



Segmentasi Pelanggan Toko Hanifah Berdasarkan Analisis RFM dengan Metode *K-Means Clustering*

Yerri Kurnia Febrina ¹, Riyan Saputra ², Katrina Flomina G ³

¹ Manajemen Informatika, Teknologi Informasi, Politeknik Negeri Padang

² Magister Teknik Informatika, Ilmu Komputer, Universitas Putra Indonesia "YPTK" Padang

¹ Teknik Komputer, Teknologi Informasi, Politeknik Negeri Padang

¹yerrikurnia@pnp.ac.id, ²alamat.emailnya.riyan@gmail.com, ³katrina@pnp.ac.id

Abstract

Sales are the crucial aspect in a business because they directly affect revenue and market competitiveness. In the retail context, a uniform sales approach is often ineffective due to the diversity of customer characteristics and shopping behavior. This study aims to segment customers at Hanifah Store based on their purchasing behavior. The method used is Recency, Frequency, Monetary (RFM) analysis, followed by classification using the K-Means Clustering algorithm. The dataset, obtained from a point-of-sale system, consists of 2,183 unique customer records. After standardization and clustering processes, customers were successfully grouped into four clusters: regular, passive, loyal, and potential customers. The clustering results were validated using Principal Component Analysis (PCA) which demonstrated a clear separation between clusters. The findings of this study provide practical insights for formulating more efficient and targeted sales strategies, and for enhancing marketing effectiveness based on the specific characteristics of each customer group. This approach also illustrates the potential of leveraging transactional data to support business decision-making in the retail sector.

Keywords: data mining, k-means clustering, RFM, customer segmentation

Abstrak

Penjualan merupakan aspek krusial dalam bisnis karena secara langsung memengaruhi pendapatan dan daya saing di pasar. Dalam konteks ritel, pendekatan penjualan yang seragam sering kali kurang efektif mengingat keragaman karakteristik dan perilaku belanja pelanggan. Penelitian ini bertujuan untuk mengelompokkan pelanggan Toko Hanifah berdasarkan perilaku belanja mereka. Metode yang digunakan adalah analisis *Recency, Frequency, Monetary* (RFM), yang kemudian diproses lebih lanjut menggunakan algoritma *K-Means Clustering*. Dataset diperoleh dari sistem *point-of-sale* dengan total 2.183 data pelanggan unik. Setelah proses standarisasi dan pengelompokan, pelanggan berhasil diklasifikasikan ke dalam empat klaster: pelanggan reguler, pasif, loyal, dan potensial. Validasi visual dilakukan dengan *Principal Component Analysis* (PCA) yang menunjukkan sebaran klaster yang jelas. Hasil penelitian ini memberikan kontribusi praktis dalam merumuskan strategi penjualan yang lebih efisien dan terarah, serta meningkatkan efektivitas pemasaran berdasarkan karakteristik masing-masing kelompok pelanggan. Pendekatan ini juga menunjukkan potensi pemanfaatan data transaksi dalam mendukung pengambilan keputusan bisnis di sektor ritel.

Kata kunci: penambangan data; pengelompokan *k-means*; RFM; segmentasi pelanggan;

© 2025 Jurnal Pustaka AI

1. Pendahuluan

Dalam sebuah bisnis, penjualan merupakan aspek yang sangat penting karena mempengaruhi pendapatan dan daya saing pasar [1]. Meningkatkan penjualan adalah tantangan beragam yang dapat diatasi melalui berbagai strategi [2]. Dengan melakukan segmentasi pelanggan dapat secara signifikan meningkatkan strategi penjualan dengan mengarahkan prospek ke kelompok penjualan yang sesuai berdasarkan segmen pelanggan yang jelas [3], [4]. Orientasi pelanggan berdampak positif pada penjualan, ini menekankan pentingnya mengelola hubungan dengan pelanggan [5].

Toko Hanifah adalah tempat perbelanjaan kebutuhan harian yang didatangi pelanggan untuk mencari berbagai kebutuhan rumah tangga, mulai dari bahan pokok seperti beras, minyak, dan gula, hingga makanan ringan, minuman kemasan, serta produk-produk kosmetik dan perawatan tubuh. Pelanggannya berasal dari latar belakang yang beragam, dengan kebutuhan dan frekuensi belanja yang tidak selalu sama. Dalam kondisi tersebut, pendekatan penjualan yang seragam cenderung kurang efektif karena tidak mampu menjawab kebutuhan dan potensi dari masing-masing kelompok pelanggan secara optimal [6].

