



Analisis Klustering Menggunakan Algoritma DBSCAN untuk Deteksi Anomali dalam Data Transaksi Keuangan

Clustering Analysis Using DBSCAN Algorithm for Anomaly Detection in Financial Transaction Data

Buchori Alwi & Rizki Muliono

Teknik informatika, Teknik, Universitas Medan Area, Indonesia

Diterima: 16 April 2025; Direview: 20 April 2025; Disetujui: 24 April 2025

*Corresponding Email:

Abstrak

Deteksi anomali pada data transaksi keuangan menjadi aspek krusial seiring meningkatnya penggunaan e-money yang berpotensi memunculkan aktivitas mencurigakan seperti fraud dan pencucian uang. Penelitian ini menerapkan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) untuk mengelompokkan data transaksi dan mengidentifikasi anomali berdasarkan variabel jumlah transaksi, frekuensi transaksi, dan saldo akhir. Penentuan parameter optimal dilakukan dengan mengevaluasi kombinasi nilai epsilon (ϵ) dan minPts menggunakan Davies-Bouldin Index (DBI) sebagai indikator kualitas kluster. Hasil analisis menunjukkan bahwa parameter optimal adalah ϵ sebesar 0.2727 dan minPts sebesar 6, dengan nilai DBI sebesar 1.1753. DBSCAN berhasil membentuk enam cluster utama dan mendeteksi 138 data sebagai noise yang mengindikasikan potensi transaksi tidak wajar. Temuan ini membuktikan bahwa DBSCAN mampu memisahkan data normal dan mencurigakan secara efektif tanpa asumsi awal jumlah cluster, serta berkontribusi dalam pengembangan sistem deteksi anomali transaksi digital yang lebih akurat dan adaptif.

Kata Kunci: DBSCAN; Klasterisasi; Deteksi Anomali; Transaksi Keuangan; Davies-Bouldin Index.

Abstract

Anomaly detection in financial transaction data is a crucial aspect due to the increasing use of e-money, which raises the risk of suspicious activities such as fraud and money laundering. This study applies the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to cluster transaction data and identify anomalies based on three main variables: transaction amount, transaction frequency, and final balance. The optimal parameters were determined by evaluating various combinations of epsilon (ϵ) and minPts values using the Davies-Bouldin Index (DBI) as a clustering quality indicator. The analysis results indicate that the optimal parameters are ϵ of 0.2727 and minPts of 6, with a DBI score of 1.1753. DBSCAN successfully formed six main clusters and detected 138 data points as noise, indicating potentially abnormal transactions. These findings demonstrate that DBSCAN can effectively distinguish between normal and suspicious data without requiring prior assumptions on the number of clusters, contributing to the development of more accurate and adaptive digital transaction anomaly detection systems.

Keywords: DBSCAN; Clustering; Anomaly Detection; Financial Transactions; Davies-Bouldin Index

PENDAHULUAN

Deteksi anomali merupakan salah satu cabang penting dalam ilmu data mining dan machine learning yang berfokus pada identifikasi pola atau observasi yang menyimpang dari perilaku normal. Dalam ilmu komputer, anomali didefinisikan sebagai data yang berbeda secara signifikan dari mayoritas data, sering kali berkaitan dengan kejadian langka, kesalahan sistem, atau aktivitas mencurigakan [1]. Dalam konteks transaksi keuangan, deteksi anomali menjadi krusial untuk mengidentifikasi potensi penipuan, pencucian uang, dan aktivitas ilegal lainnya yang dapat merugikan konsumen dan lembaga keuangan.

Salah satu pendekatan yang banyak digunakan dalam mendeteksi anomali adalah clustering, yaitu metode unsupervised learning yang mengelompokkan data berdasarkan kemiripan atributnya tanpa memerlukan label. Clustering membantu mengidentifikasi kelompok data mayoritas (normal) dan memisahkan data minoritas (anomali) berdasarkan distribusi spasialnya [2]. Namun, metode clustering tradisional seperti K-Means sering kali kurang efektif dalam mendeteksi outlier karena kecenderungan asumsinya terhadap bentuk cluster yang globular dan jumlah cluster yang ditentukan di awal.

Sebagai alternatif, algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) dikembangkan untuk mengatasi kelemahan tersebut [3]. DBSCAN mengelompokkan data berdasarkan densitas lokal dengan dua parameter utama, epsilon (ϵ) dan minPts, sehingga mampu mendeteksi cluster berbentuk arbitrer serta secara otomatis mengidentifikasi noise atau outlier [4]. Keunggulan ini menjadikan DBSCAN ideal untuk aplikasi deteksi anomali pada data dengan distribusi kompleks, seperti data transaksi keuangan.

