VOL. 03. NO. 05 SEPTEMBER 2026

JOURNAL OF COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE (JCSAI)

ANALISIS SEGMENTASI PELANGGAN PADA DATA PENJUALAN ELEKTRONIK MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Aditya Jordy[1], Nana Suarna[2], Agus Bahtiar[3], Edi Wahyudin[4].

Program Studi Teknik Informatika[12]
Program Studi Sistem Informasi[3]
Program Studi Komputerisasi Akuntansi[4]

STMIK IKMI Cirebon
https://ikmi.ac.id/page/18/?lang=de
adityajordy06@gmail.com

(*) Corresponding Author : adityajordy06@gmail.com
Published : 17 Juli 2026

Abstract—*The rapid development of information technology has led to a significant increase in transaction data, particularly in the electronic sales sector. This transaction data contains valuable information regarding customer behavior that can be utilized to support business decision-making. However, in practice, such data is often not optimally processed, resulting in limited insights being generated. Therefore, this study aims to perform customer segmentation using the K-Means algorithm based on electronic sales transaction data. The research process begins with data preprocessing, which includes data cleaning, data transformation, and normalization to produce a dataset suitable for clustering. Subsequently, the clustering process is conducted using the K-Means algorithm by testing several values of the number of clusters (K). The optimal number of clusters is determined using the Davies-Bouldin Index (DBI), where a lower DBI value indicates better cluster quality.The results of this study show that the optimal number of clusters is obtained at K = 2 with a DBI value of 0.073, indicating that the resulting clusters have good separation. Based on the clustering results, centroid analysis is performed to identify the characteristics of each customer segment. The findings reveal that customers are divided into two main groups, namely high-purchasing customers and low-purchasing customers. In conclusion, the implementation of the K-Means algorithm integrated with preprocessing stages and evaluation using the Davies-Bouldin Index (DBI) is capable of producing accurate customer segmentation, which can be used as a basis for developing data-driven marketing strategies.*

Keywords: *Customer Segmentation, K-Means, Clustering, Davies-Bouldin Index*

Abstrak—Perkembangan teknologi informasi telah mendorong peningkatan jumlah data transaksi, khususnya pada sektor penjualan elektronik. Data transaksi tersebut mengandung informasi penting terkait perilaku pelanggan yang dapat dimanfaatkan untuk mendukung pengambilan keputusan bisnis. Namun, dalam praktiknya, data tersebut sering kali belum diolah secara optimal sehingga belum mampu memberikan insight yang signifikan. Oleh karena itu, penelitian ini bertujuan untuk melakukan segmentasi pelanggan menggunakan algoritma K-Means berdasarkan data transaksi penjualan elektronik.Tahapan penelitian diawali dengan proses preprocessing data yang meliputi pembersihan data (data cleaning), transformasi data, dan normalisasi untuk menghasilkan dataset yang siap digunakan dalam proses clustering. Selanjutnya, dilakukan proses pengelompokan data menggunakan algoritma K-Means dengan menguji beberapa nilai jumlah cluster (K). Penentuan jumlah cluster optimal dilakukan dengan menggunakan metode evaluasi Davies-Bouldin Index (DBI), dimana nilai DBI yang lebih kecil menunjukkan kualitas cluster yang lebih baik. Hasil penelitian menunjukkan bahwa jumlah cluster optimal diperoleh pada K = 2 dengan nilai DBI sebesar 0,073, yang mengindikasikan bahwa cluster yang dihasilkan memiliki tingkat pemisahan yang baik. Berdasarkan hasil clustering, dilakukan interpretasi terhadap nilai centroid untuk mengidentifikasi karakteristik masing-masing segmen pelanggan. Hasil analisis menunjukkan bahwa pelanggan terbagi menjadi dua kelompok utama, yaitu pelanggan dengan tingkat pembelian tinggi dan pelanggan dengan tingkat pembelian rendah. Dengan demikian, penelitian ini menunjukkan bahwa penerapan algoritma K-Means yang terintegrasi dengan tahapan preprocessing dan evaluasi menggunakan Davies-Bouldin Index (DBI) mampu menghasilkan segmentasi pelanggan yang akurat dan dapat digunakan sebagai dasar dalam penyusunan strategi pemasaran berbasis data.