Meskipun Toko Hanifah telah mengumpulkan data transaksi secara digital melalui sistem *point-of-sale*, informasi ini belum dimanfaatkan untuk menganalisis perilaku pelanggan secara sistematis. Di tengah persaingan yang semakin ketat dengan toko modern dan platform daring, pemanfaatan data historis pelanggan menjadi sangat penting [7]. Tanpa adanya pemetaan karakteristik pelanggan, toko berisiko kehilangan peluang untuk meningkatkan loyalitas dan efisiensi strategi penjualannya [8]. Situasi ini menuntut adanya pendekatan berbasis data yang mampu menjawab kebutuhan toko secara lebih akurat dan terukur.

Segmentasi pelanggan merupakan salah satu solusi strategis yang dapat diterapkan dalam konteks ini. Dengan mengelompokkan pelanggan secara strategis, toko dapat menentukan prioritas pendekatan, misalnya dengan memperkuat hubungan dengan pelanggan yang aktif dan setia, memberi perhatian khusus pada pelanggan yang masih memiliki potensi besar, serta mengevaluasi cara berinteraksi dengan pelanggan yang kurang terlibat [9]. Pendekatan ini tidak hanya meningkatkan efisiensi dalam strategi penjualan, tetapi juga membantu toko dalam mengambil keputusan yang lebih tepat sasaran [10].

Untuk menjawab tantangan tersebut, diperlukan metode analisis yang mampu mengidentifikasi kelompok pelanggan berdasarkan nilai transaksional mereka. Salah satu pendekatan yang relevan adalah analisis RFM (*Recency, Frequency, Monetary*), yang mengukur perilaku konsumsi pengguna berdasarkan riwayat pembelian mereka dengan memberi peringkat pelanggan [11] berdasarkan seberapa baru mereka melakukan pembelian (*Recency*), seberapa sering mereka membeli (*Frequency*), dan berapa banyak uang yang mereka belanjakan (*Monetary*) [12]. Dengan data tersebut, pelanggan dapat dikelompokkan secara lebih objektif menggunakan metode pengelompokan seperti algoritma *K-Means Clustering* [13], yang memungkinkan pengelompokan berdasarkan pola yang ditemukan dalam data [14].

Penelitian sebelumnya, berjudul “Analisis Segmentasi Pelanggan pada Bisnis dengan Menggunakan Metode *K-Means Clustering* pada Model Data RFM” berhasil menyegmentasikan pelanggan menjadi tiga kelompok berbeda menggunakan algoritma *K-Means Clustering*, *Elbow Method*, *Silhouette Scores*, dan *Davies-Bouldin Index* digunakan, mengarah pada kesimpulan bahwa tiga cluster optimal [15].

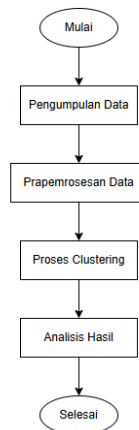
Penelitian dengan judul “*Customer Segmentation with Clustering Algorithm Based on Recency, Frequency, and Monetary (RFM) Attributes*” melakukan segmentasi pelanggan berdasarkan atribut *Recency, Frequency* dan *Monetary* (RFM) menggunakan algoritma *K-Means Clustering* berdasarkan perilaku pembelian mereka. Penelitian berhasil mengimplementasikan model RFM untuk menganalisis perilaku pelanggan. Penelitian ini menggunakan berbagai metode evaluasi, termasuk *Elbow Method*, *Silhouette Scores*, dan *Davies-Bouldin Index*, untuk menentukan jumlah segmen pelanggan yang optimal [16].

Penelitian sebelumnya, berjudul “*Customer Segmentation Using the Integration of the RecencyFrequency Monetary Model and the K-Means Cluster Algorithm*” berhasil menunjukkan bahwa menggabungkan model RFM dengan algoritma *K-Means Clustering*, ditingkatkan dengan metode *Elbow*, memberikan pendekatan yang efektif dan optimal untuk segmentasi pelanggan di ritel. Tiga kelompok yang diidentifikasi, terutama segmen bernilai tinggi, menawarkan wawasan yang dapat ditindaklanjuti bagi bisnis untuk menyesuaikan strategi layanan mereka dan mempertahankan pelanggan penting [17].

