Syntax Literate: Jurnal Ilmiah Indonesia p–ISSN: 2541-0849
e-ISSN: 2548-1398
Vol. 10, No. 4, April 2025

PENERAPAN MACHINE LEARNING DALAM PENGELOMPOKAN PELANGGAN MENGGUNAKAN K-MEANS CLUSTERING UNTUK MENINGKATKAN STRATEGI PEMASARAN PT MASPION

Bimas Ihsan Pratama[1], Verdi Yasin[2*], Irfan Junaedi[3], Anton Zulkarnain Sianipar[4], Zulhalim[5]
Sekolah Tinggi Manajemen Informatika dan Komputer Jayakarta, Indonesia[1,2,3,4,5]
Email: verdiyasin29@gmail.com[2*]

Abstrak

Penelitian ini bertujuan untuk meningkatkan efektivitas strategi pemasaran PT Maspion melalui pengelompokan pelanggan menggunakan algoritma K-Means Clustering. Dengan memanfaatkan data transaksi pelanggan, penelitian ini berhasil mengelompokkan pelanggan ke dalam empat cluster berdasarkan pola pembelian mereka. Setiap cluster dianalisis untuk mengidentifikasi karakteristik utama dan memberikan rekomendasi strategi pemasaran yang relevan, seperti diskon berbasis volume, personalisasi layanan, dan program loyalitas. Hasil penelitian menunjukkan bahwa penerapan K-Means Clustering dapat membantu PT Maspion dalam memahami kebutuhan pelanggan dengan lebih baik, meningkatkan loyalitas, dan mengoptimalkan pendapatan perusahaan. Penelitian ini memberikan kontribusi praktis bagi perusahaan serta memperkaya literatur akademik tentang penerapan machine learning dalam pemasaran.

Kata Kunci: K-Means Clustering, segmentasi pelanggan, machine learning, strategi pemasaran, PT Maspion

Abstract

This study aims to enhance PT Maspion's marketing strategy effectiveness by clustering customers using the K-Means Clustering algorithm. By leveraging customer transaction data, this research successfully grouped customers into four clusters based on their purchasing patterns. Each cluster was analyzed to identify key characteristics and provide relevant marketing strategy recommendations, such as volume-based discounts, personalized services, and loyalty programs. The results indicate that implementing K-Means Clustering helps PT Maspion better understand customer needs, increase loyalty, and optimize company revenue. This study offers practical contributions to the company and enriches academic literature on the application of machine learning in marketing.

Keywords*: K-Means Clustering, customer segmentation, machine learning, marketing strategy, PT Maspion*

Pendahuluan

Dalam era digital saat ini, penerapan teknologi machine learning telah menjadi alat yang sangat penting bagi perusahaan dalam memahami perilaku pelanggan dan merancang strategi pemasaran yang lebih efektif (Junaedy & Asfiah, 2024). Salah satu metode machine learning yang sering digunakan dalam pengelompokan data adalah K-Means Clustering. K-Means Clustering sendiri merupakan salah satu teknik unsupervised learning dalam machine learning, yang berarti bahwa data yang digunakan tidak memiliki label yang jelas (Suyal & Sharma, 2024). Metode ini memungkinkan perusahaan untuk mengelompokkan pelanggan berdasarkan karakteristik tertentu yang diambil dari data

transaksi mereka, seperti frekuensi pembelian, jumlah pembelian, serta kategori produk yang dibeli (Hadi et al., 2017). Dengan pengelompokan ini, perusahaan dapat mengidentifikasi segmen-segmen pelanggan yang memiliki perilaku serupa dan merancang kampanye pemasaran yang lebih spesifik dan tepat sasaran. Hasil dari pengelompokan ini dapat memberikan insight yang berharga mengenai pola perilaku pelanggan, sehingga memungkinkan perusahaan untuk lebih fokus dalam pengembangan produk, layanan, dan komunikasi pemasaran (Awalina & Rahayu, 2023a). Keuntungan utama dari penerapan K-Means Clustering adalah kemampuannya untuk bekerja dengan data yang besar dan kompleks, seperti yang dimiliki. Dalam skenario bisnis modern, data menjadi salah satu aset paling berharga (Alvianatinova et al., 2024). Namun, data mentah tanpa analisis yang tepat hanya akan menjadi tumpukan informasi yang tidak berarti. Oleh karena itu, perusahaan perlu menggunakan alat seperti K-Means Clustering untuk memproses data dan menghasilkan informasi yang dapat digunakan sebagai dasar pengambilan keputusan strategis (Kamila, 2021; Norshahlan et al., 2023; Sarkar et al., 2024).

PT Maspion, sebagai salah satu perusahaan manufaktur terbesar di Indonesia,. Permasalahan yang dihadapi PT Maspion saat ini adalah tidak adanya metode yang efisien dalam mengelompokkan pelanggan berdasarkan data transaksi yang dimiliki. Pengelompokan manual tidak hanya memakan waktu tetapi juga tidak akurat. Selain itu, permasalahan yang kerap muncul adalah ketika algoritma pengelompokan yang diterapkan tidak mampu mengenali preferensi dan kebutuhan pelanggan secara akurat. Ketidaktepatan dalam pengelompokan ini dapat menyebabkan strategi pemasaran yang tidak produktif, karena kampanye pemasaran menjadi kurang relevan bagi kelompok pelanggan tertentu. Hal ini berpotensi menurunkan performa sistem pemasaran, di mana hasil yang diharapkan, seperti peningkatan efektivitas dan efisiensi dalam personalisasi, tidak dapat tercapai. Akibatnya, sistem tidak hanya gagal memenuhi kebutuhan pelanggan, tetapi juga tidak dapat mengakomodasi kebutuhan PT Maspion dalam merespons perubahan pasar secara dinamis. Strategi pemasaran yang tidak tepat dapat menyebabkan ketidakpuasan pelanggan, penurunan loyalitas, dan bahkan kehilangan pelanggan potensial. Dari sudut pandang PT Maspion, ketidakmampuan dalam mengakomodasi kebutuhan ini berpotensi mengakibatkan kerugian signifikan, baik dari segi pendapatan maupun reputasi.