Kata kunci: Segmentasi Pelanggan, K-Means, Clustering, Davies-Bouldin Index

INTRODUCTION

Perkembangan teknologi informasi yang pesat telah mendorong meningkatnya aktivitas transaksi pada berbagai sektor bisnis, khususnya dalam bidang penjualan elektronik dan e-commerce. Setiap transaksi yang terjadi menghasilkan data dalam jumlah besar yang terus bertambah seiring waktu. Data tersebut mencakup berbagai informasi penting seperti jumlah pembelian, nilai transaksi, serta pola perilaku pelanggan. Namun, banyak perusahaan yang belum memanfaatkan data tersebut secara optimal sebagai dasar dalam pengambilan keputusan bisnis.

Dalam kondisi ini, diperlukan suatu pendekatan berbasis data yang mampu mengolah dan menganalisis data transaksi secara efektif. Salah satu pendekatan yang dapat digunakan adalah data mining, khususnya teknik clustering untuk segmentasi pelanggan. Segmentasi pelanggan merupakan proses pengelompokan pelanggan ke dalam beberapa kelompok berdasarkan karakteristik tertentu, sehingga perusahaan dapat memahami perilaku pelanggan dengan lebih baik. Penelitian menunjukkan bahwa segmentasi pelanggan memiliki peran penting dalam meningkatkan efektivitas strategi pemasaran dan membantu perusahaan dalam mengambil keputusan yang lebih tepat [1]; [2]Salah satu metode yang широко digunakan dalam segmentasi pelanggan adalah algoritma K-Means. Algoritma ini mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik dengan cara meminimalkan jarak antar data dalam satu cluster. K-Means memiliki keunggulan dalam hal kesederhanaan, efisiensi, serta kemampuannya dalam menangani data dalam jumlah besar [3]. Berbagai penelitian telah membuktikan bahwa K-Means efektif dalam mengelompokkan pelanggan berdasarkan pola pembelian, sehingga dapat digunakan sebagai dasar dalam penyusunan strategi pemasaran [4]; [5]

Beberapa penelitian terdahulu juga menunjukkan bahwa penerapan K-Means tidak hanya terbatas pada segmentasi pelanggan, tetapi juga dapat digunakan untuk mendukung berbagai aspek bisnis seperti analisis strategi pemasaran digital [6] serta manajemen stok berbasis pola pembelian pelanggan [7]. Selain itu, penggunaan metode optimasi seperti Elbow juga telah diterapkan untuk menentukan jumlah cluster yang optimal dalam proses clustering [8]

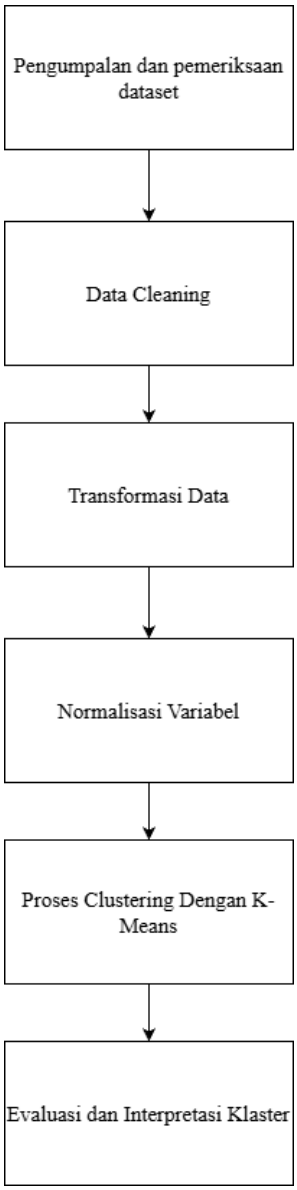
Meskipun demikian, masih terdapat beberapa permasalahan dalam penelitian-penelitian sebelumnya. Sebagian besar penelitian cenderung berfokus pada hasil akhir clustering tanpa menekankan pentingnya tahapan preprocessing data secara komprehensif. Padahal, kualitas data sangat berpengaruh terhadap hasil segmentasi yang dihasilkan [9]; [10]. Selain itu, penentuan jumlah cluster sering kali dilakukan tanpa evaluasi kuantitatif yang kuat, sehingga berpotensi menghasilkan segmentasi yang kurang optimal. Di sisi lain, interpretasi hasil clustering juga masih terbatas dan belum sepenuhnya dimanfaatkan untuk mendukung pengambilan keputusan bisnis.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk melakukan segmentasi pelanggan pada data penjualan elektronik menggunakan algoritma K-Means dengan pendekatan yang lebih terstruktur. Penelitian ini mengintegrasikan tahapan preprocessing data yang meliputi pembersihan data, transformasi, dan normalisasi sebelum dilakukan proses clustering. Selanjutnya, penentuan jumlah cluster dilakukan melalui pengujian beberapa nilai K dan dievaluasi menggunakan metode Davies-Bouldin Index (DBI) untuk memperoleh hasil yang optimal. Hasil clustering kemudian diinterpretasikan melalui analisis centroid untuk mengidentifikasi karakteristik masing-masing kelompok pelanggan.

