

PENERAPAN METODE K-MEANS CLUSTERING UNTUK IDENTIFIKASI PROFIL TRANSAKSI DAN KARAKTERISTIK NASABAH PERBANKAN

Natasya Hayudyo Murthiningtyas¹, Farah Bianca², Siti Mukaromah³, Agung Brastama Putra⁴

^{1,2,3,4}Universitas Pembangunan Nasional “Veteran” Jawa Timur, Indonesia

¹natasyahm@gmail.com, ²farahbianca05@gmail.com

³sitimukaromah.si@upnjatim.ac.id

⁴agungbp.si@upnjatim.ac.id

Received: 03-06- 2026

Revised: 20-06-2026

Approved: 27-06-2026

ABSTRACT

Perkembangan teknologi data telah mendorong sektor perbankan untuk memanfaatkan data transaksi sebagai sumber informasi dalam mendukung pengambilan keputusan. Volume transaksi yang terus meningkat menghasilkan data dalam jumlah besar yang mengandung pola dan karakteristik transaksi yang bisa diteliti lebih mendalam. Penelitian ini bertujuan untuk mengidentifikasi profil transaksi dan karakteristik kelompok transaksi melalui penerapan teknik pengelompokan K-Means. Jenis data yang digunakan adalah data transaksi perbankan sebanyak 1.348 transaksi yang terdiri atas atribut waktu transaksi, kode transaksi, sistem transaksi, nominal transaksi, dan jenis transaksi. Sebelum proses clustering, data melalui tahap preprocessing, yang mencakup seleksi atribut, transformasi data, dan normalisasi. Metode Elbow digunakan untuk menghitung jumlah cluster, dan algoritma K-Means digunakan untuk mengelompokkan. Hasil penelitian menunjukkan bahwa 1.348 data transaksi telah dikelompokkan ke dalam lima cluster (A-E). Cluster B memiliki jumlah transaksi terbesar, sedangkan Cluster E memiliki jumlah transaksi paling sedikit namun memiliki rata-rata nominal transaksi tertinggi. Hasil pengelompokan menunjukkan kemampuan metode K-Means untuk mengidentifikasi pola transaksi berdasarkan ciri-cirinya dan memberikan informasi yang bermanfaat dalam memahami profil transaksi dan mendukung proses pengambilan keputusan di lingkungan perbankan.

Kata Kunci: *K-Means Clustering, Preprocessing Data, Segmentasi Transaksi, Profil Transaksi, Perbankan.*

PENDAHULUAN

Era Globalisasi dan revolusi industri 4.0 telah mendorong transformasi digital sebagai faktor utama dalam menghadapi tantangan dan peluang di berbagai bidang, termasuk di sektor perbankan. Digitalisasi industri perbankan yang berjalan seiring pesatnya kemajuan teknologi informasi mendorong Indonesia untuk bersiap menghadapi era digital banking dimana semua layanan kedepannya akan dilakukan secara online dan tidak lagi membutuhkan kantor cabang secara fisik [1]. Dunia perbankan menghadapi tekanan yang signifikan dari perubahan ekonomi global, terutama dengan munculnya dinamika baru yang dipicu oleh faktor-faktor seperti fluktuasi suku bunga, perkembangan teknologi, dan tuntutan regulasi yang semakin ketat. Sebagai lembaga kepercayaan masyarakat yang berperan dalam menyimpan dan mengelola keuangan, bank dituntut untuk terus berinovasi dalam memberikan layanan yang lebih cepat, efisien, dan aman kepada nasabah.

Pertumbuhan volume transaksi perbankan yang sangat pesat setiap harinya menghasilkan akumulasi data dalam skala besar yang menyimpan

informasi berharga mengenai perilaku dan karakteristik nasabah. Data transaksi teller, yang mencakup ratusan hingga ribuan entri per hari, mencerminkan pola aktivitas finansial nasabah yang apabila dianalisis secara tepat dapat menghasilkan insight strategis bagi manajemen perbankan. Dalam konteks ini, pemanfaatan teknik machine learning dan data mining menjadi sangat relevan. Dengan menerapkan algoritma machine learning pada analisis nasabah, bank memperoleh daya prediktif yang lebih baik, memungkinkan bank untuk mengidentifikasi potensi nasabah deposit dengan lebih tepat dan efisien [2].