Hal ini menyebabkan strategi pemasaran yang dijalankan seringkali kurang tepat sasaran dan tidak sesuai dengan kebutuhan pelanggan yang beragam K-Means Clustering menawarkan solusi yang efektif dan efisien untuk mengatasi permasalahan tersebut (Liu & Billard, 2022; Tabianan et al., 2022; Wu et al., 2021). Dengan memanfaatkan data pelanggan yang ada, PT Maspion dapat mengembangkan kampanye pemasaran yang lebih personal dan relevan, yang pada akhirnya akan meningkatkan retensi pelanggan dan pendapatan perusahaan. Dengan menggunakan pendekatan ini, PT Maspion dapat lebih memahami preferensi dan kebutuhan pelanggan mereka, sehingga strategi pemasaran dapat disesuaikan dengan segmen pasar yang spesifik.

Penelitian ini mengembangkan penelitian sebelumnya yang berkaitan dengan segmentasi pelanggan dan peningkatan strategi pemasaran menggunakan teknik machine learning, khususnya algoritma K-Means Clustering. Sebagai contoh, Junaedy dan Asfiah (2024) menekankan peran machine learning dalam memahami perilaku pelanggan untuk merancang strategi pemasaran yang efektif, sementara Hadi et al. (2017) menunjukkan bagaimana pengelompokan data pelanggan berdasarkan karakteristik transaksi dapat menginformasikan upaya pemasaran yang lebih terarah. Kebaruan dari penelitian ini

terletak pada penerapannya yang spesifik pada PT Maspion, sebuah perusahaan manufaktur besar di Indonesia, yang meskipun menghadapi tantangan dengan segmentasi manual tradisional, kini dapat memanfaatkan K-Means Clustering untuk mengelompokkan pelanggan secara lebih akurat dan efisien.

Dalam penelitian ini, fokus utama adalah pada penerapan K-Means Clustering untuk mengelompokkan pelanggan PT Maspion berdasarkan data transaksi yang ada. Penelitian ini bertujuan untuk mengevaluasi sejauh mana penerapan algoritma ini dapat meningkatkan efektivitas strategi pemasaran perusahaan. Selain itu, penelitian ini juga akan mengeksplorasi manfaat yang dapat diperoleh PT Maspion dari pengelompokan pelanggan, seperti peningkatan personalisasi dalam kampanye pemasaran, peningkatan kepuasan pelanggan, dan peningkatan pendapatan. Diharapkan hasil dari penelitian ini dapat memberikan kontribusi tidak hanya bagi PT Maspion, tetapi juga bagi literatur akademik terkait penerapan machine learning dalam pemasaran.

Metode Penelitian

Penelitian ini dilakukan di PT Maspion, yang merupakan salah satu perusahaan manufaktur dan distribusi terkemuka di Indonesia yang belamatkan Jl. Kembang Jepun 38-40 Surabaya. Proses pengumpulan data dilakukan melalui akses ke database transaksi pelanggan PT Maspion. Waktu penelitian berlangsung selama 6 bulan, dimulai pada Oktober 2024 hingga Maret 2025. Selama periode ini, analisis data dan penerapan algoritma machine learning dilakukan secara bertahap sesuai dengan metode Agile. Penelitian ini menggunakan paradigma positivistik yang berfokus pada pendekatan kuantitatif, di mana peneliti menguji hubungan atau pola dalam data pelanggan berdasarkan transaksi yang tersedia. Paradigma ini menekankan pengukuran yang objektif dan analisis data yang sistematis melalui penerapan algoritma machine learning, yaitu K-Means Clustering, untuk memahami pengelompokan pelanggan dan pengaruhnya terhadap strategi pemasaran.

Desain penelitian ini adalah penelitian deskriptif-kuantitatif, yang bertujuan untuk mengidentifikasi dan menganalisis pola pembelian pelanggan PT Maspion dengan menggunakan data transaksi. Penerapan algoritma K-Means Clustering bertujuan untuk mengelompokkan pelanggan berdasarkan kesamaan dalam frekuensi pembelian, jumlah pembelian, dan kategori produk yang dibeli. Penelitian ini juga menggunakan metode Agile, yang memungkinkan pengembangan model dilakukan secara iteratif dan fleksibel, disesuaikan dengan hasil yang diperoleh di setiap tahapan. Teknik pengumpulan data dalam penelitian ini dilakukan melalui akses ke database transaksi pelanggan PT Maspion. Data transaksi yang mencakup variabel-variabel seperti frekuensi pembelian, jumlah pembelian, kategori produk, dan lokasi pembelian diekstraksi dan diolah untuk keperluan analisis. Data diunduh dalam format CSV untuk mempermudah proses analisis menggunakan perangkat lunak Python.

Proses analisis data dilakukan dalam beberapa tahapan berikut:

1. Preprocessing Data: Data yang telah dikumpulkan akan melalui tahap preprocessing, termasuk pembersihan data (data cleaning), penanganan data hilang, normalisasi, dan pengubahan format data agar siap untuk digunakan dalam model machine learning. Ini dilakukan untuk memastikan kualitas data yang akan digunakan.
2. Penerapan Algoritma K-Means Clustering: Data yang telah diproses dianalisis menggunakan algoritma K-Means Clustering dengan bantuan library Scikit-Learn dalam Python. Algoritma ini akan mengelompokkan pelanggan berdasarkan pola transaksi.