Dengan menerapkan kombinasi analisis RFM dan algoritma *K-Means Clustering*, penelitian ini bertujuan untuk mengelompokkan pelanggan Toko Hanifah berdasarkan perilaku belanja mereka. Pengelompokan ini diharapkan dapat memberikan dasar bagi perumusan strategi penjualan yang lebih efisien dan terarah, serta meningkatkan efektivitas pemasaran dengan pendekatan yang sesuai untuk masing-masing kelompok pelanggan. Selain itu, hasil penelitian ini juga diharapkan menjadi kontribusi praktis bagi pelaku usaha ritel dalam memanfaatkan data transaksi untuk mendukung pengambilan keputusan bisnis.

2. Metode Penelitian

Penelitian ini dilakukan melalui beberapa tahapan untuk mendapatkan hasil yang terbaik. Pendekatan dengan data mining menggunakan algoritma *K-Means Clustering* digunakan untuk pengelompokan pelanggan berdasarkan data transaksional mereka di Toko Hanifah. Berikut adalah gambaran alur penelitian yang dilalui:



Gambar 1. Alur Penelitian

Berdasarkan tahapan yang digambarkan pada Gambar 1, masing-masing tahapan dapat diartikulasikan sebagai berikut:

2.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini terdiri dari catatan transaksional yang berasal dari toko Hanifah dari bulan Agustus 2024 sampai dengan Juni 2025. Informasi tersebut bersumber dari aplikasi *point-of-sale* yang diterapkan di toko Hanifah.

2.2. Prapemrosesan Data

Fase prapemrosesan data dilakukan untuk menyiapkan dataset untuk pemrosesan selanjutnya selama fase pengelompokan, yang mencakup identifikasi metrik RFM (*Recency*, *Frekuensi*, dan *Moneter*) dan standarisasi data melalui penerapan teknik *Z-Score Standardization* untuk menjamin keseragaman dan koherensi skala data, dengan rumus sebagai berikut:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

dimana:

z = nilai z-score
 x = nilai asli
 μ = rata-rata
 σ = simpangan baku (standard deviation)

2.3. Proses Clustering

Proses *clustering* pelanggan dalam penelitian ini menggunakan algoritma *K-Means Clustering*, yang diakui sebagai salah satu teknik pembelajaran tanpa pengawasan yang sering digunakan di bidang analisis *cluster*. Identifikasi jumlah *cluster* yang optimal dicapai melalui penerapan *Elbow Method*, yang selanjutnya ditingkatkan dengan evaluasi komparatif metrik *Silhouette Score* untuk menjamin pemilihan kuantitas *cluster* yang menghasilkan pemisahan dan kekompakan maksimal di antara *cluster*. Implementasi algoritma *K-Means Clustering* dilakukan melalui beberapa tahapan sebagai berikut:

1. Menentukan jumlah *cluster* (K)

2. Inisialisasi *centroid* awal, ditetapkan dengan melalui proses seleksi manual, di mana n titik data pertama dari dataset digunakan sebagai representasi dasar untuk setiap *cluster* masing-masing.
3. Perhitungan jarak antar titik data dan *centroid* dengan menggunakan *Euclidian Distance*, dengan persamaan sebagai berikut:

$$d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (2)$$

dimana:

p, q = dua titik dalam ruang n *Euclidean*
 q_i, p_i = vektor *Euclidean*, dimulai dari titik awal
 n = n -ruang

4. Kelompokkan data ke *centroid* terdekat
5. Hitung ulang *centroid* dengan rata-rata setiap klaster
6. Ulangi Langkah 3-5 hingga *centroid* tidak berubah lagi

Dalam praktik nya, proses *clustering* akan dilakukan menggunakan bahasa pemrograman Python yang dijalankan di Google Colab. Hasil dari proses *clustering* akan digambarkan kedalam bentuk grafik sebar (*scatter plot*) dengan memanfaatkan teknik PCA (*Principal Component Analysis*) untuk memberikan gambaran dan hasil yang mudah dipahami.

2.4. Analisis Hasil

Hasil dari proses clustering yang dilakukan adalah berupa data yang telah dikelompokkan kedalam klaster, hasil ini kemudian dapat dimanfaatkan dalam memilih prioritas pendekatan kepada pelanggan.