Salah satu teknik machine learning yang banyak digunakan dalam analisis data perbankan adalah K-Means Clustering, yaitu algoritma pengelompokan berbasis unsupervised learning yang bekerja secara otomatis untuk mengidentifikasi pola tersembunyi di dalam data. Berdasarkan jenis tugas yang dapat dijalankan, data mining dapat dikelompokkan ke dalam beberapa kategori, yaitu deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi. Melalui teknik clustering, bank dapat mengelompokkan nasabah berdasarkan kemiripan karakteristik tertentu, seperti pola transaksi, tingkat pendapatan, preferensi produk, hingga risiko kredit, tanpa memerlukan label awal pada data tersebut. Dengan segmentasi berbasis clustering, bank dapat menyusun strategi pemasaran yang lebih terarah dan personalisasi [3].

Segmentasi nasabah berdasarkan pola transaksi memungkinkan manajemen perbankan untuk memahami distribusi perilaku operasional secara lebih terstruktur, mengidentifikasi kelompok nasabah berdasarkan nilai dan frekuensi transaksi. Penerapan K-Means Clustering dalam penelitian ini juga memungkinkan identifikasi transaksi yang anomali apabila dibandingkan dengan pola cluster yang telah terbentuk, sehingga mendukung fungsi pemantauan risiko operasional secara lebih sistematis. Bank X menghadapi tantangan dalam mengidentifikasi transaksi yang menyimpang dari pola normal, terutama karena keterbatasan ketersediaan data berlabel dalam jumlah besar yang dibutuhkan untuk proses analisis secara konvensional. Untuk mengatasi permasalahan tersebut, diperlukan pendekatan analisis berbasis clustering yang mampu mengolah data transaksi dalam jumlah besar dan mendeteksi anomali secara adaptif terhadap pola data yang kompleks [4].

Dalam upaya memahami pola perilaku transaksi nasabah secara lebih mendalam, studi ini menggunakan pendekatan metode K-Means Clustering pada data transaksi teller di Bank X sebagai objek studi kasus. Metode K-Means adalah metode yang termasuk dalam algoritma clustering berbasis jarak yang membagi data ke dalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut numerik [5]. Penerapan K-Means Clustering diawali dengan tahap pra-pemrosesan data, yang meliputi pembersihan data dari nilai kosong dan duplikat.

METODE PENELITIAN

Studi ini menggunakan metode K-Means Clustering untuk mengkategorikan data transaksi pelanggan berdasarkan kemiripan karakteristik transaksi yang ada. Metode ini dipilih karena termasuk teknik unsupervised learning yang mampu melakukan pengelompokan data tanpa memerlukan data berlabel sebagai acuan. Hasil dari proses clustering digunakan untuk mengidentifikasi profil transaksi serta karakteristik nasabah perbankan, sehingga dapat memberikan informasi yang bermanfaat bagi pihak

bank dalam memahami pola dan perilaku transaksi nasabah secara lebih sistematis. Adapun tahapan penelitian yang dilakukan meliputi proses pengumpulan data, preprocessing data, penentuan jumlah cluster yang akan digunakan, implementasi algoritma K-Means, serta analisis dan penafsiran terhadap hasil pengelompokan yang didapatkan.

Pengumpulan Data

Langkah pertama penelitian dilaksanakan dengan cara mengumpulkan data transaksi dari nasabah. Data yang digunakan berasal dari sistem perbankan yang merekam berbagai aktivitas transaksi nasabah secara rutin. Informasi transaksi yang tersimpan dalam sistem perbankan atau data warehouse dapat dimanfaatkan untuk menganalisis kebutuhan pelanggan bank dan mengamati pola perilaku pelanggan [6]. Dataset tersebut memuat sejumlah atribut yang berkaitan dengan transaksi, seperti tanggal transaksi, waktu transaksi, kode transaksi, nominal transaksi, jenis transaksi (debit atau kredit), serta atribut lain yang mendukung proses analisis. Setelah data terkumpul, dilakukan proses seleksi data untuk menentukan atribut yang memiliki keterkaitan dengan tujuan penelitian. Atribut yang dinilai tidak memberikan kontribusi signifikan terhadap proses pengelompokan akan dieliminasi, sedangkan atribut yang mampu menggambarkan karakteristik transaksi nasabah akan dipertahankan. Tujuan dari langkah ini adalah untuk memperbaiki mutu data yang digunakan, sehingga hasil clustering yang diperoleh dapat lebih optimal dan representatif dalam menggambarkan karakteristik nasabah.