3. Validasi Model: Hasil pengelompokan akan dievaluasi menggunakan Silhouette Score untuk menilai seberapa baik pelanggan dikelompokkan. Jika diperlukan, iterasi ulang dilakukan untuk meningkatkan hasil pengelompokan.
4. Visualisasi Data: Hasil pengelompokan kemudian divisualisasikan dalam bentuk grafik atau diagram untuk mempermudah pemahaman dan analisis lebih lanjut.
5. Implementasi Hasil: Hasil akhir pengelompokan akan diimplementasikan dalam strategi pemasaran PT Maspion untuk meningkatkan personalisasi kampanye dan relevansi promosi terhadap segmen pelanggan yang berbeda.

Hasil dan Pembahasan

Preprocessing Data

Tahapan preprocessing data dilakukan untuk mempersiapkan data transaksi pelanggan PT Maspion agar dapat dianalisis secara optimal. Langkah pertama adalah pemilihan atribut, di mana hanya kolom yang relevan, seperti *Kd. Cust* (kode pelanggan), *Customer* (nama pelanggan), *Nama Barang* (deskripsi produk), *Qty* (jumlah barang yang dibeli), *Harga* (harga per unit barang), dan *Subtotal* (total nilai transaksi), dipertahankan untuk analisis. Kolom lain yang tidak relevan dihapus untuk menyederhanakan data. Berikut ini adalah gambar yang menunjukkan data dengan kolom-kolom relevan setelah proses pemilihan atribut.

Kd. Cust	Customer	Nama Barang	Qty	Harga	Subtotal
022A2153	ROBIN CHANDRA	FAUCET SOCK TS 4"	24	25.150,00	603.600,00
00240832	ANUGERAH DAMAI BERSAMA	TEE DV 10"	10	395.840,00	3.958.400,00
00203621	KALINDO INTILESTARI	GUTTER PROFILE HR 6" X 4 M	500	92.970,00	46.485.000,00
00240664	BANGUN KARYA REJEKI,PT	SOCK DV 4 X 3	150	7.870,00	1.180.500,00
00240664	BANGUN KARYA REJEKI,PT	GUTTER UNION DV 220(GU.04R-220	92	15.380,00	1.414.960,00
00203911	MAKTAL DISTRIBUSI PRATAMA,PT	GUTTER OUTLET DV HR 6(GO.03R6)	55	14.440,00	794.200,00
00203911	MAKTAL DISTRIBUSI PRATAMA,PT	GUTTER UNION DV HR 6(GU.04R-6)	300	2.480,00	744.000,00
022A2153	ROBIN CHANDRA	KNEE DV 10"	2	291.560,00	583.120,00
022A2153	ROBIN CHANDRA	TEE-Y DV 6" X 90	12	90.670,00	1.088.040,00
00241035	CAHAYA GUNA SUKSES	GUTTER CAP DV 140 MM	400	6.130,00	2.452.000,00
022A2153	ROBIN CHANDRA	TEE TS 3"	16	30.310,00	484.960,00
022A2153	ROBIN CHANDRA	TEE TS 3"	16	30.310,00	484.960,00
00241035	CAHAYA GUNA SUKSES	PIPA MOF ABU 14" AW X 4 M	4	3.862.040,00	15.448.160,00
00221077	ATIKA MANDIRI	S12,5 PIPA TJ 90 MM X 6 M EFF	100	374.820,00	37.482.000,00
00221077	ATIKA MANDIRI	RUBBER RING 90 MM (3" AW)	100	1,00	100,00

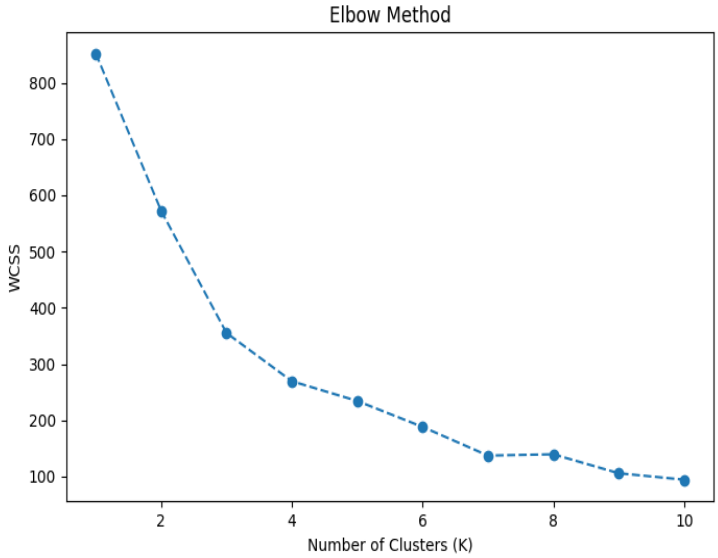
Gambar 1. Pemilihan Atribut

Selanjutnya, data diperiksa untuk mendeteksi keberadaan nilai kosong atau tidak valid. Baris-baris data yang mengandung nilai kosong pada kolom penting seperti *Kd. Cust* (kode pelanggan), *Customer* (nama pelanggan), *Nama Barang* (deskripsi produk), *Qty* (jumlah barang yang dibeli), *Harga* (harga per unit barang), dan *Subtotal* (total nilai transaksi), dihapus untuk menjaga kualitas dan keakuratan hasil analisis. Setelah data dibersihkan, dilakukan agregasi data berdasarkan pelanggan, dengan

menghitung total kuantitas barang yang dibeli, rata-rata harga barang per unit, dan total penjualan untuk setiap pelanggan (Butsianto & Mayangwulan, 2020; Ikotun et al., 2023; Sirunyan et al., 2020). Hasil agregasi ini memberikan ringkasan pola transaksi pelanggan yang dapat digunakan sebagai dasar untuk analisis clustering. Proses preprocessing ini memastikan bahwa data yang digunakan bersih, relevan, dan siap untuk mendukung langkah analisis lebih lanjut.