3. Hasil dan Pembahasan

3.1. Pengumpulan Data

Dataset pada penelitian ini didapatkan dari aplikasi point-of-sale yang diterapkan di toko Hanifah. Data tersebut merupakan catatan transaksi dari Agustus 2024 – Juni 2025 yang berjumlah 85.670. Seluruh transaksi telah diolah dan diakumulasikan berdasarkan masing-masing pelanggan, sehingga menghasilkan 2.183 baris data pelanggan unik. Setiap baris dalam dataset merepresentasikan satu pelanggan dan mencakup informasi seperti pada Tabel 1 dan Tabel 2.

Tabel 1. Datatype dari dataset

#	Column	DType
0	customer	object
1	tanggal_transaksi	datetime64[ns]
2	nilai_transaksi	int64
3	frekuensi_transaksi	int64

Tabel 2. Ringkasan dari dataset

#	customer	tanggal_transaksi	nilai_transaksi	frekuensi_transaksi
0	BNI	2024-10-23 17:28:16	30000	2
1	BOY	2024-10-13 13:14:47	339500	10
2	BNA	2023-09-09 17:50:08	2232000	24
3	AANG	2024-10-22 18:49:40	1020500	32
4	ABUL	2024-10-10 04:02:23	26000	1
...	...	...	...	...
2179	ZODI	2024-08-30 22:33:26	18000	1
2180	ZUIN	2024-10-04 04:54:33	17000	1

2181	ZURI	2024-09-05 09:23:11	21000	1
2182	ZUIA	2024-09-02 22:50:17	99000	3
2183	ZULL	2024-09-01 03:39:50	120000	4

3. 2. Prapemrosesan Data

Dalam tahap prapemrosesan data, fase pertama yang dilakukan adalah menentukan nilai metrik RFM dari dataset yang telah dikumpulkan sebelumnya. Analisis RFM dilakukan pada saat penelitian ini berlangsung di tanggal 09 Juni 2025, yang menghasilkan metrik RFM seperti Tabel 3 berikut:

Tabel 3. Dataset setelah melalui proses analisis RFM

#	customer	recency	frequency	monetary
0	BNI	228	2	30000
1	BOY	238	10	339500
2	BNA	272	24	2232000
3	AANG	229	32	1020500
4	ABUL	241	1	26000
...	...	...	...	...
2179	ZODI	282	1	18000
2180	ZUIN	247	1	17000
2181	ZURI	276	1	21000
2182	ZUIA	279	3	99000
2183	ZULL	280	4	120000

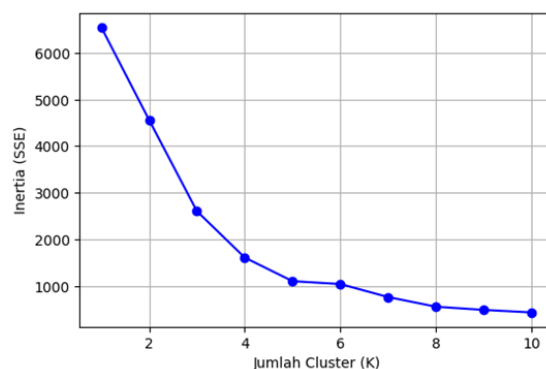
Setelah melalui analisis RFM, dataset tersebut akan melalui proses standarisasi dengan metode *Z-Score Standardization*. Upaya ini dilakukan karena fitur yang memiliki skala besar akan mendominasi hasil clustering. Hasil yang didapatkan setelah melakukan standarisasi dapat dilihat pada Tabel 4 dibawah ini.

Tabel 4. Dataset setelah melalui *Z-Score Standardization*

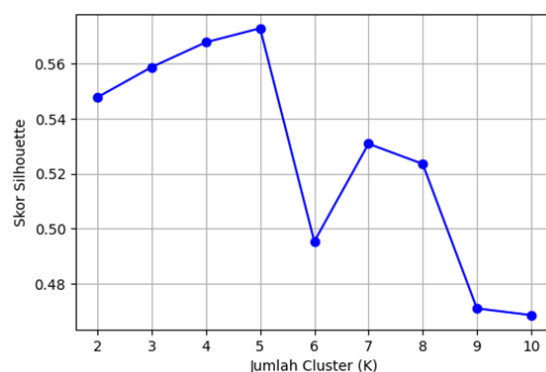
#	customer	recency	frequency	monetary
0	BNI	0.978317	-0.358352	-0.214438
1	BOY	1.088907	-0.205448	-0.127633
2	BNA	1.464915	0.062135	0.403158
3	AANG	0.989376	0.215039	0.063368
4	ABUL	1.122084	-0.377465	-0.215560
...	...	...	...	...
2179	ZODI	1.575506	-0.377465	-0.217804
2180	ZUIN	1.188439	-0.377465	-0.218084
2181	ZURI	1.509151	-0.377465	-0.216962
2182	ZUIA	1.542329	-0.339239	-0.195086
2183	ZULL	1.553388	-0.320126	-0.189196