Penelitian sebelumnya telah membuktikan efektivitas DBSCAN dalam berbagai konteks. Boer et al. menggunakan DBSCAN untuk mengelompokkan wilayah rawan bencana di Sumatera Utara [5]. Penelitian tersebut berhasil membentuk dua cluster utama dengan nilai silhouette coefficient sebesar 0,46, yang menunjukkan kualitas klaster yang cukup baik dalam memetakan area risiko berdasarkan densitas kejadian bencana. Penelitian ini membuktikan bahwa DBSCAN mampu mengelola data spasial dengan distribusi densitas yang tidak seragam.

Di bidang lain, Prihatmono menerapkan DBSCAN untuk melakukan klusterisasi pengunjung mall berdasarkan usia dan pendapatan [6]. Dari 200 data yang dianalisis, DBSCAN berhasil membentuk 4 cluster utama dan mengidentifikasi 1 cluster sebagai noise (outlier). Hasil penelitian tersebut menunjukkan bahwa DBSCAN efektif dalam membedakan kelompok pengunjung dengan karakteristik yang berbeda, tanpa memerlukan asumsi awal mengenai jumlah cluster.

Dalam konteks transaksi keuangan, penelitian oleh Polgan et al. mengembangkan model deteksi anomali menggunakan Support Vector Machine (SVM) berbasis data mining [7]. Meskipun pendekatan ini efektif, peneliti menyoroti perlunya eksplorasi metode lain seperti DBSCAN untuk menangani data dengan distribusi kompleks dan deteksi outlier secara lebih fleksibel. Hal ini menunjukkan adanya peluang untuk menerapkan DBSCAN sebagai alternatif yang lebih adaptif.

Berdasarkan kajian tersebut, penerapan DBSCAN untuk deteksi anomali pada data transaksi keuangan berbasis e-money di Indonesia masih relatif terbatas. Celah penelitian (*research gap*) ini menjadi landasan bagi penelitian ini untuk mengeksplorasi penggunaan DBSCAN dalam mendeteksi transaksi tidak wajar.

Penelitian ini bertujuan untuk menerapkan algoritma DBSCAN dalam mendeteksi anomali pada data transaksi keuangan, dengan fokus pada variabel jumlah transaksi, frekuensi transaksi, dan saldo akhir. Penentuan parameter optimal dilakukan dengan menggunakan evaluasi Davies-Bouldin Index (DBI). Secara teoretis, penelitian ini berkontribusi dalam memperkaya literatur tentang penerapan DBSCAN dalam deteksi anomali transaksi keuangan. Secara praktis, penelitian ini diharapkan memberikan solusi untuk mendeteksi aktivitas mencurigakan secara lebih akurat, sehingga mendukung upaya pengawasan dan keamanan transaksi digital di Indonesia.

METODE PENELITIAN

Pra pemrosesan

Pra pemrosesan adalah dalam konteks pemrosesan data adalah sekumpulan teknik atau langkah yang diterapkan ke data mentah sebelum masuk ke pemodelan. Tujuan dari pra pemrosesan adalah untuk membersihkan data, normalisasi, dan pengisian data kosong sehingga data dapat diolah dengan lebih efektif dan tepat [8]. Salah satu teknik yang digunakan peneliti untuk menormalisasi data adalah z-score

$$n = \frac{x - \mu}{\sigma}$$

Dimana:

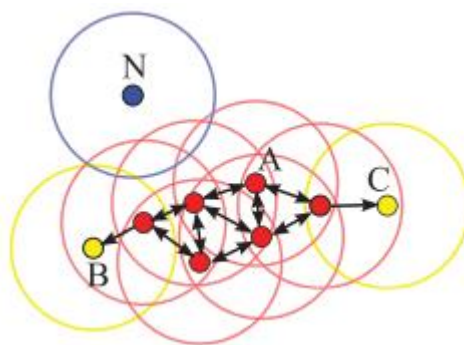
n = hasil penjumlahan

μ = rata-rata dalam sebuah kolom data

σ = standar deviasi data

DBSCAN (Density-Based Spatial Clustering of Application with Noise)

DBSCAN merupakan algoritma yang mengelompokkan titik-titik data berdasarkan tingkat kepadatan dan mengabaikan data noise atau outlier [9]. Menurut [10] DBSCAN dapat mendeteksi cluster dengan data yang sedikit dan berubah-ubah. Untuk menemukan kluster DBSCAN menghitung jarak epsilon(ϵ) dan menghitung minPts yang ada disekitarnya [11]. Semakin kecil nilai epsilon(ϵ) maka semakin banyak kluster yang terbentuk [12]. Algoritma ini juga cocok untuk mengelompokkan data sedang, dan mampu membedakan bentuk kluster dengan ukuran yang berbeda [13]. DBSCAN. Bisa dilihat gambar 1. DBSCAN memiliki 3 pengelompokkan N merupakan noise point, A merupakan core point, B dan C adalah border point. core point adalah titik yang memenuhi syarat kurang dari nilai epsilon dan minPts yang sesuai, border point adalah titik yang memenuhi epsilon namun tidak memenuhi minPts sedangkan noise adalah titik yang tidak memenuhi epsilon maupun minPts [14].