Preprocessing Data

Sebelum proses clustering dilakukan, tahap preprocessing dilaksanakan untuk memperbaiki mutu data terlebih dahulu. Tahapan ini terdiri atas seleksi atribut, transformasi data, dan normalisasi data.

Seleksi Atribut

Proses seleksi atribut dilakukan dengan mengevaluasi setiap atribut pada dataset berdasarkan tingkat relevansinya terhadap tujuan analisis, yaitu mengidentifikasi profil dan karakteristik transaksi nasabah. Atribut yang bersifat sebagai pengidentifikasi unik suatu transaksi (*identifier*), seperti nomor transaksi atau kode teller, tidak digunakan dalam proses *clustering* karena atribut tersebut tidak mencerminkan karakteristik atau pola transaksi, melainkan hanya berfungsi sebagai penanda administratif. Sebaliknya, atribut yang mampu menggambarkan perilaku transaksi nasabah dipertahankan karena memberikan kontribusi terhadap perhitungan jarak antar data pada algoritma K-Means.

Kriteria seleksi atribut yang digunakan dalam penelitian ini mengacu pada dua pertimbangan utama:

1. Relevansi atribut terhadap tujuan penelitian yaitu atribut dipertahankan apabila mampu merepresentasikan pola dan karakteristik transaksi (misalnya waktu transaksi, sistem transaksi, nominal, dan jenis transaksi).
2. Kontribusi atribut terhadap proses pengelompokan atribut yang bersifat unik per baris data (*identifier*) atau tidak memiliki variasi nilai yang

bermakna dieliminasi, karena tidak memberikan informasi tambahan pada perhitungan jarak Euclidean.

Hasil penerapan kriteria seleksi atribut di atas pada data penelitian ditampilkan pada Tabel 1 (subbagian 3.2.1).

Transformasi Data

Atribut kategorikal pada dataset, yaitu *Sys* (sistem transaksi) dan *Tipe* (jenis transaksi), ditransformasikan ke dalam bentuk numerik menggunakan teknik Label Encoding. Teknik ini memetakan setiap kategori unik pada suatu atribut ke dalam nilai integer tunggal, sehingga atribut tersebut dapat digunakan dalam perhitungan jarak Euclidean pada algoritma K-Means.

Untuk atribut *Tipe* yang bersifat biner, fungsi transformasi didefinisikan sebagai:

$$f(x) = \begin{cases} 1, & x = \text{CR (Kredit)} \\ 0, & x = \text{DR (Debit)} \end{cases} \quad (1)$$

Keterangan:

- x = nilai kategori atribut *Tipe*
- $f(x)$ = nilai numerik hasil transformasi

Untuk atribut *Sys* yang memiliki m kategori berbeda $C = \{c_1, c_2, \dots, c_m\}$ (misalnya BOR, INV, GEN), *Label Encoding* dilakukan dengan memberikan indeks integer unik pada setiap kategori:

$$g(c_j) = j, \quad j = 1, 2, \dots, m \quad (2)$$

Keterangan:

- c_j = kategori ke- j pada atribut *Sys*
- $g(c_j)$ = kode integer hasil *encoding* untuk kategori c_j
- m = jumlah kategori unik pada atribut

2.2.3 Normalisasi Data

Untuk menyamakan rentang nilai antar atribut numerik — khususnya atribut *Amount* yang memiliki rentang nilai jauh lebih besar dibandingkan atribut lain — dilakukan normalisasi menggunakan Min-Max Normalization. Metode ini mentransformasikan nilai atribut ke dalam rentang [0,1] berdasarkan nilai minimum dan maksimum atribut tersebut pada dataset, sehingga tidak ada atribut yang mendominasi perhitungan jarak Euclidean pada algoritma K-Means. Formula *Min-Max Normalization* dirumuskan sebagai berikut:

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (3)$$

Keterangan:

- x'_i = nilai atribut ke- i setelah dinormalisasi
- x_i = nilai atribut ke- i sebelum normalisasi
- x_{min} = nilai minimum atribut pada dataset
- x_{max} = nilai maksimum atribut pada dataset

Hasil *preprocessing* menghasilkan sebanyak 1.348 data transaksi yang siap digunakan pada proses *clustering*.