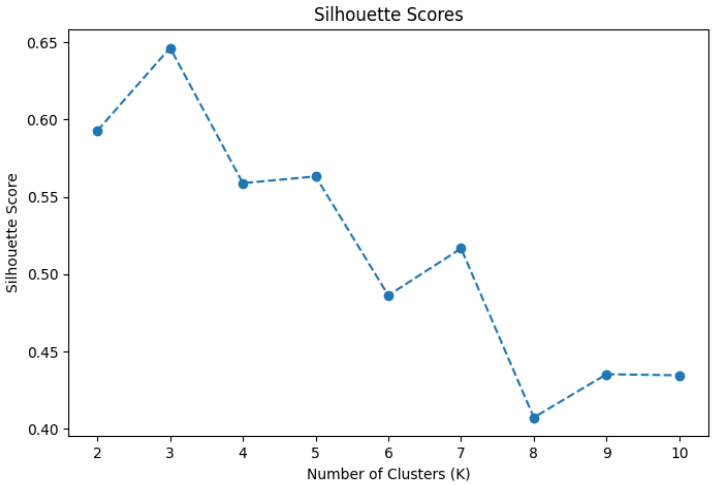
Penentuan Jumlah Cluster (Optimal K)

Jumlah cluster (*K*) yang optimal perlu ditentukan untuk menghasilkan pengelompokan yang efektif menggunakan algoritma *K-Means Clustering* (Ahmed et al., 2020; Atira & Nurina Sari, 2023; Awalina & Rahayu, 2023b). Dalam penelitian ini, digunakan dua metode, yaitu Elbow Method dan Silhouette Score, untuk menentukan jumlah cluster optimal. Elbow Method mengevaluasi WCSS (*Within-Cluster Sum of Squares*) untuk berbagai jumlah cluster (*K*). Grafik yang dihasilkan menunjukkan hubungan antara jumlah cluster dan nilai WCSS. Titik optimal ditentukan pada titik siku (*elbow point*), di mana penurunan nilai WCSS mulai melambat. Grafik hasil analisis menggunakan Elbow Method dapat dilihat pada Gambar 2.



Gambar 2. Grafik Elbow Method

Pada Gambar 2, terlihat bahwa penurunan nilai WCSS signifikan pada awalnya, namun mulai melambat setelah jumlah cluster tertentu. Titik siku (*elbow point*) yang menjadi indikator jumlah cluster optimal terlihat pada *K = 4*. Pada titik ini, penambahan jumlah cluster tidak lagi memberikan pengurangan WCSS yang signifikan, sehingga *K = 4* dipilih sebagai jumlah cluster optimal untuk analisis ini. Silhouette Score digunakan untuk memvalidasi hasil dari Elbow Method. Skor ini mengukur sejauh mana data dalam suatu cluster lebih dekat satu sama lain dibandingkan dengan data dalam cluster lain. Grafik hasil analisis menggunakan Silhouette Score dapat dilihat pada Gambar 3.



Gambar 3. *Grafik Silhouette Score*

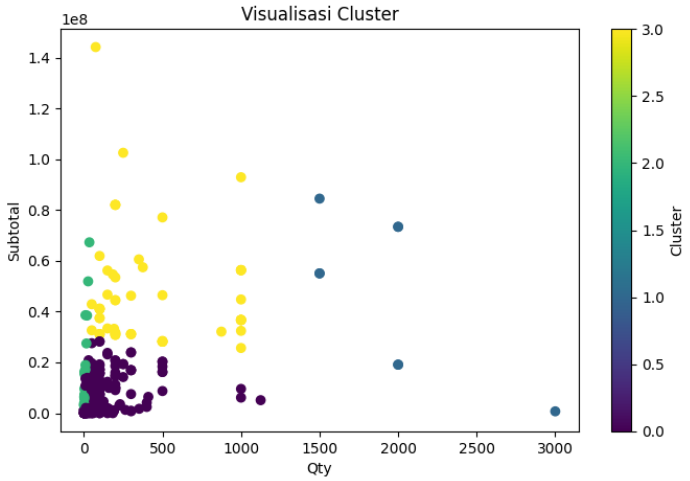
Berdasarkan Gambar 3, terlihat bahwa nilai Silhouette Score tertinggi dicapai pada $K = 4$. Hal ini mendukung hasil dari Elbow Method, di mana pengelompokan pelanggan menjadi 4 segmen memberikan hasil yang paling efektif dalam memisahkan pola transaksi mereka.

Penerapan Algoritma K-Means Clustering

Setelah menentukan jumlah cluster optimal ($K = 4$), algoritma K-Means Clustering diterapkan pada data transaksi pelanggan. Proses ini melibatkan:

a) Inisialisasi centroid sebanyak jumlah cluster ($K = 4$).
b) Pengelompokan data berdasarkan jarak terdekat dari centroid.
c) Iterasi hingga posisi centroid tidak berubah atau mencapai jumlah iterasi maksimum.

Hasil dari algoritma K-Means Clustering menunjukkan bahwa data pelanggan berhasil dikelompokkan ke dalam empat segmen yang berbeda. Informasi cluster ditambahkan ke dataset sebagai kolom baru bernama Cluster. Visualisasi hasil clustering dapat dilihat pada Gambar 4.