Gambar 1. Ilustrasi model cluster DBSCAN [15]

DBSCAN memiliki dua parameter utama yaitu:

Epsilon (ϵ): Yaitu jarak maksimum antara titik dalam sebuah kluster

MinPts: Jumlah minimum tetangga suatu titik dalam sebuah kluster

Adapun tahapan dari DBSCAN:

1. menentukan nilai *epsilon* dan minPts
2. menghitung semua jarak *epsilon* dengan rumus *Euclidian distance*

$$E = \sqrt{\sum_{i=1}^i (x_i - xp_i)^2}$$

E = Jarak antar titik

x_i = kordinat titik tujuan

x_{p_i} = kordinat pusat

3. Klaster terbentuk berdasarkan jika memenuhi syarat *epsilon* dan minPts
4. Menguji klaster menggunakan Davies-Bouldin Index (DBI) dengan persamaan [16]:

$$DBI = \frac{1}{K} \sum_{i=1}^K \max_{j \neq i} \frac{s_i + s_j}{d_{ij}}$$

K = Jumlah kluster

S = Rata-rata seluruh titik masing-masing kluster i

D = Jarak antara pusat kluster

Semakin kecil nilai evaluasi maka klastering akan semakin baik

Perhitungan Manual DBSCAN

Dataset yang digunakan peneliti bersumber dari Kaggle. Dataset terdiri dari 15 variabel diantaranya, Id_Transaksi, id_akun, jumlah_transaksi, tanggal_transaksi, tipe_transaksi, lokasi, id_perangkat, ip_address, id_penjual, channel, umur_konsumen, pekerjaan, durasi_transaksi, frekuensi_transaksi, dan saldo_akhir. Variabel bisa dilihat pada tabel 1.

Tabel 1. Variabel dataset

No	Id_Transaksi	id_akun	Jumlah_transaksi	tanggal_transaksi	tipe_transaksi	lokasi	id_perangkat	ip_address
1	TX000001	AC00128	14.09	2023-04-11 16:29:14	Debit	San Diego	D000380	162.198.218.92
2	TX000002	AC00455	376.24	2023-06-27 16:44:19	Debit	Houston	D000051	13.149.61.4
3	TX000003	AC00019	126.29	2023-07-10 18:16:08	Debit	Mesa	D000235	215.97.143.157
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2510	TX002510	AC00095	28.63	2023-08-21 17:08:50	Debit	San Diego	D000095	58.1.27.124
2511	TX002511	AC00118	185.97	2023-02-24 16:24:46	Debit	Denver	D000634	21.190.11.223



251	TX00251	AC000	243.08	2023-02-14	Credit	Jacksonville	D000215	59.127.135.25
2	2	09		16:21:23				

No	id_penjual	channel	umur_konsumen	perkerjaan	durasi_transaksi	frekuensi_transaksi	saldo_akhir
1	M015	ATM	70	Doctor	81	1	5112.21
2	M052	ATM	68	Doctor	141	1	13758.91
3	M009	Online	19	Student	56	1	1122.35
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2510	M087	Branch	56	Retired	146	1	3382.91
2511	M041	Online	23	Student	19	1	1776.91
2512	M041	Online	24	Student	93	1	131.25

Sebelum dilakukan perhitungan. Peneliti menghapus beberapa variabel yang tidak relevan dengan pembahasan penelitian. Di tabel 2 diperlihatkan tabel yang relevan dengan penelitian ini. Selanjutnya tabel tersebut akan dilakukan pengukuran jarak menggunakan *Euclidian distance* :

Tabel 2. Dataset transaksi yang relevan				
no	Id_Transaksi	Jumlah_Transaksi (\$)	Saldo_Akhir (\$)	Frekuensi_transaksi
1	TX002082	225.10	1547.71	1
2	TX000434	85.35	2173.76	1
3	TX001725	16.27	10971.32	1
4	TX001252	95.78	271.64	1
5	TX000094	442.12	1375.80	1
6	TX000903	328.34	2840.45	2
7	TX002055	156.86	14453.35	1
8	TX000768	327.80	10585.48	4
9	TX000600	435.50	6981.03	1
10	TX001230	17.21	914.16	2