Dengan pendekatan tersebut, diharapkan penelitian ini tidak hanya menghasilkan segmentasi pelanggan yang akurat secara matematis, tetapi juga mampu memberikan insight yang relevan dalam mendukung pengambilan keputusan, khususnya dalam penyusunan strategi pemasaran berbasis data.

MATERIALS AND METHODS

Prosedur penelitian ini merinci langkah-langkah operasional yang dilakukan dalam proses pengolahan dan analisis data untuk menghasilkan segmentasi pelanggan menggunakan algoritma K-Means. Seluruh prosedur dirancang untuk memastikan bahwa data yang digunakan memenuhi syarat kualitas, memiliki format yang sesuai, dan dapat diolah secara optimal oleh metode *clustering*. Berikut enam tahapan utama yang dilakukan dalam penelitian ini.



Gambar 1. Alur Penelitian

Teknik analisis data dalam penelitian ini dilakukan melalui serangkaian tahapan yang dirancang secara sistematis untuk memastikan bahwa proses segmentasi pelanggan dilakukan dengan akurat dan sesuai standar analisis data modern. Tahap pertama adalah *data cleaning*, yaitu proses untuk memastikan bahwa dataset yang digunakan bebas dari kesalahan pencatatan, duplikasi data, nilai kosong, atau nilai ekstrem yang dapat memengaruhi hasil *clustering*. Tahap ini mencakup pemeriksaan manual serta penggunaan operator khusus di RapidMiner seperti Replace Missing Values, yang membantu mengisi nilai kosong menggunakan metode imputasi yang sesuai.

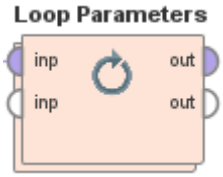
Tahap kedua adalah *data transformation*, yaitu mengonversi variabel non-numerik menjadi numerik menggunakan operator Nominal to Numerical. Transformasi ini penting karena algoritma K-Means hanya dapat memproses data dalam bentuk numerik. Dengan demikian, variabel seperti *Loyalty Member* harus dikonversi ke format angka agar dapat dihitung menggunakan jarak *Euclidean*.

Tahap ketiga adalah *normalisasi*, di mana setiap variabel di normalisasi menggunakan metode *Min-Max* atau *Z-Transformation* untuk memastikan bahwa seluruh atribut berada dalam rentang nilai yang seragam. Normalisasi ini penting agar variabel dengan skala yang lebih besar, seperti *Total Price*, tidak mendominasi variabel lain seperti *Quantity* atau *Age.* Tahap keempat adalah *clustering*, yaitu pengelompokan data menggunakan algoritma K-Means untuk menemukan struktur klaster berdasarkan kemiripan pola belanja pelanggan.