Penentuan Jumlah Cluster

Sebelum memulai proses *clustering*, metode Elbow digunakan untuk menghitung jumlah cluster yang ideal. Metode elbow membantu dalam menentukan berapa banyak cluster yang cocok untuk dataset tertentu [7]. Metode ini digunakan untuk menentukan nilai Sum of Squared Error (SSE) pada variasi jumlah cluster. Jumlah cluster yang dipilih adalah titik yang

menunjukkan penurunan nilai SSE mulai melambat atau membentuk pola siku (*elbow point*).

Implementasi Algoritma K-Means

Setelah jumlah *cluster* optimal diperoleh, dilakukan proses pengelompokan menggunakan algoritma K-Means. Algoritma ini bekerja dengan membagi data ke dalam beberapa kelompok berdasarkan tingkat kemiripan karakteristik transaksi. Langkah-langkah algoritma K-Means adalah sebagai berikut:

1. Menentukan jumlah kelompok (K).
2. Memilih centroid awal secara acak.
3. Mengukur jarak setiap data ke centroid menggunakan jarak Euclidean.
4. Mengelompokkan data berdasarkan kedekatan jarak.
5. Menghitung ulang centroid baru berdasarkan rata-rata anggota *cluster*.
6. Mengulangi langkah 3–5 sampai posisi centroid stabil (konvergen).

Perhitungan jarak Euclidean antara data dan centroid dirumuskan sebagai berikut:

$$(4) \quad d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Keterangan:

- $d(x, y)$ = jarak Euclidean antara data x dan centroid y
- x_i = nilai atribut data ke- i
- y_i = nilai atribut centroid ke- i
- n = jumlah atribut (karakteristik) yang digunakan

Analisis Hasil Clustering

Tahap akhir penelitian dilakukan dengan mengevaluasi dan menginterpretasikan hasil pengelompokan yang telah terbentuk. Setiap cluster dianalisis berdasarkan karakteristik transaksi yang dimiliki oleh anggota di dalamnya. Proses analisis dilakukan dengan meninjau dan membandingkan rata-rata nilai untuk masing-masing atribut pada setiap cluster, sehingga dapat diperoleh gambaran mengenai profil transaksi serta karakteristik nasabah pada setiap kelompok. Selanjutnya, hasil clustering dimanfaatkan untuk mengidentifikasi segmen nasabah berdasarkan tingkat aktivitas transaksinya, seperti kelompok dengan aktivitas tinggi, sedang, maupun rendah. Perbankan dapat menggunakan informasi ini untuk mempertimbangkan strategi pelayanan mereka, menetapkan program yang sesuai dengan kebutuhan nasabah, serta mendukung proses pengambilan keputusan bisnis yang lebih efektif.

HASIL DAN PEMBAHASAN

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data transaksi teller yang diambil dari sistem operasional PT Bank Negara Indonesia (Persero) Tbk selama kegiatan magang. Data yang digunakan terdiri atas 1.348 transaksi yang memuat informasi waktu transaksi, kode transaksi, sistem transaksi, nominal

transaksi, serta jenis transaksi berupa kredit (CR) dan debit (DR). Sebelum dilakukan proses clustering, data terlebih dahulu melalui tahap pemeriksaan dan pembersihan untuk memastikan tidak terdapat data yang tidak lengkap maupun inkonsisten. Selanjutnya data ditransformasikan ke dalam format yang dapat diproses oleh algoritma K-Means Clustering.