Gambar 4. Visualisasi Cluster Pelanggan

Gambar 4 menunjukkan hasil clustering pelanggan menggunakan algoritma K-Means Clustering dengan jumlah cluster optimal sebanyak 4. Sumbu X menunjukkan

jumlah barang yang dibeli (*Qty*), sedangkan sumbu Y menunjukkan total nilai transaksi (*Subtotal*). Warna yang berbeda pada grafik merepresentasikan cluster yang berbeda. Dari visualisasi ini, dapat disimpulkan bahwa:

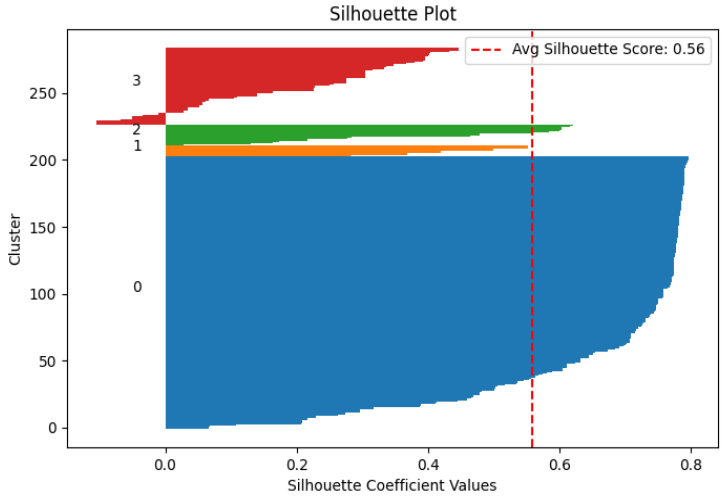
a) Pelanggan dengan jumlah pembelian (*Qty*) rendah namun memiliki total nilai transaksi (*Subtotal*) tinggi (misalnya, cluster hijau) kemungkinan besar membeli produk dengan harga per unit yang tinggi.
b) Pelanggan dengan jumlah pembelian tinggi namun total nilai transaksi sedang (misalnya, cluster kuning) menunjukkan pembelian produk dengan harga rata-rata per unit yang lebih rendah.
c) Cluster lainnya (ungu dan biru) mewakili segmen pelanggan dengan pola pembelian yang berbeda, seperti pembelian dalam jumlah kecil dengan total nilai transaksi rendah.

Pengelompokan ini dapat digunakan untuk merancang strategi pemasaran yang lebih efektif, seperti memberikan diskon kepada pelanggan dengan transaksi tinggi untuk meningkatkan loyalitas atau menawarkan promosi kepada pelanggan di cluster dengan pembelian rendah.

Evaluasi Hasil Clustering

Evaluasi hasil clustering dilakukan untuk menilai kualitas pengelompokan yang dihasilkan oleh algoritma K-Means. Dalam penelitian ini, digunakan Silhouette Score sebagai metrik evaluasi. Silhouette Score mengukur sejauh mana data dalam suatu cluster lebih dekat dengan anggota cluster yang sama dibandingkan dengan cluster lain. Nilai Silhouette Score berkisar antara -1 hingga 1:

a) Nilai mendekati 1 menunjukkan kualitas clustering yang baik, di mana data dalam satu cluster lebih seragam dan jauh dari cluster lain.
b) Nilai mendekati 0 menunjukkan bahwa data berada di batas antara dua cluster.
c) Nilai negatif menunjukkan kesalahan dalam pengelompokan.



Gambar 5. Silhouette Plot untuk Cluster Optimal

Gambar 5 memperlihatkan Silhouette Plot, yang menggambarkan distribusi nilai koefisien Silhouette untuk setiap cluster. Setiap bar horizontal pada plot merepresentasikan nilai Silhouette Score untuk satu data poin. Cluster dengan warna berbeda menunjukkan distribusi data dalam cluster yang berbeda. Hasil evaluasi menunjukkan bahwa Silhouette Score untuk jumlah cluster optimal (K = 4) adalah 0.56.

Nilai ini mengindikasikan bahwa pengelompokan data memiliki kualitas yang cukup baik, dengan sebagian besar data berada dalam cluster yang sesuai. Garis merah vertikal pada plot menunjukkan rata-rata nilai Silhouette Score, yang memberikan gambaran keseluruhan kualitas pengelompokan data.

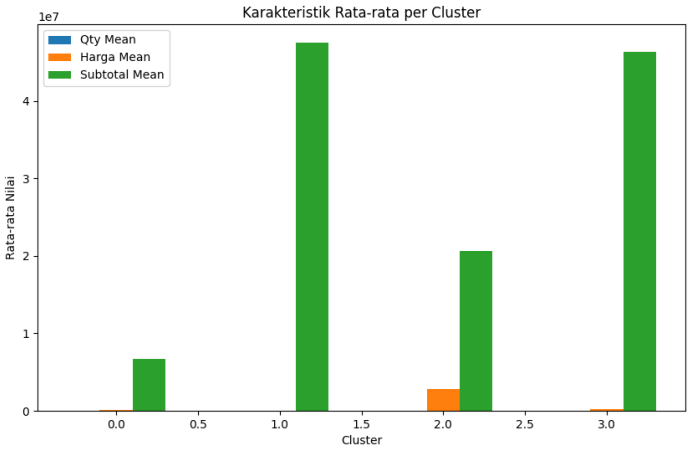
Analisis Karakteristik Cluster

Setelah melakukan pengelompokan data pelanggan menggunakan algoritma K-Means, langkah selanjutnya adalah menganalisis karakteristik dari masing-masing cluster. Analisis ini dilakukan untuk memahami pola transaksi yang ada dalam setiap cluster, seperti jumlah pembelian rata-rata, harga rata-rata, dan total nilai transaksi. Tabel 1 menunjukkan ringkasan karakteristik utama dari setiap cluster, termasuk rata-rata dan total dari Qty, Harga, dan Subtotal.