Tahap kelima adalah *evaluasi klaster*, yang dilakukan menggunakan operator Performance, khususnya *Davies-Bouldin Index* (DBI). DBI digunakan untuk mengukur validitas klaster, di mana nilai DBI yang rendah menunjukkan bahwa klaster yang terbentuk memiliki jarak antar-klaster yang baik serta keseragaman data dalam klaster. Tahap terakhir adalah *interpretasi klaster*, yaitu menganalisis nilai *centroid* pada setiap klaster untuk memahami karakteristik setiap kelompok pelanggan. Interpretasi ini memberikan wawasan yang berguna bagi bisnis dalam merancang strategi pemasaran yang tepat berdasarkan profil klaster. Berikut dibawah ini merupakan diagram alur analisis data.

RESULTS AND DISCUSSION

Dilanjutkan menggunakan Operator Loop Parameters digunakan untuk menguji beberapa kombinasi parameter secara otomatis, khususnya nilai K pada algoritma K-Means. Operator ini menjalankan proses *clustering* berulang kali, mengganti nilai K setiap iterasi sesuai rentang yang ditentukan, lalu menampilkan hasil evaluasi seperti *Davies-Bouldin Index* (DBI) untuk setiap nilai K. Dengan cara ini, peneliti dapat menemukan jumlah klaster terbaik berdasarkan nilai DBI terendah.



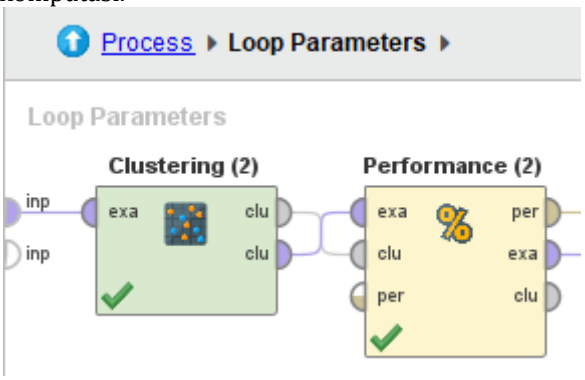
Gambar 2 Operator Loop Parameters

Berikut ini tabel pengaturan parameter operator loop parameter yang digunakan nya

Tabel 1 Parameter Loop Parameters

Nama Parameter	Nilai pada Pengaturan	Fungsi / Keterangan
Edit Parameter Settings…	*Clustering* (K-Means).k	Menentukan parameter yang akan diuji, yaitu nilai K pada algoritma K-Means.
Min	2	Nilai K terkecil yang diuji dalam proses iteratif.
Max	10	Nilai K terbesar yang diuji.
Steps	10	Jumlah iterasi atau banyaknya nilai K yang diuji (2–10).
Scale	Linear	Rentang nilai meningkat secara linear dari Min ke Max.
log performance	✔ (aktif)	Menyimpan hasil evaluasi tiap iterasi (misalnya DBI).
enable parallel execution	✔ (aktif)	Menjalankan iterasi secara paralel agar proses lebih cepat.
error handling	fail on error	Jika terjadi kesalahan pada satu iterasi, proses dihentikan.

Operator Loop Parameters digunakan untuk menjalankan proses pengujian nilai K secara otomatis dalam algoritma K-Means. Berdasarkan pengaturan parameter, nilai K diuji mulai dari K = 2 hingga K = 10 dengan langkah linear, dan setiap iterasi akan dicatat menggunakan fitur log *performance*. Parameter Edit Parameter Settings diarahkan pada K-Means.k, sehingga pada setiap iterasi nilai K berubah secara otomatis sesuai rentang yang telah ditentukan. Pengaturan lain seperti *invert selection* dan *include special attributes* tidak diaktifkan karena tidak diperlukan dalam proses ini, sedangkan opsi *enable parallel execution* dapat mempercepat waktu komputasi.



Gambar 3 Operator didalam Loop Parameters

Di dalam operator Loop Parameters terdapat dua operator inti, yaitu K-Means dan Performance (*Davies-Bouldin Index*). Saat proses looping berjalan, operator K-Means akan dieksekusi berulang kali dengan nilai K yang berubah pada setiap iterasi, sementara operator Performance menghitung nilai *Davies-Bouldin Index* (DBI) untuk setiap hasil *clustering*. Dengan konfigurasi ini, Loop Parameters tidak hanya mengotomatisasi pengujian K, tetapi juga menghasilkan perbandingan kualitas klaster berdasarkan nilai DBI. Nilai DBI terendah dari seluruh iterasi kemudian digunakan sebagai dasar untuk menentukan jumlah klaster terbaik sebelum menjalankan model K-Means final.
Di bawah ini disajikan hasil eksekusi operator Loop Parameters yang memperlihatkan nilai DBI dari setiap iterasi pengujian nilai K.