Detail transaksi per cluster
 Klik kartu cluster di atas untuk filter 1348 baris

#	CLUSTER	NO	JAM	TELLER	TRAN CODE	SYS	AMOUNT	TIPE
1	Cluster A	107706	08:19	89080	11010	BOR	Rp 225.000.000	CR
2	Cluster A	107707	09:43	89080	1060	INV	Rp 1.800.000	DR
3	Cluster B	107708	10:07	89080	1010	INV	Rp 6.310.000	CR
4	Cluster B	107709	10:25	89080	20010	GEN	Rp 10.000	CR
5	Cluster B	107710	10:28	89080	1010	INV	Rp 100.000	CR
6	Cluster B	107711	10:29	89080	1010	INV	Rp 100.000	CR
7	Cluster B	107712	10:35	89080	1060	INV	Rp 1.800.000	DR
8	Cluster B	107713	10:43	89080	1010	INV	Rp 4.000.000	CR
9	Cluster B	107714	10:52	89080	1060	INV	Rp 48.420.000	DR
10	Cluster B	107715	11:02	89080	1010	INV	Rp 196.132.000	CR
11	Cluster B	107716	11:09	89080	1060	INV	Rp 1.800.000	DR
12	Cluster B	107717	11:19	89080	1010	INV	Rp 7.500.000	CR
13	Cluster B	107718	11:36	89080	1060	INV	Rp 1.800.000	DR
14	Cluster B	107719	11:44	89080	1060	INV	Rp 1.800.000	DR
15	Cluster C	107720	12:25	89080	11010	BOR	Rp 68.495.000	CR

Gambar 1. Dataset Transaksi Nasabah

Preprocessing Data

Untuk memastikan kualitas data, tahap preprocessing dilakukan terhadap data transaksi sebelum data tersebut diproses lebih lanjut dalam tahap clustering yang telah disesuaikan dengan kebutuhan analisis. Tahapan preprocessing dilakukan melalui proses pemeriksaan data, seleksi atribut, transformasi data, dan normalisasi.

Seleksi Atribut

Tidak seluruh atribut yang terdapat pada dataset digunakan dalam proses clustering. Atribut yang hanya berfungsi sebagai identitas transaksi, seperti nomor transaksi dan kode teller, tidak digunakan karena tidak memiliki pengaruh terhadap pola transaksi yang akan dianalisis. Tabel 1 dapat menunjukkan atribut-atribut yang digunakan dalam proses clustering.

Tabel 1. Atribut Proses Clustering

Atribut	Keterangan
Jam Transaksi	Waktu transaksi berlangsung
Tran Code	Kode jenis transaksi

Atribut	Keterangan
Sys	Sistem transaksi
Amount	Nominal transaksi
Tipe	Kredit (CR) atau Debit (DR)

Transformasi Data

Beberapa atribut yang sifatnya kategorikal perlu diubah menjadi angka supaya bisa diolah oleh algoritma K-Means. Transformasi dilakukan pada atribut tipe transaksi dan sistem transaksi. Sebagai contoh, atribut tipe transaksi dikonversi menjadi nilai numerik dengan ketentuan CR bernilai 1 dan DR bernilai 0. Proses transformasi dilakukan agar seluruh atribut dapat digunakan dalam perhitungan jarak antar data.

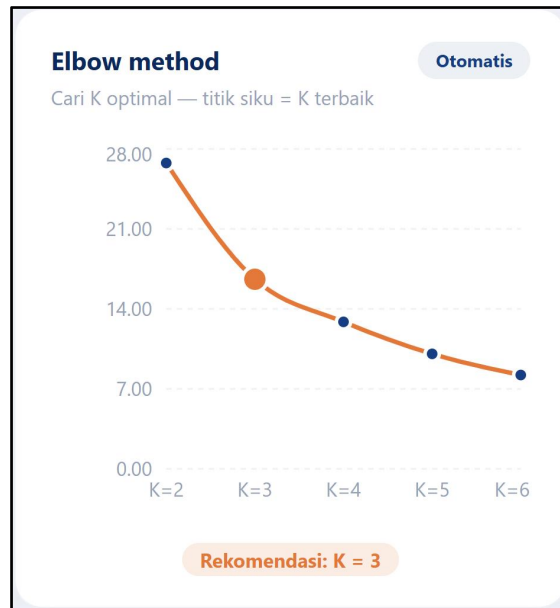
Normalisasi Data

Normalisasi dilakukan untuk menyamakan rentang nilai antar atribut, khususnya pada atribut nominal transaksi yang memiliki perbedaan nilai cukup besar. Langkah ini dilakukan agar tidak ada atribut tertentu yang mendominasi perhitungan jarak Euclidean dalam algoritma K-Means. Hasil preprocessing menghasilkan sebanyak 1.348 data transaksi yang siap digunakan pada proses clustering.

Penentuan Jumlah Cluster Dengan Metode Elbow

Jumlah cluster yang optimal ditentukan dengan menggunakan metode Elbow, yaitu dengan menghitung nilai Sum of Squared Error (SSE) untuk beberapa variasi jumlah cluster. Metode ini bertujuan untuk menemukan jumlah cluster yang dapat menghasilkan pengelompokan data secara optimal.