3. 3. Proses Clustering

Dalam metode *K-Means Clustering* hal pertama kali yang harus dilakukan yaitu menentukan jumlah kluster yang ingin dibentuk. Dalam penelitian ini, untuk menentukan jumlah kluster yang optimal, dilakukan pendekatan dengan menggunakan metode *Elbow* dan *Shilhouette*. Kedua metode ini dilakukan guna untuk mendapatkan perbandingan yang terbaik dalam menentukan jumlah kluster.



(a)



(b)

Gambar 2. Grafik sebar hasil analisis (a) metode *Elbow* (b) metode *Silhouette*

Gambar 2(a) memperlihatkan hasil analisis menggunakan metode *Elbow*, yang menunjukkan bahwa pada jumlah kluster sebanyak 5, terjadi penurunan nilai yang tidak signifikan dibandingkan dengan jumlah kluster sebelumnya. Hal ini mengindikasikan bahwa jumlah kluster yang optimal berdasarkan pendekatan *Elbow* adalah 4. Sementara itu, Gambar 2(b) menampilkan hasil analisis menggunakan metode *Silhouette*, yang menunjukkan bahwa konfigurasi dengan 5 kluster menghasilkan nilai rata-rata koefisien *Silhouette* tertinggi, sehingga secara teknis dianggap sebagai jumlah kluster yang paling optimal. Namun demikian, dengan mempertimbangkan bahwa konfigurasi 4 kluster masih menunjukkan performa yang cukup baik pada grafik sebar koefisien *Silhouette*, serta mengacu pada hasil metode *Elbow*, maka dalam penelitian ini diputuskan untuk menggunakan 4 kluster sebagai konfigurasi akhir dalam proses pengelompokan data.

Proses clustering dengan algoritma *K-Means Clustering* dilakukan dengan menggunakan bahasa pemrograman Python dan Google Colab. Hasil dari clustering membagi dataset menjadi 4 kluster dimana kluster 0 terdiri dari 1298 pelanggan, kluster 1 terdiri dari 867 pelanggan, kluster 2 terdiri dari 2 pelanggan, dan kluster 3 terdiri dari 17 pelanggan. Setiap anggota dari kluster tersebut dapat dilihat pada tabel dibawah.

Tabel 5. Anggota kluster 0

#	customer	recency	frequency	monetary
0	BNI	228	2	30000
1	BOY	238	10	339500
2	BNA	272	24	2232000
3	AANG	229	32	1020500
4	ABUL	241	1	26000
...	...	...	...	...
1293	ZODI	282	1	18000
1294	ZUIN	247	1	17000

1295	ZURI	276	1	21000
1296	ZUIA	279	3	99000
1297	ZULL	280	4	120000

Tabel 6. Anggota klaster 1

#	customer	recency	frequency	monetary
0	M002	5	63	5739000
1	M006	9	40	792000
2	M007	111	24	1683500
3	M008	5	165	3608500
4	M009	33	12	178000
...	...	...	...	...
862	M131	6	2	36000
863	M133	6	1	26000
864	M135	6	8	296000
865	M136	6	2	41000
866	M137	5	2	66000

Tabel 7. Anggota klaster 2

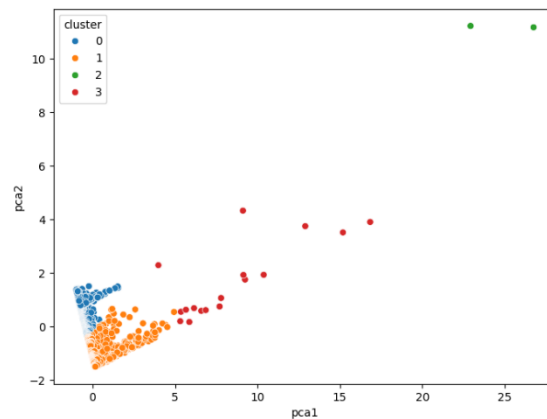
#	customer	recency	frequency	monetary
0	BHER	230	644	85331000
1	M134	6	497	112030500

Tabel 8. Anggota klaster 3

#	customer	recency	frequency	monetary
0	BEYL	228	458	22061000
1	M004	5	921	26824500
2	M024	43	313	7461000
3	M063	6	359	10203500
4	M091	5	279	8818000
...	...	...	...	...
12	M047	10	810	25716000
13	M102	5	435	39698000
14	M109	6	377	23082000
15	M184	5	464	8837000
16	MAEL	227	262	7607000

3. 4. Analisis Hasil

Sebaran dari data pelanggan yang telah diproses akan divisualisasikan kedalam grafik PCA (*Principal Component Analysis*) dengan mereduksi dimensi tanpa menghilangkan komponen yang penting. Visualisasi ini dilakukan untuk mendapatkan pemahaman yang lebih terhadap hasil segmentasi oleh algoritma *K-Means Clustering*.



Gambar 3. Visualisasi PCA

Visualisasi ini menunjukkan bahwa metode pengelompokan berhasil mengidentifikasi karakteristik unik dari setiap segmen pelanggan. Temuan penting dalam penelitian ini adalah:

1. Klaster 2 (VIP): meskipun hanya terdiri dari dua pelanggan, namun nilai transaksi dan frekuensinya sangat tinggi (contoh: pelanggan dengan frekuensi 644 kali dan nilai lebih dari 85 juta). Ini menandakan bahwa sebagian kecil pelanggan memberikan kontribusi signifikan terhadap pendapatan toko.
2. Klaster 0 dan 1 (reguler dan pasif): bersama-sama mencakup 99,1% populasi pelanggan. Segmentasi ini mengindikasikan potensi optimalisasi pendekatan pemasaran untuk pelanggan pasif agar dapat naik kelas menjadi pelanggan aktif atau loyal.
3. Klaster 3 (potensial): terdiri dari pelanggan dengan volume transaksi tinggi namun belum konsisten, menunjukkan peluang besar dalam strategi retensi melalui promosi atau program loyalitas.

Dengan hasil ini, Toko Hanifah dapat mengarahkan promosi dengan pendekatan yang lebih presisi, misalnya dengan menawarkan reward khusus kepada pelanggan klaster 2 dan 3, serta membina engagement terhadap pelanggan di klaster 1.

4. Kesimpulan

Implementasi algoritma K-Means Clustering dalam mensegmentasi pelanggan berdasarkan nilai metrik RFM berhasil mengelompokkan 2.183 pelanggan ke dalam 4 klaster yang memiliki karakteristik berbeda. Keberhasilan ini ditunjukkan secara kuantitatif melalui:

1. Visualisasi PCA yang memperlihatkan sebaran klaster yang terpisah secara jelas.
2. Identifikasi pelanggan prioritas, dengan klaster 2 yang hanya berisi 2 pelanggan namun mencakup total nilai transaksi di atas 197 juta rupiah (lebih dari 10% total nilai).
3. Proporsi klaster yang tidak seimbang mengindikasikan kebutuhan penyesuaian strategi marketing: Klaster 0 (59,4%), Klaster 1 (39,7%), Klaster 2 (0,09%), dan Klaster 3 (0,78%).

Dengan hasil tersebut, penelitian ini telah menjawab permasalahan pada latar belakang mengenai perlunya pendekatan diferensiasi pelanggan di Toko Hanifah. Rekomendasi berbasis klaster yang dihasilkan dari penelitian ini dapat digunakan untuk menyusun strategi penjualan yang lebih efisien dan berbasis data, seperti pemberian insentif pada klaster potensial dan loyal, serta strategi reaktivasi untuk pelanggan pasif.

Daftar Rujukan

- [1] M. Marolt, H. D. Zimmermann, and A. Pucihar, "Social Media Use and Business Performance in SMEs: The Mediating Roles of Relational Social Commerce Capability and Competitive Advantage," *Sustain.*, vol. 14, no. 22, 2022, doi: 10.3390/su142215029.
- [2] C. Zhang *et al.*, "PromotionLens: Inspecting Promotion Strategies of Online E-commerce via Visual Analytics," Aug. 2022, [Online]. Available: <http://arxiv.org/abs/2208.01404>
- [3] A. Ullah, M. I. Mohmand, H. Hussain, S. Johar, and I. Khan, "Customer Analysis Using Machine Learning-Based," *J. Sensors*, 2023.
- [4] M. S. E. Kasem, M. Hamada, and I. Taj-Eddin, "Customer profiling, segmentation, and sales prediction using AI in direct marketing," *Neural Comput. Appl.*, vol. 36, no. 9, pp. 4995–5005, Mar. 2024, doi: 10.1007/s00521-023-09339-6.
- [5] G. Pan, M. Liu, L. M. Tseng, and Z. Geng, "Cultural intelligence and sales performance in online insurance marketing: evidence from a Chinese insurance firm," *Humanit. Soc. Sci. Commun.*, vol. 10, no. 1, Dec. 2023, doi: 10.1057/s41599-023-01623-z.
- [6] P. Singh, L. Arora, and A. Choudhry, "Consumer Behavior in the Service Industry: An Integrative Literature Review and Research Agenda," *Sustain.*, vol. 15, no. 1, Jan. 2023, doi: 10.3390/su15010250.
- [7] O. Dogan, "A Recommendation System in E-Commerce with Profit-Support Fuzzy Association Rule Mining (P-FARM)," *J. Theor. Appl. Electron. Commer. Res.*, vol. 18, no. 2, pp. 831–847, Jun. 2023, doi: 10.3390/jtaer18020043.
- [8] R. Vadrucio, A. Seghezzi, and A. Tumino, "Smart technologies for retailing: who, what, where and why," *Int. J. Retail Distrib. Manag.*,

- vol. 52, no. 13, pp. 79–102, 2024, doi: 10.1108/IJRD-06-2023-0408.
- [9] K. Tabianan, S. Velu, and V. Ravi, “K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data,” *Sustain.*, vol. 14, no. 12, Jun. 2022, doi: 10.3390/su14127243.
- [10] L. (Xuehui) Gao, E. de Haan, I. Melero-Polo, and F. J. Sese, “Winning your customers’ minds and hearts: Disentangling the effects of lock-in and affective customer experience on retention,” *J. Acad. Mark. Sci.*, vol. 51, no. 2, pp. 334–371, 2023, doi: 10.1007/s11747-022-00898-z.
- [11] G. Mena, K. Coussement, K. W. De Bock, A. De Caigny, and S. Lessmann, “Exploiting time-varying RFM measures for customer churn prediction with deep neural networks,” *Ann. Oper. Res.*, vol. 339, no. 1–2, pp. 765–787, 2024, doi: 10.1007/s10479-023-05259-9.
- [12] H. Mo and T. Huang, “Cross-border E-commerce Business Data Processing in the Background of Digital Economy,” *Appl. Math. Nonlinear Sci.*, vol. 9, no. 1, Jan. 2024, doi: 10.2478/amns-2024-0535.
- [13] Y. Cao, P. Leung, and A. Monod, “k-means clustering for persistent homology,” *Adv. Data Anal. Classif.*, Mar. 2024, doi: 10.1007/s11634-023-00578-y.
- [14] A. Punhani, N. Faujdar, K. K. Mishra, and M. Subramanian, “Binning-Based Silhouette Approach to Find the Optimal Cluster Using K-Means,” *IEEE Access*, vol. 10, pp. 115025–115032, 2022, doi: 10.1109/ACCESS.2022.3215568.
- [15] S. F. Djun, I. G. A. Gunadi, and S. Sariyasa, “Analisis Segmentasi Pelanggan pada Bisnis dengan Menggunakan Metode K-Means Clustering pada Model Data RFM,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 5, no. 4, pp. 354–364, 2024, doi: 10.35746/jtim.v5i4.434.
- [16] M. F. Fadhillah, A. Lovely, A. Suyoso, and I. Puspitasari, “Customer Segmentation with Clustering Algorithm Based on Recency , Frequency , and Monetary (RFM) Attributes Segmentasi Pelanggan dengan Algoritma Clustering Berdasarkan Atribut Recency , Frequency dan Monetary (RFM),” vol. 5, no. January, pp. 48–56, 2025.
- [17] A. Alamsyah *et al.*, “Customer Segmentation Using the Integration of the Recency Frequency Monetary Model and the K-Means Cluster Algorithm,” *Sci. J. Informatics*, vol. 9, no. 2, pp. 189–196, 2022, doi: 10.15294/sji.v9i2.39437.