Loop Parameters (9 rows, 3 columns)

iteration	Clusteri...	Davies ...
1	2	0.179
2	3	0.195
3	4	0.207
4	5	0.224
5	6	0.220
6	7	0.203
7	8	0.210
8	9	0.204
9	10	0.193

Gambar 4 Hasil operator Loop Parameters

Tabel hasil Loop Parameters menunjukkan performa *clustering* untuk beberapa nilai K (jumlah

klaster) berdasarkan indeks Davies-Bouldin (DBI). Semakin rendah nilai DBI, semakin baik kualitas klaster karena menunjukkan jarak antar-klaster yang semakin besar dan struktur klaster yang semakin kompak. Berdasarkan hasil pada tabel, nilai DBI terendah berada pada K = 2 dengan nilai 0.179, hal ini menunjukan bahwa dua klaster memberikan struktur paling kompak dan paling terpisah sehingga diplih sebagai model terbaik.

Setelah proses Loop Parameters menghasilkan beberapa nilai *Davies-Bouldin Index* (DBI) untuk masing-masing kandidat jumlah klaster, tahap berikutnya adalah menjalankan operator K-Means menggunakan nilai K terbaik. Berdasarkan hasil Loop Parameters, nilai DBI terendah berada pada K = 2, sehingga pada tahap ini operator K-Means dikonfigurasi untuk membentuk dua klaster.

Pada proses ini, data hasil *No*rmalisasi dimasukkan ke operator K-Means, dan parameter k = 2 ditetapkan pada bagian pengaturan operator. Selain itu, metode pengukuran jarak (distance measure) tetap menggunakan Euclidean Distance, karena metode ini sesuai untuk data numerik hasil transformasi dan standardisasi sebelumnya.

Gambar 5 Operator K-Means

Dibawah ini merupakan pengaturan parameter pada operator K-Means

Tabel 2 Parameter K-Means

Parameter	Nilai	Penjelasan Singkat
k	2	Jumlah klaster optimal berdasarkan hasil Loop Parameters
distance measure	Euclidean Distance	Mengukur jarak antar titik dalam ruang multidimensional
max runs	10 (default)	Jumlah iterasi untuk stabilisasi *centroid*
max optimization steps	100 (default)	Pembatas iterasi untuk konvergensi model

Pada tahap ini, operator K-Means membentuk dua klaster berdasarkan kedekatan jarak antar data. Proses iteratif dilakukan untuk menentukan titik *centroid* yang paling stabil, sehingga dua kelompok pelanggan dengan karakteristik yang berbeda dapat terbentuk dengan jelas. Di bawah ini merupakan hasil pengelompokan akhir yang dihasilkan oleh operator K-Means dengan jumlah klaster optimal (K = 2).

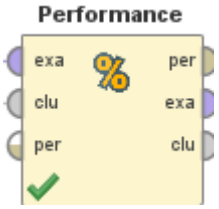
Tabel 3 Hasil Operator K-Means

Row No.	**id**	**Cluster**	***Loyalty Member = No***	**...**	***Add-on Total***
1	1	cluster_1	1	...	40.210
2	2	cluster_0	1	...	26.090
3	3	cluster_0	1	...	0
4	4	cluster_0	0	...	60.160
5	5	cluster_0	0	...	35.560
6	6	cluster_0	1	...	65.780
7	7	cluster_1	1	...	0
8	8	cluster_1	1	...	75.330
...	...	...	...	...	43.050
10000	10000	cluster_0	1	...	0

Tabel 3 menampilkan contoh output hasil pengelompokan dari operator K-Means setelah nilai K terbaik diterapkan (K = 2). Setiap baris data mewakili pelanggan dan telah diberi label klaster pada kolom *cluster* (“cluster_0” atau “cluster_1”). Hasil ini menunjukkan bahwa algoritma telah berhasil mengelompokkan pelanggan berdasarkan kesamaan karakteristik seperti *Loyalty Member*, *Age*, *Total Price*, *Unit Price*, *Quantity*, dan *Add-on Total*.