Hasil pengujian yang ditampilkan pada Gambar 2 menunjukkan bahwa nilai SSE mengalami penurunan yang cukup signifikan pada rentang K=2 hingga K=3. Setelah titik tersebut, penurunan nilai SSE mulai melambat dan cenderung stabil.



Gambar 2. Hasil *Elbow Method*

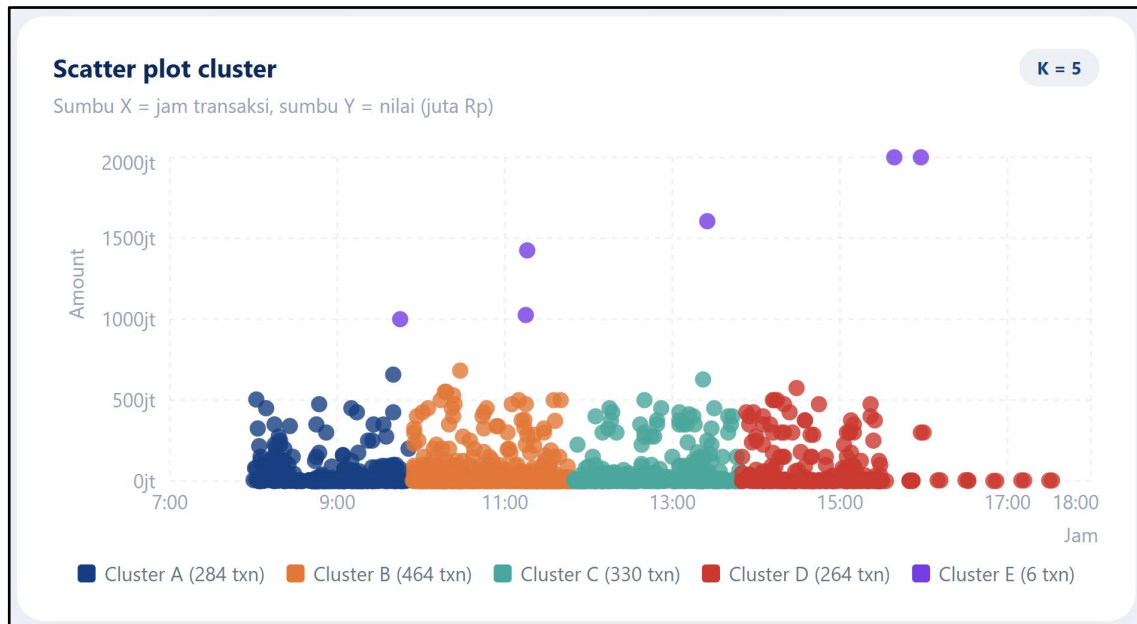
Implementasi Algoritma K-Means Clustering

Meskipun metode Elbow merekomendasikan K=3 sebagai jumlah cluster optimal secara statistik, penelitian ini menggunakan K=5 berdasarkan pertimbangan kebutuhan bisnis perbankan. Penentuan K=5 didasarkan pada kebutuhan segmentasi operasional yang lebih granular, yaitu memisahkan kelompok transaksi reguler bernilai rendah, menengah, dan tinggi. Implementasi algoritma K-Means menghasilkan lima kelompok transaksi yang terdiri atas Cluster A, Cluster B, Cluster C, Cluster D, dan Cluster E. Distribusi jumlah anggota pada masing-masing cluster ditunjukkan pada Tabel 2.

Tabel 2. Distribusi Data pada Setiap Cluster

Cluster	Jumlah Transaksi
Cluster A	285
Cluster B	464
Cluster C	330
Cluster D	264
Cluster E	5
Total	1.348

Berdasarkan hasil clustering, Cluster B memiliki jumlah anggota terbesar dengan 464 transaksi atau sekitar 34,42% dari total data. Sebaliknya, Cluster E merupakan cluster dengan jumlah anggota paling sedikit yaitu sebanyak 5 transaksi.



Gambar 3. Visualisasi Hasil Clustering Menggunakan Scatter Plot

Hasil visualisasi menunjukkan bahwa setiap cluster memiliki pola sebaran yang berbeda berdasarkan pada waktu transaksi dan nominal transaksi. Perbedaan tersebut menunjukkan bahwa algoritma K-Means berhasil mengelompokkan transaksi ke dalam kelompok yang memiliki karakteristik yang sama.