Tabel 1. Ringkasan Karakteristik Cluster

Cluster	Rata-rata Qty	Rata-rata Harga	Rata-rata Subtotal	Total Qty	Total Harga	Total Subtotal
0	117.38.00	136,088.83	6,716,809.42	23,829	2,762,603,200	1,363,547,026
1	1,937.5	15,500.0	475,675.0	15,5	222,600,000	3,805,400,000
2	9.31	282,628.12	20,692,335.63	149	45,258,500	3,310,773,700
3	416.68	243,175.43	46,246,484.12	23,751	1,381,610,000	26,430,996,300

Tabel 1 menyajikan ringkasan karakteristik utama dari setiap cluster, termasuk rata-rata dan total dari metrik Qty, Harga, dan Subtotal. Hasil analisis menunjukkan bahwa setiap cluster memiliki pola transaksi yang unik. Cluster 0 mencerminkan pelanggan dengan jumlah pembelian sedang (rata-rata 117 unit), harga per unit tinggi (rata-rata 136,088.83), dan total nilai transaksi sedang. Cluster 1 mencakup pelanggan dengan jumlah pembelian besar (rata-rata 1,937 unit) tetapi harga per unit rendah (rata-rata 15,500.0), sehingga menghasilkan total nilai transaksi yang cukup besar. Sebaliknya, Cluster 2 menunjukkan pelanggan dengan jumlah pembelian kecil (rata-rata 9 unit), namun dengan harga per unit yang sangat tinggi (rata-rata 282,628.12), menghasilkan nilai transaksi yang signifikan. Sementara itu, Cluster 3 diisi oleh pelanggan dengan jumlah pembelian cukup tinggi (rata-rata 416 unit), harga per unit tinggi (rata-rata 243,175.43), dan total nilai transaksi terbesar. Gambar 6 memperlihatkan grafik bar chart untuk rata-rata Qty, Harga, dan Subtotal di setiap cluster.



Gambar 6. Karakteristik Rata-rata per Cluster

Gambar 6. memperlihatkan visualisasi berupa grafik bar chart untuk rata-rata Qty, Harga, dan Subtotal di setiap cluster. Analisis ini memberikan wawasan penting untuk strategi pemasaran. Sebagai contoh, pelanggan di Cluster 2 dapat diberikan promosi khusus karena mereka cenderung membeli barang dengan harga tinggi tetapi dalam jumlah kecil. Sementara itu, program loyalitas dapat dikembangkan untuk pelanggan di Cluster 3, karena mereka memiliki kontribusi nilai transaksi tertinggi. Hasil ini menunjukkan bahwa dengan memahami karakteristik setiap cluster, perusahaan dapat merancang strategi pemasaran yang lebih efektif dan sesuai dengan kebutuhan pelanggan di masing-masing kelompok.

Rekomendasi Strategi Pemasaran

Berdasarkan analisis karakteristik cluster yang telah dilakukan, direkomendasikan strategi pemasaran yang disesuaikan dengan pola transaksi masing-masing cluster. Strategi ini bertujuan untuk meningkatkan nilai transaksi, loyalitas pelanggan, dan efektivitas promosi. Pada Cluster 0, yang terdiri dari pelanggan dengan jumlah pembelian sedang dan harga per unit tinggi, disarankan untuk memberikan insentif berupa diskon berbasis jumlah pembelian guna mendorong pembelian lebih rutin. Selain itu, program eksklusif seperti "Membership Premium" dapat ditawarkan untuk menjaga loyalitas pelanggan dengan kontribusi nilai transaksi yang konsisten.

Sementara itu, Cluster 1, yang mencakup pelanggan dengan jumlah pembelian besar tetapi harga per unit rendah, dapat dioptimalkan dengan menawarkan diskon volume untuk mendorong transaksi dalam jumlah besar. Selain itu, perusahaan dapat memperkenalkan produk dengan nilai tambah lebih tinggi untuk meningkatkan margin keuntungan dari kelompok pelanggan ini. Untuk Cluster 2, yang memiliki pelanggan dengan jumlah pembelian kecil tetapi harga per unit sangat tinggi, disarankan untuk meluncurkan program eksklusif seperti undangan ke acara khusus atau program VIP. Layanan personalisasi, seperti konsultasi langsung atau after-sales service, juga dapat diterapkan untuk meningkatkan loyalitas pelanggan premium ini.

Pada Cluster 3, yang terdiri dari pelanggan dengan jumlah pembelian tinggi dan nilai transaksi terbesar, fokus strategi adalah menjaga loyalitas dengan program reward seperti cashback atau poin yang dapat digunakan untuk pembelian berikutnya. Selain itu, promosi produk yang relevan dengan kategori barang yang sering dibeli oleh pelanggan ini dapat dilakukan untuk meningkatkan frekuensi transaksi.

Dengan mengimplementasikan strategi pemasaran yang terfokus pada karakteristik masing-masing cluster, perusahaan dapat meningkatkan efisiensi pemasaran, memperkuat hubungan dengan pelanggan, serta mengoptimalkan pendapatan secara keseluruhan. Strategi berbasis data ini diharapkan dapat memberikan hasil yang maksimal sesuai dengan kontribusi setiap cluster terhadap bisnis.

Pembahasan

Interpretasi Hasil Clustering

Hasil clustering yang diperoleh dari analisis menggunakan algoritma K-Means menunjukkan bahwa data pelanggan dapat dikelompokkan menjadi empat cluster utama berdasarkan pola transaksi mereka. Setiap cluster memiliki karakteristik unik yang mencerminkan perilaku pembelian pelanggan. Pada Cluster 0, pelanggan memiliki jumlah pembelian sedang, namun harga per unit barang yang dibeli relatif tinggi. Hal ini

menunjukkan bahwa pelanggan dalam cluster ini cenderung membeli produk dengan nilai tinggi secara sporadis.