Dari tabel ini terlihat bahwa setiap klaster memiliki pola belanja yang berbeda, misalnya beberapa pelanggan dengan nilai *Total Price* tinggi masuk ke klaster tertentu, sementara pelanggan

dengan *Add-on Total* rendah atau *No*l mengelompok pada klaster lain. Output ini menjadi dasar untuk *analisis lebih lanjut* pada tahap evaluasi dan interpretasi klaster. Tahapan selanjutnya setelah proses pengelompokan menggunakan K-Means adalah evaluasi kualitas klaster dengan operator *Performance*. Operator ini digunakan untuk menghitung metrik evaluasi internal, khususnya *Davies-Bouldin Index* (DBI), yang berfungsi menilai seberapa baik klaster terbentuk. Output dari operator ini berupa skor DBI, di mana nilai yang lebih rendah menunjukkan klaster yang lebih kompak dan lebih terpisah satu sama lain. Pada tahap ini, operator *Performance* memastikan bahwa hasil *clustering* tidak hanya terbentuk secara teknis, tetapi juga memiliki kualitas struktur klaster yang baik.



Gambar 6 Operator Cluster Distance Performance

Parameter pada operator *Performance (Cluster Distance Performance)* diatur menggunakan *Davies-Bouldin Index* (DBI) sebagai kriteria utama untuk menilai kualitas klaster, di mana nilai DBI yang lebih rendah menandakan pemisahan klaster yang lebih baik dan struktur klaster yang lebih kompak. Opsi main criterion only dipilih agar RapidMiner hanya menghitung DBI tanpa metrik tambahan sehingga hasil evaluasi lebih fokus. Parameter *No*rmalize diaktifkan untuk memastikan bahwa perbedaan skala antar atribut tidak memengaruhi perhitungan jarak klaster. Sementara itu, pengaturan maximize tetap digunakan karena RapidMiner secara internal men*Yes*uaikan logika evaluasi DBI, sehingga konfigurasi ini sudah sesuai dan menghasilkan perhitungan performa klaster yang akurat. Berdasarkan pengaturan tersebut, berikut ini ditampilkan hasil evaluasi nilai *Davies-Bouldin Index* dari proses *clustering* yang telah dilakukan.

Tabel 4 Hasil Operator Cluster Distance Performance

Criterion	Value
Davies Bouldin	0.073

Tabel 4.13 menunjukkan hasil evaluasi kualitas klaster menggunakan indeks Davies–Bouldin (DBI). Nilai DBI yang diperoleh adalah sebesar 0.073. Nilai ini tergolong sangat rendah, yang berarti klaster yang terbentuk memiliki jarak antar-klaster yang jauh serta struktur internal yang kompak. Dengan kata lain, setiap klaster memiliki karakteristik yang jelas dan tidak tumpang tindih satu sama lain. Semakin kecil nilai DBI, semakin baik kualitas pengelompokan, sehingga nilai 0.073 menunjukkan bahwa hasil *clustering* pada model ini sangat optimal dan layak digunakan untuk tahap interpretasi serta penarikan kesimpulan.

Tabel tersebut menunjukkan nilai *Davies-Bouldin Index* (DBI) yang dihasilkan dari proses evaluasi klaster menggunakan algoritma K-Means. Nilai DBI sebesar 0.073 menandakan bahwa kualitas klaster yang terbentuk sangat baik. Semakin kecil nilai DBI, semakin rapat anggota dalam satu klaster dan semakin jauh jarak antar klaster yang lain. Dengan nilai yang mendekati *No*l, dapat disimpulkan bahwa pemisahan antar klaster sudah optimal dan hasil segmentasi pelanggan tergolong sangat baik.