Kemudian lakukan perhitungan jarak menggunakan rumus *Euclidian distance*. Berikut adalah contoh nya:

$$E(1,2) = \sqrt{(85.35 - 225.10)^2 + (2173.76 - 1547.71)^2 + (1 - 1)^2} = 641,458$$

$$E(1,3) = \sqrt{(16.27 - 225.10)^2 + (10971.32 - 1547.71)^2 + (1 - 1)^2} = 9.425,923$$

$$E(1,4) = \sqrt{(95.78 - 225.10)^2 + (271.64 - 1547.71)^2 + (1 - 1)^2} = 4037.29$$

$$E(1,5) = \sqrt{(442.12 - 225.10)^2 + (1375.80 - 1547.71)^2 + (1 - 1)^2} = 277.06$$

$$E(1,6) = \sqrt{(328.34 - 225.10)^2 + (2840.45 - 1547.71)^2 + (2 - 1)^2} = 1296.35$$



$$E(1,7) = \sqrt{(156.86 - 225.10)^2 + (14453.35 - 1547.71)^2 + (1 - 1)^2} = 12917.67$$

$$E(1,8) = \sqrt{(327.80 - 225.10)^2 + (10585.48 - 1547.71)^2 + (4 - 1)^2} = 9043.20$$

$$E(1,9) = \sqrt{(435.50 - 225.10)^2 + (6981.03 - 1547.71)^2 + (1 - 1)^2} = 5442.34$$

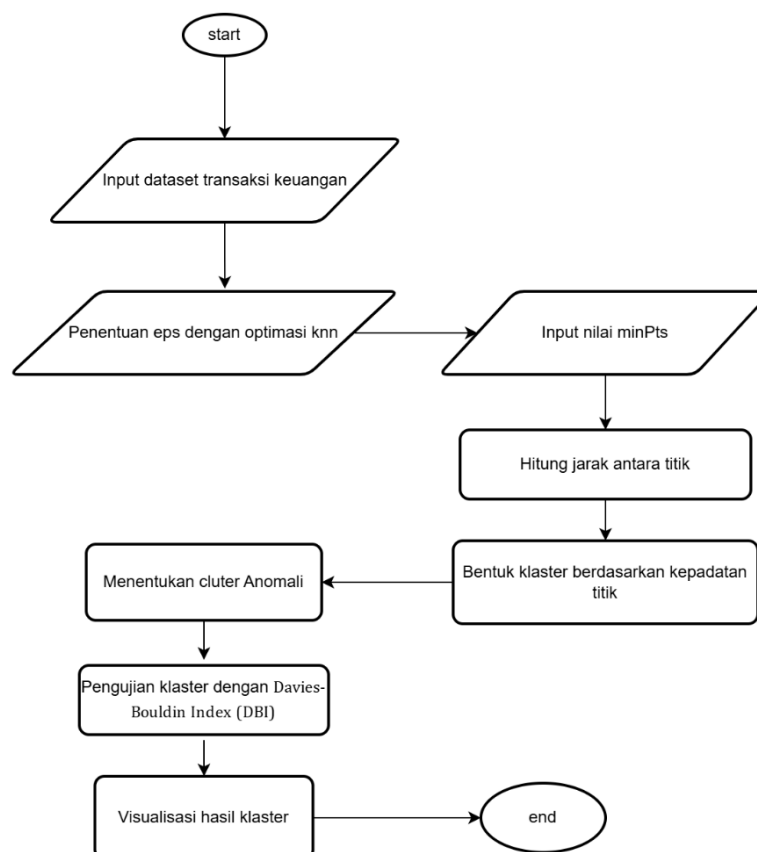
$$E(1,10) = \sqrt{(17.21 - 225.10)^2 + (914.16 - 1547.71)^2 + (2 - 1)^2} = 667.02$$

Sebagai contoh kita menerapkan $\epsilon = 6600$ dan $\text{minPts} = 4$ maka *core point* dan noise yang didapat berupa:

- **Klaster 1:** terdapat di titik {1,2,4,5,6,9,10} dimana transaksi dan jumlah saldo berada dalam rentang normal.
- **Anomali(noise):** terdapat di titik {3,7,8} Dimana transaksi dan jumlah saldo berada dalam rentang yang tidak normal.