Analisis Karakteristik Cluster

Karakteristik masing-masing cluster dianalisis berdasarkan nilai rata-rata transaksi, nilai minimum, nilai maksimum, serta jumlah transaksi kredit dan debit.

Tabel 3. Karakteristik Setiap Cluster

Cluster	Jumlah Data	Rata-rata	Nilai Minimum	Nilai Maksimum
Cluster A	285	Rp54 juta	Rp10.000	Rp1,00 miliar
Cluster B	464	Rp55 juta	Rp10.000	Rp682 juta
Cluster C	330	Rp56 juta	Rp10.000	Rp628 juta
Cluster D	264	Rp69 juta	Rp10.000	Rp575 juta
Cluster E	5	Rp1,61 miliar	Rp1,03 miliar	Rp2,00 miliar

Berdasarkan hasil analisis, Cluster A hingga Cluster D memiliki rata-rata nilai transaksi yang relatif berdekatan. Sementara itu, Cluster E memiliki rata-rata transaksi yang jauh lebih besar dibandingkan cluster lainnya. Kondisi ini menunjukkan bahwa Cluster E merepresentasikan kelompok transaksi bernilai tinggi (high value transaction) meskipun jumlah transaksinya sangat sedikit. Cluster D memiliki rata-rata transaksi terbesar kedua setelah Cluster E, yaitu

sebesar Rp69 juta. Hal ini menunjukkan bahwa cluster tersebut berisi transaksi dengan nominal relatif tinggi dibandingkan cluster reguler lainnya.

Analisis Distribusi Transaksi Kredit dan Debit

Selain berdasarkan nominal transaksi, karakteristik cluster juga dianalisis berdasarkan komposisi transaksi kredit dan debit.

Tabel 4. Distribusi Transaksi Kredit dan Debit

Cluster	Kredit (CR)	Debit (DR)
Cluster A	194	91
Cluster B	303	161
Cluster C	232	98
Cluster D	162	102
Cluster E	2	3

Visualisasi distribusi transaksi kredit dan debit ditunjukkan pada Gambar 4.



Gambar 4. Distribusi Transaksi Kredit dan Debit

Hasil analisis menunjukkan bahwa Cluster A, Cluster B, Cluster C, dan Cluster D didominasi oleh transaksi kredit. Dominasi transaksi kredit menunjukkan bahwa sebagian besar transaksi pada kelompok tersebut merupakan aktivitas penerimaan dana. Berbeda dengan cluster lainnya, Cluster E memiliki jumlah transaksi debit yang lebih banyak dibandingkan transaksi kredit. Meskipun jumlah anggota cluster ini sangat sedikit, nilai transaksi yang dimiliki tergolong sangat besar sehingga dapat dikategorikan sebagai kelompok transaksi khusus yang memiliki karakteristik berbeda dari kelompok lainnya.

PEMBAHASAN

Dari hasil penelitian yang telah dilakukan, algoritma K-Means mampu mengelompokkan data transaksi menjadi beberapa kelompok berdasarkan kemiripan karakteristik transaksi. Hasil clustering menunjukkan adanya perbedaan pola transaksi pada setiap cluster baik dari sisi jumlah transaksi, nilai transaksi, maupun komposisi transaksi kredit dan debit. Cluster A hingga Cluster D dapat dikategorikan sebagai kelompok transaksi reguler karena memiliki jumlah anggota yang relatif besar dengan nilai transaksi rata-rata yang tidak berbeda jauh. Sementara itu, Cluster E merupakan kelompok transaksi bernilai

tinggi yang memiliki jumlah anggota sangat sedikit namun menunjukkan rata-rata nominal transaksi terbesar. Hasil segmentasi yang diperoleh dapat membantu pihak perbankan dalam memahami karakteristik transaksi yang terjadi pada sistem operasional. Selain itu, informasi mengenai kelompok transaksi bernilai tinggi dapat digunakan sebagai salah satu pertimbangan dalam proses pengambilan keputusan bisnis, pemantauan aktivitas transaksi, serta penyusunan strategi layanan yang lebih sesuai dengan karakteristik transaksi yang dimiliki masing-masing kelompok.

