Condorcet Clustering,” *Mathematics*, Vol. 10, No. 5, 2022, Doi: 10.3390/Math10050718.
- [16] R. J. Kuo, Y. R. Zheng, And T. P. Q. Nguyen, “Metaheuristic-Based Possibilistic Fuzzy K-Modes Algorithms For Categorical Data Clustering,” *Inf. Sci. (N. Y.)*, Vol. 557, Pp. 1–15, 2021, Doi: 10.1016/J.Ins.2020.12.051.
- [17] M. I. Zamzani, A. Andini, K. Rahayu, And N. Muhasdi, “Analisis Segmentasi Pelanggan UMKM Toko Kue Albynha Menggunakan Metode RFM Dan K-Means,” Vol. 2, No. 1, Pp. 1–11, 2026.
- [18] N. Wakhidah, “Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering)”.
- [19] D. Maulida And M. Dachyar, “Customer Loyalty Strategy Using Customer Lifetime Value (A Case Study Of Jamu And Herbal Products SME),” No. Clv, Pp. 1123–1135, 2022.
- [20] A. Aranta, “Analisis Pemilihan Cluster Optimal Dalam Segmentasi,” Vol. 18, No. 2, Pp. 152–163, 2021.
- [21] M. Method, F. Indriyani, And E. Irfiani, “Clustering Data Penjualan Pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means (Clustering Sales Data At Outdoor Equipment Stores Using K-,” Vol. 7, No. November, Pp. 109–113, 2019.
- [22] A. D. Savitri, F. A. Bachtiar, And N. Y. Setiawan, “Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan (Studi Kasus : Belle Crown Malang),” Vol. 2, No. 9, Pp. 2957–2966, 2018.