CONCLUSION

Berdasarkan rangkaian tahapan penelitian yang telah dilakukan, mulai dari pengumpulan dan pemeriksaan dataset, data cleaning, transformasi data, Normalisasi, proses *clustering* dengan K-Means, hingga evaluasi dan interpretasi klaster, maka diperoleh beberapa kesimpulan sebagai berikut:

1. Penelitian berhasil membentuk segmentasi pelanggan menggunakan algoritma K-Means dengan kualitas klaster yang sangat baik. Nilai *Davies-Bouldin Index* (DBI) sebesar 0.073 menunjukkan tingkat pemisahan klaster yang optimal, di mana setiap klaster memiliki jarak yang jauh dan struktur yang kompak. Hal ini menandakan bahwa seluruh tahapan preprocessing meliputi Select Attributes, Replace Missing Values, Nominal to Numerical, dan Normalize telah memberikan dampak signifikan terhadap stabilitas model.

2. Model *clustering* menghasilkan dua klaster utama pelanggan dengan karakteristik yang berbeda.

a. Cluster_0 (High-Value Segment) terdiri dari pelanggan dengan nilai transaksi lebih tinggi, kuantitas pembelian lebih banyak, kecenderungan memilih produk premium, dan memiliki potensi bisnis yang besar.

b. Cluster_1 (Low-Value Segment) merupakan pelanggan dengan nilai transaksi dan kuantitas pembelian lebih rendah serta lebih sensitif terhadap harga.

3. *Centroid* klaster menunjukkan pola pembelian yang jelas, terutama pada atribut *Total Price*, *Quantity*, dan *Unit Price*. Hal ini menegaskan bahwa variabel-variabel tersebut merupakan penentu utama dalam membedakan perilaku pelanggan.

4. Scatter plot klaster memperkuat validitas segmentasi, di mana distribusi kedua klaster tampak jelas dan tidak saling tumpang tindih. Visualisasi ini konsisten dengan hasil DBI yang sangat rendah.

5. Secara keseluruhan, penelitian ini membuktikan bahwa algoritma K-Means dapat digunakan secara efektif sebagai alat analisis pemasaran berbasis data, serta mampu memberikan wawasan strategis yang dapat dimanfaatkan dalam penyusunan promosi, program loyalitas, hingga strategi upselling dan *cross-selling*.

REFERENCE

[1] N. Gautam and N. Kumar, "Customer Segmentation Using K-Means Clustering for Developing Sustainable Marketing Strategies," 2022.
[2] J. Zhao, "Customer Segmentation Application Based on K-Means," 2022.
[3] A. B. Gopakumar, G. Krishnan, and N. L. Rozario, "Customer Segmentation Using K-Means Algorithm," 2022.
[4] B. Apriyanto and S. L. M. Sitio, "Penerapan K-Means dalam Menganalisis Pola Pembelian Pelanggan Pada Data Transaksi E-Commerce," 2023.
[5] C. H. Ardana, A. A. A. Aisyah Ainur Khoyum, and M. Faisal, "Segmentasi Pelanggan Penjualan Online Menggunakan Metode K-Means Clustering," 2023.
[6] Anugra, S. Al Tarif, Muh. S. Natsir, and R. A. Djamro, "Analisis Strategi Pemasaran Digital Dalam Memperluas Jangkauan Pemasaran Produk Toko Algifari Dengan Menggunakan K-Means," 2023.
[7] Vina, N. Rahaningsih, and I. Ali, "Segmentasi Data Transaksi Penjualan Toko Online Vastyle Menggunakan Algoritma K-Means Clustering," 2023.
[8] Y. Adryan, F. Rizqi, I. Bayu, R. Yosafat, and Saprudin, "Analisis Segmentasi Pelanggan Mall Menggunakan Metode K-Means Clustering dan Elbow untuk Optimasi Strategi Pemasaran," *Jurnal Riset dan Inovasi Informatika*, 2023.
[9] A. Leto, R. Aminullah, and A. D. Rahajoe, "Customer Data Management Analysis for Customer Segmentation Using K-Means Clustering Method," 2023.
[10] K. Tabianan, S. Velu, and V. Ravi, "K Means clustering approach for intelligent customer segmentation using customer purchase behavior data," *Sustainability*, vol. 14, no. 12, p. 7243, 2022, doi: 10.3390/su14272443.