KESIMPULAN

Berdasarkan hasil analisis dan pengelompokan data transaksi dengan penerapan K-Means Clustering terhadap 1.348 data transaksi perbankan berhasil membentuk lima cluster dengan karakteristik yang berbeda. Cluster B merupakan kelompok dengan jumlah transaksi tertinggi, sedangkan Cluster E kelompok dengan jumlah transaksi terendah. Meskipun demikian, Cluster E memiliki rata-rata nominal transaksi paling tinggi dibandingkan cluster lainnya sehingga dapat merepresentasikan kelompok transaksi bernilai besar. Sementara itu, Cluster A, B, C, dan D menunjukkan karakteristik transaksi yang relatif serupa dengan rata-rata nominal transaksi yang tidak berbeda secara signifikan. Sebagian besar cluster didominasi oleh transaksi kredit (CR), yang mengindikasikan bahwa aktivitas penerimaan dana lebih banyak terjadi dibandingkan aktivitas pengeluaran dana.

Hasil segmentasi membuktikan bahwa algoritma K-Means mampu mengidentifikasi profil dan pola transaksi secara terstruktur tanpa memerlukan label awal, sehingga informasi yang dihasilkan dapat dijadikan dasar pengambilan keputusan operasional di lingkungan perbankan.

Secara keseluruhan, hasil penelitian membuktikan bahwa penerapan metode K-Means Clustering dapat membantu proses identifikasi profil transaksi dan karakteristik kelompok transaksi sehingga informasi yang dihasilkan dapat digunakan untuk membantu dalam pengambilan keputusan di lingkungan perbankan. Perbedaan karakteristik yang terbentuk pada setiap cluster menunjukkan bahwa algoritma K-Means mampu memberikan segmentasi transaksi yang dapat digunakan untuk memahami pola aktivitas transaksi secara lebih terstruktur.

DAFTAR PUSTAKA

- [1] M. A. Suharbi dan H. Margono, "Kebutuhan transformasi bank digital Indonesia di era revolusi industri 4.0," *Fair Value J. Ilm. Akunt. Dan Keuang.*, vol. 4, no. 10, hlm. 4749–4759, Mei 2022, doi: 10.32670/fairvalue.v4i10.1758.
- [2] M. R. A. Riyyasy, W. N. Aghniya, dan H. Tantyoko, "Penerapan Algoritma Machine Learning Untuk Memprediksi Term Deposit Nasabah Perbankan," *LEDGER: Journal Informatic and Information Technology*, vol. 2, no. 3, pp. 146–156, Agu. 2023. [Online]. Available: LEDGER (open access).
- [3] D. Irawan, G. Wijaya, dan T. T. Warisaji, "Penerapan Algoritma K-Means Clustering untuk Segmentasi Nasabah Bank," *BIOS J. Teknol. Inf. Dan Rekayasa Komput.*, vol. 6, no. 1, hlm. 47–53, Mar 2025, doi: 10.37148/bios.v6i1.162.

- [4] F. M. Ayudewi dan U. R. Rafifah, "PENERAPAN DETEKSI ANOMALI MENGGUNAKAN MODEL ISOLATION FOREST UNTUK ANALISIS FRAUD PADA TRANSAKSI KEUANGAN AKUNTANSI," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 10, no. 3, 2026, doi: 10.36040/jati.v10i3.18334.
- [5] M. Norshahlan, H. Jaya, dan R. Kustini, "Penerapan Metode Clustering Dengan Algoritma K-means Pada Pengelompokan Data Calon Siswa Baru," *J. Sist. Inf. Triguna Dharma JURSI TGD*, vol. 2, no. 6, hlm. 1042, Nov 2023, doi: 10.53513/jursi.v2i6.9148.
- [6] I. Irawan, R. Rahman, dan A. Wibowo, "PENGELOMPOKAN TRANSAKSI KARTU DEBIT PERBANKAN MENGGUNAKAN ALGORITMA K-MEANS," *J. Sist. Inf. Dan Inform. Simika*, vol. 8, no. 1, hlm. 99–110, Jan 2025, doi: 10.47080/simika.v8i1.3558.
- [7] N. A. Maori dan E. Evanita, "Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering," *Simetris J. Tek. Mesin Elektro Dan Ilmu Komputer.*, vol. 14, no. 2, hlm. 277–288, Nov 2023, doi: 10.24176/simet.v14i2.9630.