Cluster 1 mencerminkan pelanggan dengan jumlah pembelian yang sangat besar namun harga per unit yang lebih rendah. Pola ini mengindikasikan bahwa pelanggan dalam cluster ini kemungkinan besar adalah pelanggan grosir atau pembeli dalam jumlah besar untuk barang dengan harga terjangkau. Sebaliknya, Cluster 2 terdiri dari pelanggan dengan jumlah pembelian kecil, tetapi harga per unit barang sangat tinggi. Cluster ini menunjukkan segmen pelanggan premium yang cenderung membeli barang bernilai tinggi secara terbatas, namun dengan kontribusi signifikan terhadap total nilai transaksi. Sementara itu, Cluster 3 terdiri dari pelanggan dengan jumlah pembelian cukup tinggi dan harga per unit barang yang juga relatif tinggi. Pelanggan dalam cluster ini berkontribusi sebagai penyumbang utama terhadap total nilai transaksi, yang menunjukkan bahwa mereka adalah pelanggan utama perusahaan.

Hasil pengelompokan ini memberikan wawasan penting bagi perusahaan untuk memahami segmen pelanggan mereka dengan lebih baik. Setiap cluster dapat dimanfaatkan sebagai dasar dalam merancang strategi pemasaran yang lebih efektif, sesuai dengan pola pembelian dan preferensi pelanggan dalam cluster tersebut. Interpretasi ini tidak hanya membantu perusahaan dalam mengoptimalkan pendapatan, tetapi juga memperkuat hubungan dengan pelanggan melalui pendekatan pemasaran yang terfokus.

Implikasi Terhadap Strategi Pemasaran PT Maspion

Hasil analisis clustering memberikan wawasan yang berharga untuk mendukung strategi pemasaran yang lebih efektif dan efisien di PT Maspion. Dengan memahami karakteristik pelanggan di setiap cluster, perusahaan dapat merancang pendekatan yang lebih terfokus untuk memenuhi kebutuhan dan preferensi masing-masing segmen pelanggan. Implikasi utama dari hasil clustering terhadap strategi pemasaran PT Maspion adalah sebagai berikut:

Pada Cluster 0, yang terdiri dari pelanggan dengan jumlah pembelian sedang namun harga per unit tinggi, PT Maspion dapat memanfaatkan segmen ini untuk meningkatkan loyalitas pelanggan dengan menawarkan program diskon berbasis volume pembelian. Strategi ini bertujuan untuk mendorong pembelian lebih rutin, sehingga pelanggan ini dapat memberikan kontribusi lebih besar terhadap total pendapatan. Selain itu, perusahaan dapat mengembangkan program keanggotaan eksklusif untuk menarik minat pelanggan dalam cluster ini.

Cluster 1, yang mencakup pelanggan dengan jumlah pembelian besar namun harga per unit rendah, memberikan peluang bagi PT Maspion untuk menawarkan promosi khusus dalam bentuk diskon volume atau paket bundling. Strategi ini tidak hanya akan memperkuat hubungan dengan pelanggan grosir tetapi juga meningkatkan efisiensi operasional melalui transaksi dalam jumlah besar. Selain itu, perusahaan dapat mengidentifikasi peluang untuk menawarkan produk dengan nilai tambah yang lebih tinggi guna meningkatkan margin keuntungan pada segmen ini.

Sementara itu, Cluster 2, yang terdiri dari pelanggan dengan jumlah pembelian kecil tetapi harga per unit yang sangat tinggi, menuntut strategi yang lebih personal. PT Maspion dapat meluncurkan program eksklusif seperti undangan ke acara VIP atau memberikan layanan personalisasi seperti konsultasi produk dan layanan purna jual. Tujuan utama strategi ini adalah untuk meningkatkan loyalitas pelanggan premium yang berkontribusi besar terhadap nilai transaksi meskipun jumlah pembelian mereka terbatas.

Pada Cluster 3, yang terdiri dari pelanggan dengan jumlah pembelian tinggi dan nilai transaksi terbesar, fokus utama perusahaan adalah menjaga loyalitas mereka. PT Maspion dapat meluncurkan program reward seperti poin loyalitas, cashback, atau diskon khusus untuk pembelian berulang. Selain itu, perusahaan dapat memanfaatkan data transaksi dari cluster ini untuk merekomendasikan produk yang relevan dengan kebutuhan pelanggan, sehingga meningkatkan peluang cross-selling dan up-selling.

Implikasi strategis dari analisis clustering ini adalah kemampuan PT Maspion untuk mengalokasikan sumber daya pemasaran secara lebih efisien dan efektif. Dengan memahami kebutuhan dan preferensi setiap cluster, perusahaan dapat mengurangi biaya pemasaran yang tidak terarah dan sekaligus meningkatkan kepuasan pelanggan. Pendekatan berbasis data ini juga memungkinkan PT Maspion untuk membangun hubungan jangka panjang dengan pelanggan melalui strategi pemasaran yang relevan dan terpersonalisasi.

Kesimpulan

Penelitian ini bertujuan untuk mengelompokkan pelanggan PT Maspion menggunakan algoritma K-Means guna memberikan rekomendasi strategi pemasaran yang lebih efektif dan berbasis data. Berdasarkan hasil analisis, data pelanggan berhasil dikelompokkan ke dalam empat cluster dengan karakteristik yang berbeda. Cluster 0 terdiri dari pelanggan dengan jumlah pembelian sedang dan harga per unit tinggi, yang menunjukkan adanya segmen pelanggan dengan kontribusi transaksi yang stabil. Cluster 1 mencerminkan pelanggan dengan jumlah pembelian sangat besar tetapi harga per unit rendah, yang kemungkinan besar merupakan pelanggan grosir atau distributor. Cluster 2 menunjukkan pelanggan premium yang membeli barang bernilai tinggi meskipun jumlah pembelian kecil. Cluster 3, sebagai penyumbang terbesar nilai transaksi, terdiri dari pelanggan dengan jumlah pembelian tinggi dan harga per unit yang juga tinggi.