Diagram Alur Algoritma DBSCAN

Berikut ini merupakan diagram alur dari algoritma DBSCAN dalam menentukan jumlah kluster



Gambar 2. Flowchart Algoritma DBSCAN

HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini didapat dari Kaggle. Berdasarkan dari contoh dataset di tabel 2, terdapat 3 variabel utama (x_1) jumlah transaksi, (x_2) saldo akhir, (x_3) frekuensi transaksi. Secara keseluruhan dataset memiliki 2512 baris data. Berikut adalah analisis deskriptif:

Tabel 3. Analisa deskriptif transaksi keuangan

variabel	Maximum	minimum	Rata-rata	Simpangan baku
Jumlah_Transaksi (x_1)	1919.110 \$	0.260 \$	297.593	291.946
Saldo_Akhir(x_2)	14977.990 \$	101.250 \$	5114.302	3900.942
Frekuensi_transaksi(x_3)	5	1	1.124	0.602

Normalisasi Data

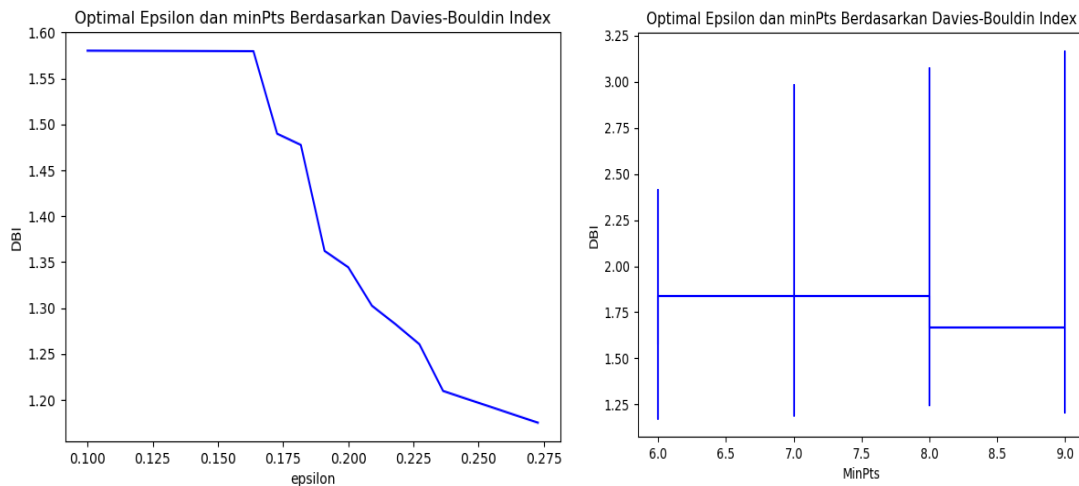
Pada prakteknya data-data akan dilakukan normalisasi dikarenakan adanya atribut yang memiliki perbedaan skala yang jauh. Perbedaan data yang sangat jauh akan sangat mempengaruhi hasil perhitungan jarak antara titik. Hasil perhitungan dari normalisasi menggunakan z-score dapat dilihat pada tabel 4.

Tabel 4. Hasil normalisasi data transaksi

No	Jumlah_Transaksi (\$)	frekuensi_Transaksi	Saldo_Akhir (\$)
1	-0.971275	1	-0.000537
2	0.269440	1	2.216472
3	-0.586882	1	-1.023534
⋮	⋮	⋮	⋮
2511	-0.382420	1	-0.855705
2512	-0.186763	1	-1.277652

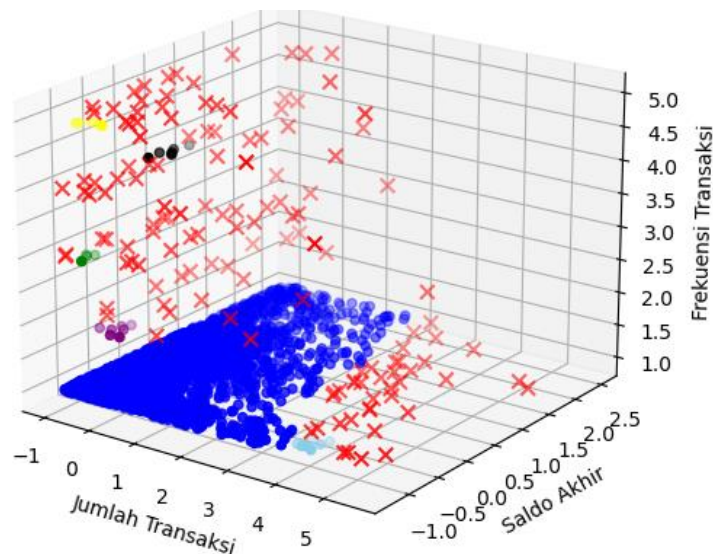
Algoritma DBSCAN

Dalam penentuan nilai epsilon(ϵ) dan minPts peneliti menggunakan Davies-Bouldin Index (DBI) sebagai evaluasi satu persatu nilai. Hasilnya dapat dilihat pada gambar 2 dan 3:



Gambar 2. Nilai epsilon(ϵ) dan minPts

Seperti yang terlihat di grafik Setelah melakukan operasi evaluasi Davies-Bouldin Index (DBI) didapatkan nilai epsilon(ϵ) = 0.2727 dan minPts=6 dikarenakan memiliki nilai DBI yang paling kecil, setelah diketahui nilai epsilon(ϵ) dan minPts adapun visualisasi penerapan algoritma DBSCAN pada gambar 3.



Gambar 3. Visualisasi hasil DBSCAN

Tanda silang merah pada visualisasi diatas menandakan *noise* sedangkan titik warna biru, hitam, ungu, biru muda, hijau dan kuning manandakan data yang terklaster. Hasil pengujian algoritma DBSCAN berdasarkan nilai epsilon(ϵ) dan minPts yang telah ditentukan sebelumnya, didapatkan 6 klaster dan 1 *noise*.

Evaluasi hasil

Hasil evaluasi klastering bertujuan untuk mengetahui kualitas klaster. Hasil evaluasi menggunakan Davies-Bouldin Index (DBI) didapatkan nilai sebesar 1.1753. Untuk hasil pengelompokan yang lebih detail dapat dilihat pada tabel 5.

Tabel 5. Hasil pengelompokkan DBSCAN

Cluster	Id_transaksi
0 (noise) terdiri dari 138 data	'TX001619', 'TX001888', 'TX001610', 'TX002263', 'TX000033', 'TX000177', 'TX002262', 'TX000395', 'TX001122', 'TX001789', 'TX000548', 'TX001439', 'TX000533', 'TX000477', 'TX000805', 'TX001532', 'TX000039', 'TX001048', 'TX001275', 'TX001985', 'TX001515', 'TX002150', 'TX000415', 'TX001310',...
Cluster 1 terdiri dari 2344 data	'TX000881', 'TX002255', 'TX001633', 'TX000521', 'TX000619', 'TX001355', 'TX001267', 'TX002260', 'TX001184', 'TX000025', 'TX001642', 'TX001206', 'TX001423', 'TX000192', 'TX000264', 'TX001445', 'TX002012', 'TX000908', 'TX001051', 'TX000205', 'TX00240,
Cluster 2 terdiri dari 8 data	'TX001354', 'TX002202', 'TX000476', 'TX001282', 'TX002273', 'TX001918', 'TX001953', 'TX002404'
Cluster 3 terdiri dari 5 data	'TX001043', 'TX001284', 'TX000027', 'TX000509', 'TX000686'
Cluster 4 terdiri dari 4 data	'TX001083', 'TX001666', 'TX002373', 'TX000594'
Cluster 5 terdiri dari 6 data	'TX000670', 'TX001058', 'TX000267', 'TX000552', 'TX002054', 'TX001689'
Cluster 6 terdiri dari 7 data	'TX002493', 'TX001469', 'TX001507', 'TX001169', 'TX000565', 'TX001300', 'TX002272'

Untuk *noise* terdapat perbedaan dengan klaster 1 sampai 6, klaster 1 berwarna biru, klaster 2 biru muda, klaster 3 kuning, klaster 4 hijau, klaster 5 hitam, dan klaster 6 ungu. Perbedaannya dapat terlihat Ketika melakukan operasi rata-rata pada setiap variabel didalam *cluster* terdiri dari 3 variabel utama (x_1) jumlah transaksi, (x_2) saldo akhir, (x_3) frekuensi transaksi.. Berikut nilai disajikan dalam tabel 6.

Tabel 6. Rata-rata variabel setiap *cluster*

Noice	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
-------	-----------	-----------	-----------	-----------	-----------	-----------



X ₁	599.709	276.975	1455.367	168.814	120.877	91.658	292.178
X ₂	2.884	1	1	5	3	4	2
X ₃	6610.947	5065.454	1366.428	597.086	804.795	6321.751	903.751

Seperti yang terlihat didalam tabel, meskipun *noise* hanya memiliki 138 data dibandingkan dengan klaster 1 yang memiliki 2344 data, nilai rata-rata ketiga variabel *noise* lebih besar dari pada yang klaster 1. Untuk klaster lainnya dikarenakan jarak data terlalu rapat sehingga data terhitung sebagai klaster.

Berdasarkan penerapan algoritma DBSCAN pada data transaksi keuangan dengan parameter optimal epsilon (ϵ) sebesar 0.2727 dan minPts sebanyak 6, diperoleh 6 cluster utama serta 1 kelompok noise (anomali). Visualisasi clustering menunjukkan distribusi data yang cukup variatif, di mana mayoritas data tergabung dalam Cluster 1 (2344 data), sedangkan cluster lainnya berisi data dengan jumlah yang lebih kecil, yakni Cluster 2 (8 data), Cluster 3 (5 data), Cluster 4 (4 data), Cluster 5 (6 data), dan Cluster 6 (7 data). Sementara itu, noise terdeteksi sebanyak 138 data yang menunjukkan karakteristik signifikan berbeda dari data mayoritas.

Analisis rata-rata variabel dalam setiap cluster (Jumlah Transaksi, Saldo Akhir, dan Frekuensi Transaksi) mengungkapkan pola yang menarik:

- **Cluster 1 (Normal Dominan):** Memiliki rerata jumlah transaksi dan saldo akhir yang relatif moderat, serta frekuensi transaksi rendah (rata-rata 1 kali), mencerminkan transaksi rutin dari mayoritas pengguna e-money.
- **Cluster 2 & Cluster 3 (Anomali Ringan):** Cluster kecil dengan rerata jumlah transaksi yang tinggi dan saldo akhir yang lebih besar dari Cluster 1. Hal ini mengindikasikan adanya pengguna dengan perilaku transaksi besar namun masih dalam batas wajar.
- **Cluster 4, 5, dan 6 (Kelompok Khusus):** Berisi data dengan frekuensi transaksi lebih tinggi dan variasi saldo akhir yang fluktuatif, mencerminkan pola transaksi pengguna bisnis atau akun dengan aktivitas spesifik.
- **Noise (Anomali Kuat):** Ditemukan bahwa data dalam kelompok noise memiliki rerata jumlah transaksi dan saldo akhir yang jauh lebih tinggi dibandingkan cluster lain, dengan rerata saldo akhir mencapai 6610.947. Hal ini menjadi indikasi kuat adanya transaksi mencurigakan yang patut diwaspadai sebagai potensi fraud atau aktivitas tidak wajar.

Nilai Davies-Bouldin Index (DBI) sebesar 1.1753 menunjukkan bahwa pemisahan antar cluster cukup optimal, meskipun masih terdapat potensi overlap pada beberapa cluster kecil. DBI < 2 umumnya dianggap cukup baik dalam konteks clustering berbasis densitas untuk data transaksi yang memiliki distribusi variatif.

Hasil penelitian ini sejalan dengan temuan Boer et al. (2023) yang menunjukkan efektivitas DBSCAN dalam mengelompokkan data spasial dengan distribusi densitas yang tidak seragam. Demikian pula, studi Prihatmono (2025) membuktikan bahwa DBSCAN mampu memisahkan cluster mayoritas dari outlier tanpa memerlukan asumsi jumlah cluster di awal.

Namun, dibandingkan dengan penelitian Polgan et al. (2024) yang menggunakan SVM, keunggulan DBSCAN terletak pada fleksibilitas dalam mendeteksi anomali dengan distribusi kompleks tanpa supervised learning. Ini menjadi kontribusi teoretis penting bahwa DBSCAN lebih adaptif dalam konteks transaksi e-money yang fluktuatif.

SIMPULAN

Penelitian ini menunjukkan bahwa algoritma DBSCAN efektif dalam mendeteksi anomali pada data transaksi keuangan berbasis e-money dengan hasil clustering yang mampu membedakan data normal dan mencurigakan tanpa memerlukan asumsi awal jumlah cluster. Dengan parameter optimal epsilon sebesar 0.2727 dan minPts sebanyak 6, DBSCAN berhasil mengelompokkan 2.512 data menjadi enam cluster utama dan satu kelompok noise berisi 138 data yang memiliki karakteristik signifikan berbeda, terutama pada variabel jumlah transaksi, saldo akhir, dan frekuensi transaksi. Hasil evaluasi menggunakan Davies-Bouldin Index (DBI) sebesar 1.1753 mendukung kualitas pemisahan cluster yang memadai, membuktikan bahwa DBSCAN mampu mengidentifikasi pola transaksi tidak wajar secara akurat. Secara teoretis, penelitian ini memperkaya kajian tentang deteksi anomali berbasis clustering densitas, sedangkan secara praktis memberikan kontribusi bagi lembaga keuangan dan regulator dalam memantau aktivitas transaksi digital guna mencegah potensi penyimpangan finansial.

DAFTAR PUSTAKA

- [1] N. R. Prasad, S. Almanza-Garcia, and T. T. Lu, "Anomaly detection," *Computers, Materials and Continua*, vol. 14, no. 1, pp. 1–22, Jul. 2009, doi: 10.1145/1541880.1541882;SERIALTOPIC:TOPIC:ACM-PUBTYPE>JOURNAL;PAGE:STRING:ARTICLE/CHAPTER.
- [2] A. Sharma, A. Kumar, and A. K. Vaish, "Market anomalies and investor behaviour," *inderscienceonline.com*, vol. 12, no. 1, pp. 62–81, 2022, doi: 10.1504/AJFA.2022.121768.

 <http://journal.mahesacenter.org/index.php/incoding>

 mahesainstitut@gmail.com 132



- [3] M. Ester, H. Kriegel, J. Sander, X. X.- kdd, and undefined 1996, "A density-based algorithm for discovering clusters in large spatial databases with noise," *cdn.aaai.orgM Ester, HP Kriegel, J Sander, X Xukdd, 1996•cdn.aaai.org, 1996, Accessed: May 16, 2025. [Online]. Available: https://cdn.aaai.org/KDD/1996/KDD96-037.pdf?source=post_page-----*
- [4] C. Schubert, "Green nudges: Do they work? Are they ethical?," *Ecological Economics*, vol. 132, pp. 329–342, Feb. 2017, doi: 10.1016/J.ECOLECON.2016.11.009.
- [5] K. M. Boer, "Iklan Partai Politik Dan Politik Media," *Jurnal Visi Komunikasi*, vol. 13, no. 02, pp. 296–310, 2014, [Online]. Available: <https://publikasi.mercubuana.ac.id/index.php/viskom/article/view/1660>
- [6] A. R. Prihatmono, "Pilkada Calon Tunggal Di Kabupaten Pati Tahun 2017: Suatu Tinjauan Oligarkisme Parpol," 2019.
- [7] C. S. Otiva, P. E. Haes, T. I. Fajri, H. Eldo, and M. L. Hakim, "Implementasi Teknologi Informasi pada UMKM: Tantangan dan Peluang," *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 815–821, 2024.
- [8] N. A. Permatasari, Y. H. Chrisnanto, and A. K. Ningsih, "Segmentasi Kasus Data Kematian Covid 19 Di Jawa Barat Menggunakan Algoritma DBSCAN," *IJESPG (International Journal of Engineering, Economic, Social Politic and Government)*, vol. 1, no. 4, pp. 119–128, 2023.
- [9] I. N. Simbolon and P. D. Friskila, "ANALISIS DAN EVALUASI ALGORITMA DBSCAN SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE) PADA TUBERKULOSIS," vol. 12, no. 3, 2024.
- [10] M. T. Anwar, W. Hadikurniawati, E. Winarno, and A. Supriyanto, "Wildfire Risk Map Based on DBSCAN Clustering and Cluster Density Evaluation," *Advance Sustainable Science, Engineering and Technology*, vol. 1, no. 1, pp. 1–7, 2019, doi: 10.26877/asset.v1i1.4876.
- [11] H. Lu, "Evaluating the Performance of SVM , Isolation Forest , and DBSCAN for Anomaly Detection," vol. 04012, pp. 1–6, 2025.
- [12] M. S. Al-Batah, E. R. Al-Kwaldeh, M. A. Wahed, M. Alzyoud, and N. Al-Shanableh, "Enhancement over DBSCAN Satellite Spatial Data Clustering," *Journal of Electrical and Computer Engineering*, vol. 2024, 2024, doi: 10.1155/2024/2330624.
- [13] I. B. F. Arafat, M. A. Hariyadi, I. B. Santoso, and C. Crysdian, "Clustering Gempabumi di Wilayah Regional VII Menggunakan Pendekatan DBSCAN," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 4, pp. 823–830, 2023, doi: 10.25126/jtiik.20241046918.
- [14] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "Why and How You Should (Still) Use DBSCAN," *ACM Transactions on Database Systems*, vol. 42, no. 3, pp. 1–21, 2017.
- [15] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "Why and How You Should (Still) Use DBSCAN," *ACM Transactions on Database Systems*, vol. 42, no. 3, pp. 1–21, 2017.
- [16] J. Xiao, J. Lu, and X. Li, "Davies Bouldin Index based hierarchical initialization K-means," *Intelligent Data Analysis*, vol. 21, no. 6, pp. 1327–1338, 2017, doi: 10.3233/IDA-163129.