Hasil clustering ini memberikan wawasan yang signifikan bagi PT Maspion untuk mengembangkan strategi pemasaran yang disesuaikan dengan kebutuhan dan pola pembelian pelanggan di setiap cluster. Strategi yang dirancang meliputi program loyalitas, promosi diskon berbasis volume, personalisasi layanan, hingga rekomendasi produk yang relevan. Dengan implementasi strategi yang terarah, perusahaan diharapkan dapat meningkatkan efisiensi pemasaran, loyalitas pelanggan, dan nilai transaksi secara keseluruhan.

DAFTAR PUSTAKA

Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. In *Electronics (Switzerland)* (Vol. 9, Issue 8). https://doi.org/10.3390/electronics9081295

Alvianatinova, V., Ali, I., Rahaningsih, N., & Bahtiar, A. (2024). Penerapan Algoritma K-Means Clustering Dalam Pengelompokan Data Penjualan Supermarket Berdasarkan Cabang (Branch). *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(2), 1529–1535. https://doi.org/10.36040/jati.v8i2.8993

Atira, A., & Nurina Sari, B. (2023). Penerapan Silhouette Coefficient, Elbow Method dan Gap Statistics untuk Penentuan Cluster Optimum dalam Pengelompokkan Provinsi di Indonesia Berdasarkan Indeks Kebahagiaan. *Jurnal Ilmiah Wahana Pendidikan*, *9*(17).

Awalina, E. F. L., & Rahayu, W. I. (2023a). Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail. *Jurnal Teknologi Dan Informasi*, *13*(2), 122–137. https://doi.org/10.34010/jati.v13i2.10090

Awalina, E. F. L., & Rahayu, W. I. (2023b). Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail. *Jurnal Teknologi Dan Informasi*, *13*(2). https://doi.org/10.34010/jati.v13i2.10090

Butsianto, S., & Mayangwulan, N. T. (2020). Penerapan Data Mining Untuk Prediksi Penjualan Mobil Menggunakan Metode K-Means Clustering. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, *3*(3). https://doi.org/10.32672/jnkti.v3i3.2428

Hadi, F., Octari Rahmadia, D., Hadi Nugraha, F., Putri Bulan, N., Mustakin, & Monalisa, S. (2017). Penerapan K-Means Clustering Berdasarkan RFM Mofek Sebagai Pemetaan dan Pendukung Strategi Pengelolaan Pelanggan (Studi Kasus: PT. Herbal Penawar Alwahidah Indonesia Pekanbaru). *SITEKIN: Jurnal Sains, Teknologi Dan Industri*, *15*(1), 69–76.

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*. https://doi.org/10.1016/j.ins.2022.11.139

Junaedy, N. A. A., & Asfiah, N. (2024). Hubungan antara AI, Machine Learning, Dan Implikasinya Terhadap Responsivitas Bisnis. *Jurnal Bisnis Inovatif Dan Digital*, *1*(3), 81–91. https://doi.org/10.61132/jubid.v1i3.189

Kamila, C. (2021). Systematic Literature Review: Penggunaan Algoritma K-Means Untuk Clustering di Indonesia dalam Bidang Pendidikan. *INTECH*, *2*(1). https://doi.org/10.54895/intech.v2i1.866

Liu, F., & Billard, L. (2022). Partition of Interval-Valued Observations Using Regression. *Journal of Classification*, *39*(1). https://doi.org/10.1007/s00357-021-09394-5

Norshahlan, M., Jaya, H., & Kustini, R. (2023). Penerapan Metode Clustering Dengan Algoritma K-means Pada Pengelompokan Data Calon Siswa Baru. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, *2*(6). https://doi.org/10.53513/jursi.v2i6.9148

Sarkar, M., Puja, A. R., & Chowdhury, F. R. (2024). Optimizing Marketing Strategies with RFM Method and K-Means Clustering-Based AI Customer Segmentation Analysis. *Journal of Business and Management Studies*, *6*(2). https://doi.org/10.32996/jbms.2024.6.2.5

Sirunyan, A. M., Tumasyan, A., Adam, W., Ambrogi, F., Bergauer, T., Dragicevic, M., Erö, J., Del Valle, A. E., Flechl, M., Frühwirth, R., Jeitler, M., Krammer, N., Krätschmer, I., Liko, D., Madlener, T., Mikulec, I., Rad, N., Schieck, J., Schöfbeck, R., … Trembath-Reichert, S. (2020). Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques. *Journal of Instrumentation*, *15*(6). https://doi.org/10.1088/1748-0221/15/06/P06005

Suyal, M., & Sharma, S. (2024). A Review on Analysis of K-Means Clustering Machine Learning Algorithm based on Unsupervised Learning. *Journal of Artificial Intelligence and Systems*, *6*(1), 85–95. https://doi.org/10.33969/ais.2024060106

Tabianan, K., Velu, S., & Ravi, V. (2022). K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data. *Sustainability (Switzerland)*, *14*(12). https://doi.org/10.3390/su14127243

Wu, S., Yau, W. C., Ong, T. S., & Chong, S. C. (2021). Integrated Churn Prediction and Customer Segmentation Framework for Telco Business. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3073776

Copyright holder:
Bimas Ihsan Pratama, Verdi Yasin, Irfan Junaedi, Anton Zulkarnain Sianipar, Zulhalim (2025)

First publication right:
Syntax Literate: Jurnal Ilmiah Indonesia

This article is licensed under:

