

Clustering Pola Penggunaan Energi pada Smart Home Menggunakan DBSCAN Berbasis Data Time Series Sensor

Arif Susilo*, Asep Arwan Sulaeman, Nur Suci Rahayu

Fakultas Teknik, Program Studi Teknik Informatika, Universitas Pelita Bangsa, Bekasi
Ruko Bekasi Mas, Jl. Ahmad Yani, Marga Jaya, Kec. Bekasi Sel., Kota Bks, Jawa Barat, Indonesia
Email: ^{1,*}arif.susilo@pelitabangsa.ac.id, ²asep.arwan@pelitabangsa.ac.id, ³n.cahya@gmail.com

Email Penulis Korespondensi: arif.susilo@pelitabangsa.ac.id

Submitted: 29/06/2026; Accepted: 23/07/2026; Published: 24/07/2026

Abstrak—Peningkatan penggunaan perangkat Internet of Things (IoT) pada lingkungan smart home menghasilkan data konsumsi energi yang bersifat kontinu dan kompleks, sehingga diperlukan metode yang mampu mengidentifikasi pola penggunaan energi secara efektif. Penelitian ini bertujuan mengelompokkan pola penggunaan energi rumah tangga menggunakan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) berbasis data time series sensor. Dataset yang digunakan merupakan Smart Home Energy Consumption Dataset yang terdiri atas sekitar 90.000 observasi dengan enam variabel utama, yaitu Energy Consumption, Peak Hours Usage, Household Size, Average Temperature, Has AC, dan Weekday. Tahapan penelitian meliputi seleksi atribut, pembersihan data, normalisasi menggunakan StandardScaler, penentuan parameter DBSCAN menggunakan K-Distance Graph, proses clustering, serta evaluasi menggunakan Silhouette Score. Hasil eksperimen menunjukkan bahwa parameter terbaik diperoleh pada $\epsilon = 0,38$ dan MinPts = 5, yang menghasilkan 89 cluster, 541 data noise (0,60%), dan Silhouette Score sebesar 0,0690. Analisis karakteristik cluster menunjukkan bahwa konsumsi energi, penggunaan energi pada jam puncak, kepemilikan pendingin ruangan, jumlah penghuni, dan suhu lingkungan merupakan faktor yang membedakan pola penggunaan energi antar kelompok rumah tangga. Hasil penelitian menunjukkan bahwa DBSCAN mampu mengidentifikasi pola penggunaan energi dan mendeteksi data penciliran tanpa memerlukan penentuan jumlah cluster di awal proses analisis, sehingga berpotensi mendukung pengembangan sistem manajemen energi cerdas pada lingkungan smart home.

Kata Kunci: DBSCAN; Smart Home; Clustering; Konsumsi Energi; Time Series; Internet of Things

Abstract—The increasing adoption of Internet of Things (IoT) devices in smart homes has generated continuous and complex energy consumption data, requiring effective clustering techniques to identify household energy usage patterns. This study aims to cluster household energy consumption patterns using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm based on sensor time-series data. The study utilized the Smart Home Energy Consumption Dataset, consisting of approximately 90,000 observations with six main variables: Energy Consumption, Peak Hours Usage, Household Size, Average Temperature, Has AC, and Weekday. The research workflow included feature selection, data cleaning, data normalization using StandardScaler, parameter determination through the K-Distance Graph, DBSCAN clustering, and clustering evaluation using the Silhouette Score. Experimental results indicated that the optimal parameters were $\epsilon = 0.38$ and MinPts = 5, producing 89 clusters, 541 noise observations (0.60%), and a Silhouette Score of 0.0690. Cluster characteristic analysis revealed that energy consumption, peak-hour energy usage, air conditioner ownership, household size, and ambient temperature were the primary factors distinguishing household energy usage patterns. The findings demonstrate that DBSCAN effectively identifies household energy consumption patterns while detecting outliers without requiring the number of clusters to be predefined, making it a promising approach for supporting intelligent energy management systems in smart home environments.

Keywords: DBSCAN; Smart Home; Energy Consumption; Clustering; Time Series; Internet of Things

1. PENDAHULUAN

Perkembangan Internet of Things (IoT) telah mendorong implementasi smart home sebagai solusi untuk meningkatkan efisiensi penggunaan energi, kenyamanan penghuni, serta otomatisasi berbagai perangkat rumah tangga [1], [2]. Smart home memanfaatkan berbagai sensor, seperti sensor daya listrik, suhu, kelembaban, cahaya, dan pergerakan, yang menghasilkan data secara kontinu dalam bentuk time series [3], [4]. Data tersebut menggambarkan perilaku konsumsi energi dari berbagai perangkat rumah tangga pada interval waktu tertentu sehingga dapat dimanfaatkan untuk memahami pola penggunaan energi secara lebih mendalam [5], [6]. Meningkatnya jumlah perangkat IoT menyebabkan volume data sensor yang dihasilkan menjadi sangat besar dan kompleks [7], [8]. Karakteristik data time series yang bersifat dinamis, memiliki fluktuasi temporal, serta mengandung noise menjadikan analisis secara manual sulit dilakukan [9], [10]. Oleh karena itu, diperlukan teknik data mining yang mampu mengelompokkan pola penggunaan energi berdasarkan karakteristik konsumsi tanpa memerlukan label data [11], [12]. Hasil pengelompokan tersebut dapat digunakan sebagai dasar dalam pengembangan sistem manajemen energi, deteksi anomali konsumsi listrik, optimasi penjadwalan perangkat elektronik, serta penyusunan strategi penghematan energi pada lingkungan smart home [13], [14].

Salah satu metode clustering yang banyak digunakan adalah K-Means [15], [16], namun metode ini memiliki beberapa keterbatasan, antara lain mengharuskan jumlah cluster ditentukan di awal, sensitif terhadap inisialisasi centroid, dan kurang mampu menangani data yang mengandung noise maupun cluster dengan bentuk tidak beraturan [17], [18]. Permasalahan tersebut sering dijumpai pada data sensor smart home karena pola konsumsi energi setiap penghuni bersifat dinamis dan tidak selalu membentuk distribusi berbentuk bola (spherical

cluster) [19], [20]. Akibatnya, penggunaan metode berbasis centroid kurang optimal dalam merepresentasikan pola konsumsi energi yang sebenarnya.

Sebagai alternatif, algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) menawarkan pendekatan berbasis kepadatan (density-based clustering) yang mampu menemukan cluster dengan bentuk arbitrer tanpa harus menentukan jumlah cluster terlebih dahulu [21], [22]. Selain itu, DBSCAN memiliki kemampuan mengidentifikasi data pencilan (outlier) yang sering muncul akibat kesalahan pembacaan sensor, gangguan komunikasi IoT, maupun aktivitas penggunaan energi yang tidak lazim [23], [24], [25]. Karakteristik tersebut menjadikan DBSCAN lebih sesuai untuk menganalisis data time series sensor pada lingkungan smart home dibandingkan metode clustering konvensional [26].

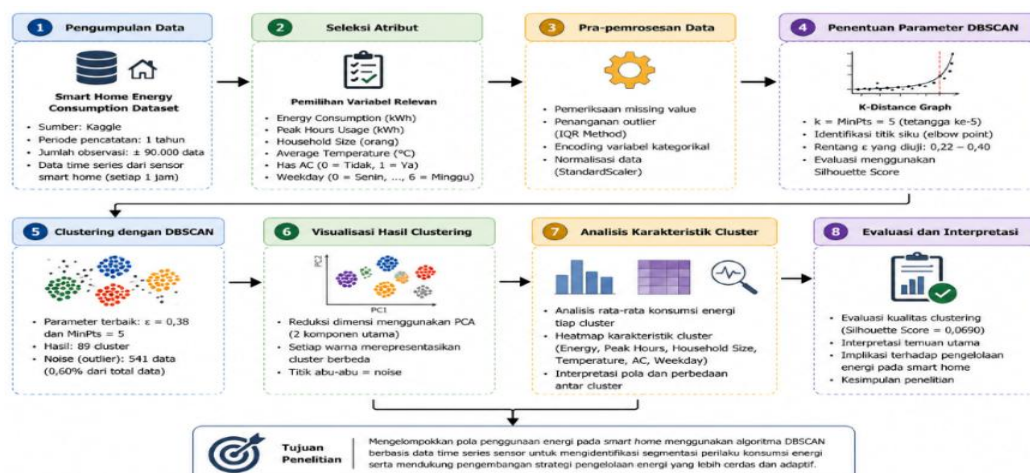
Meskipun berbagai penelitian telah membahas prediksi konsumsi energi menggunakan metode machine learning maupun deep learning, sebagian besar penelitian berfokus pada forecasting konsumsi energi. Penelitian mengenai identifikasi pola perilaku penggunaan energi melalui pendekatan clustering berbasis data time series sensor masih relatif terbatas, khususnya dengan memanfaatkan DBSCAN untuk menemukan pola konsumsi tanpa asumsi jumlah cluster. Padahal, informasi mengenai pola konsumsi energi memiliki nilai penting dalam memahami kebiasaan penghuni rumah, mengidentifikasi periode beban listrik tinggi, serta mendukung implementasi sistem Home Energy Management System (HEMS) yang adaptif.

Berbagai penelitian terkini telah memanfaatkan teknik machine learning dan deep learning untuk mendukung analisis konsumsi energi pada lingkungan smart home, terutama untuk prediksi beban listrik, deteksi anomali, dan optimasi penggunaan energi. Sebagian besar penelitian menggunakan pendekatan supervised learning seperti Random Forest, Support Vector Machine, Long Short-Term Memory (LSTM), maupun Gated Recurrent Unit (GRU), dengan fokus utama pada forecasting konsumsi energi. Di sisi lain, penelitian yang memanfaatkan pendekatan unsupervised learning untuk mengeksplorasi pola perilaku konsumsi energi masih relatif terbatas. Beberapa penelitian telah menerapkan algoritma K-Means atau Hierarchical Clustering, namun metode tersebut masih memiliki keterbatasan dalam menangani data berdensitas tidak seragam, cluster berbentuk arbitrer, dan keberadaan data pencilan. Oleh karena itu, penerapan DBSCAN pada data time series sensor smart home menjadi alternatif yang menjanjikan karena mampu membentuk cluster berdasarkan kepadatan data tanpa memerlukan penentuan jumlah cluster di awal serta secara alami mengidentifikasi observasi yang bersifat noise. Kondisi tersebut menunjukkan bahwa masih terdapat ruang penelitian untuk mengembangkan analisis pola konsumsi energi berbasis density-based clustering pada data IoT smart home.

Berdasarkan permasalahan dan kesenjangan penelitian tersebut, tujuan penelitian ini adalah mengidentifikasi pola penggunaan energi rumah tangga pada lingkungan smart home menggunakan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) berbasis data time series sensor. Penelitian ini bertujuan memperoleh konfigurasi parameter DBSCAN yang optimal menggunakan K-Distance Graph, menghasilkan cluster yang mampu merepresentasikan karakteristik konsumsi energi rumah tangga, serta mengevaluasi kualitas clustering menggunakan Silhouette Score. Hasil penelitian diharapkan memberikan informasi yang lebih komprehensif mengenai karakteristik perilaku konsumsi energi sehingga dapat mendukung pengembangan sistem manajemen energi cerdas pada lingkungan smart home.

2. METODOLOGI PENELITIAN

Penelitian ini bertujuan mengidentifikasi pola penggunaan energi pada lingkungan smart home menggunakan algoritma DBSCAN dengan pendekatan **Knowledge Discovery in Databases (KDD)** yang meliputi pengumpulan data, preprocessing, transformasi data, clustering, evaluasi, dan interpretasi hasil. Tahapan penelitian disajikan pada **Gambar 1**.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Dataset

Penelitian ini menggunakan Household Energy Consumption Dataset yang diperoleh dari Kaggle dengan sekitar 90.000 observasi data konsumsi energi rumah tangga berbasis time series. Variabel yang digunakan meliputi Date, Energy_Consumption_kWh, Household_Size, Avg_Temperature_C, Peak_Hours_Usage_kWh, dan Has_AC. Seluruh variabel digunakan untuk mengidentifikasi pola penggunaan energi rumah tangga menggunakan algoritma DBSCAN.

Tabel 1. Variabel Penelitian

Variabel	Tipe Data	Keterangan
Household_ID	String	Primary identifier
Date	Date	Tanggal pencatatan data
Energy_Consumption_kWh	Numeric	Konsumsi energi (kWh)
Household_Size	Numeric	Jumlah penghuni rumah
Avg_Temperature_C	Numeric	Suhu rata-rata (°C)
Has_AC	Kategorikal	Kepemilikan pendingin ruangan
Peak_Hours_Usage_kWh	Numeric	Konsumsi energi pada jam puncak (kWh)

2.2 Seleksi Atribut

Berdasarkan proses seleksi atribut diperoleh enam variabel yang digunakan sebagai input algoritma DBSCAN, yaitu Energy_Consumption_kWh, Peak_Hours_Usage_kWh, Household_Size, Avg_Temperature_C, Has_AC, dan Weekday.

Tabel 2. Hasil Seleksi Atribut

Atribut	Status	Alasan
Household_ID	Dihapus	Hanya sebagai identitas data
Date	Ditransformasi	Digunakan untuk memperoleh atribut Weekday
Hour	Dihapus	Seluruh data bernilai 0 sehingga tidak memiliki variasi
Day	Dihapus	Tidak merepresentasikan pola konsumsi energi
Month	Dihapus	Seluruh data berada pada bulan yang sama
Energy_Consumption_kWh	Dipertahankan	Variabel utama konsumsi energi
Peak_Hours_Usage_kWh	Dipertahankan	Merepresentasikan konsumsi pada jam puncak
Household_Size	Dipertahankan	Faktor karakteristik rumah tangga
Avg_Temperature_C	Dipertahankan	Faktor lingkungan
Has_AC	Dipertahankan	Setelah Label Encoding
Weekday	Dipertahankan	Merepresentasikan pola mingguan

2.3 Data Preprocessing

Tahap preprocessing bertujuan meningkatkan kualitas data sebelum proses clustering dilakukan. Langkah pertama adalah melakukan pemeriksaan terhadap **missing value** pada setiap atribut. Apabila ditemukan nilai yang hilang, proses imputasi dilakukan menggunakan nilai median agar distribusi data tetap terjaga.

Tabel 3. Tahapan Data Preprocessing

Tahapan	Hasil
Pemeriksaan missing value	Tidak ditemukan missing value yang memengaruhi analisis
Pemeriksaan data duplikat	Data diperiksa sebelum proses clustering
Label Encoding	Variabel Has_AC diubah menjadi numerik (0=No, 1=Yes)
Standardisasi	Menggunakan StandardScaler
Ekstraksi fitur waktu	Date → Weekday

Selanjutnya dilakukan pemeriksaan data duplikat untuk memastikan tidak terdapat observasi yang tercatat lebih dari satu kali. Setelah itu, seluruh atribut numerik dinormalisasi menggunakan metode **StandardScaler** sehingga setiap variabel memiliki skala yang sebanding. Proses normalisasi penting dilakukan karena algoritma DBSCAN menggunakan perhitungan jarak antar data dalam proses pembentukan cluster. Normalisasi dilakukan menggunakan Persamaan (1).

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

Dalam perhitungan statistik, x merupakan nilai data yang diamati, μ menunjukkan nilai rata-rata dari keseluruhan data, sedangkan σ merepresentasikan standar deviasi yang menggambarkan tingkat penyebaran data terhadap nilai rata-ratanya. Ketiga komponen tersebut digunakan untuk menganalisis posisi dan karakteristik suatu nilai data dalam distribusi statistik tertentu.

2.4 Transformasi Data Time Series

Atribut Date terlebih dahulu dikonversi ke dalam format datetime untuk memudahkan proses ekstraksi informasi temporal. Selanjutnya, dari atribut tersebut diekstraksi fitur Weekday yang merepresentasikan hari dalam satu minggu (Senin hingga Minggu). Pemilihan atribut Weekday bertujuan untuk menangkap pola konsumsi energi yang dipengaruhi oleh aktivitas harian penghuni, seperti perbedaan penggunaan energi pada hari kerja dan akhir pekan. Sebaliknya, atribut Hour, Day, dan Month tidak digunakan dalam proses clustering. Atribut Hour dihapus karena seluruh observasi memiliki nilai yang sama sehingga tidak memberikan variasi informasi. Atribut Day tidak digunakan karena hanya merepresentasikan urutan tanggal tanpa mencerminkan karakteristik pola konsumsi energi, sedangkan atribut Month dihilangkan karena seluruh data berasal dari bulan yang sama sehingga tidak memiliki kemampuan diskriminatif dalam proses pengelompokan.

Setelah proses transformasi selesai, seluruh atribut yang digunakan sebagai masukan algoritma DBSCAN terdiri atas `Energy_Consumption_kWh`, `Peak_Hours_Usage_kWh`, `Household_Size`, `Avg_Temperature_C`, `Has_AC`, dan `Weekday`. Keenam atribut tersebut dipilih karena mampu merepresentasikan karakteristik utama konsumsi energi rumah tangga, meliputi besarnya konsumsi energi, intensitas penggunaan energi pada jam puncak, karakteristik penghuni, kondisi lingkungan, kepemilikan perangkat pendingin ruangan, serta pola penggunaan berdasarkan hari pengamatan. Seluruh atribut numerik kemudian dinormalisasi menggunakan `StandardScaler` sehingga memiliki skala yang sebanding sebelum digunakan pada proses clustering dengan algoritma DBSCAN.

2.5 Clustering Menggunakan DBSCAN

Proses clustering dilakukan menggunakan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) yang mengelompokkan data berdasarkan tingkat kepadatan (density), sehingga mampu membentuk cluster dengan bentuk yang tidak beraturan (arbitrary shape) serta mendeteksi data penciran (noise). Algoritma ini menggunakan dua parameter utama, yaitu epsilon (ϵ) sebagai radius pencarian tetangga dan `MinPts` sebagai jumlah minimum titik dalam radius ϵ untuk membentuk core point. Penentuan kedua parameter tersebut sangat memengaruhi kualitas hasil clustering. Pada penelitian ini, jarak antar data dihitung menggunakan Euclidean Distance sebagaimana ditunjukkan pada Persamaan (2).

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (2)$$

Dalam perhitungan jarak antar data, x_i merepresentasikan nilai data ke- i , sedangkan x_j menunjukkan nilai data ke- j . Kedua nilai tersebut digunakan untuk membandingkan karakteristik atau perbedaan antarobjek berdasarkan atribut yang dianalisis. Sementara itu, n menyatakan jumlah atribut yang digunakan dalam proses perhitungan dan analisis data secara keseluruhan. Nilai parameter ϵ ditentukan menggunakan K-Distance Graph, sedangkan beberapa kombinasi nilai `MinPts` diuji untuk memperoleh konfigurasi parameter yang menghasilkan cluster terbaik.

2.6 Evaluasi Hasil Clustering

Kualitas hasil clustering dievaluasi menggunakan beberapa indikator, yaitu jumlah cluster yang terbentuk, jumlah observasi yang teridentifikasi sebagai noise (outlier), dan nilai Silhouette Score. Jumlah cluster digunakan untuk mengetahui struktur pengelompokan yang dihasilkan oleh algoritma DBSCAN, sedangkan jumlah noise menunjukkan kemampuan algoritma dalam mengidentifikasi observasi yang tidak memenuhi kriteria kepadatan sehingga tidak dimasukkan ke dalam cluster tertentu. Selanjutnya, Silhouette Score digunakan sebagai metrik evaluasi internal untuk mengukur kualitas hasil clustering berdasarkan tingkat kohesi (cohesion) antar anggota dalam cluster yang sama dan tingkat separasi (separation) antar cluster yang berbeda. Nilai Silhouette Score berada pada rentang -1 hingga 1 , di mana nilai yang semakin mendekati 1 menunjukkan kualitas cluster yang semakin baik, sedangkan nilai yang mendekati 0 menunjukkan adanya tumpang tindih antar cluster.

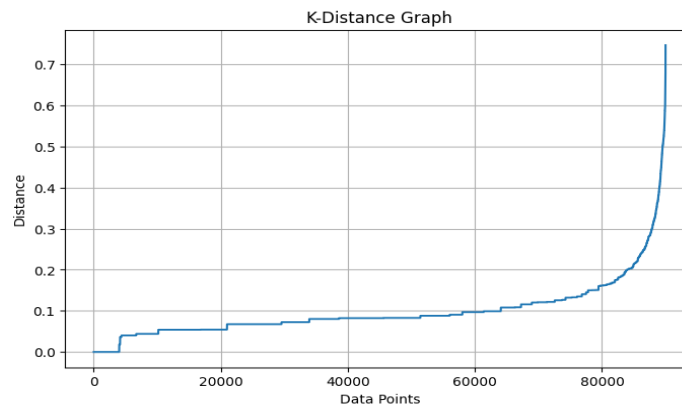
Sebelum proses clustering dilakukan, parameter epsilon (ϵ) ditentukan menggunakan K-Distance Graph. Grafik ini dibangun dengan menghitung jarak setiap observasi terhadap tetangga ke- k (k -nearest neighbors), kemudian mengurutkan jarak tersebut secara menaik. Titik perubahan kurva yang membentuk elbow point digunakan sebagai acuan dalam menentukan nilai ϵ yang optimal karena merepresentasikan batas perubahan kepadatan data. Selanjutnya, beberapa kombinasi nilai ϵ dan `MinPts` diuji untuk memperoleh konfigurasi parameter yang menghasilkan kualitas clustering terbaik berdasarkan nilai Silhouette Score, jumlah cluster yang terbentuk, dan jumlah data noise.

Hasil clustering kemudian divisualisasikan menggunakan Principal Component Analysis (PCA). Metode PCA digunakan untuk mereduksi dimensi data multivariat menjadi dua komponen utama (principal components) yang mampu mempertahankan sebagian besar variasi informasi pada dataset. Proses reduksi dimensi ini bertujuan mempermudah interpretasi hasil clustering tanpa mengubah struktur utama hubungan antar data. Visualisasi PCA juga digunakan untuk memverifikasi apakah pola pengelompokan yang dihasilkan oleh algoritma DBSCAN telah sesuai dengan karakteristik kepadatan data, sehingga dapat mendukung proses interpretasi hasil dan analisis karakteristik masing-masing cluster.

3. HASIL DAN PEMBAHASAN

3.1 Penentuan Parameter DBSCAN

Penentuan parameter merupakan tahapan penting dalam penerapan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) karena kualitas hasil clustering sangat dipengaruhi oleh nilai parameter epsilon (ϵ) dan minimum samples (MinPts). Parameter ϵ menentukan radius pencarian tetangga (neighborhood radius), sedangkan MinPts menentukan jumlah minimum titik yang harus berada dalam radius tersebut agar suatu data dapat dikategorikan sebagai core point. Pemilihan parameter yang kurang tepat dapat menyebabkan terbentuknya terlalu banyak cluster, terlalu sedikit cluster, atau meningkatnya jumlah data yang dikategorikan sebagai noise. Pada penelitian ini, penentuan nilai ϵ dilakukan menggunakan K-Distance Graph dengan jumlah tetangga (k-nearest neighbors) sebanyak lima sesuai dengan nilai MinPts = 5. Grafik K-Distance diperoleh dengan menghitung jarak terhadap tetangga ke-5 dari setiap observasi, kemudian mengurutkannya secara menaik. Titik perubahan kurva (elbow point) digunakan sebagai acuan awal dalam menentukan rentang nilai ϵ yang akan diuji.



Gambar 2. Grafik K-Distance

Keterangan Gambar 2. Grafik K-Distance menunjukkan hubungan antara urutan data dan jarak ke tetangga terdekat ke-5. Kurva mulai mengalami peningkatan tajam pada kisaran nilai ϵ sekitar 0,30–0,40, sehingga rentang tersebut digunakan sebagai acuan dalam proses pengujian parameter DBSCAN. Berdasarkan hasil visualisasi pada Gambar 2, pengujian parameter dilakukan menggunakan beberapa kombinasi nilai ϵ , yaitu 0,22–0,40, dengan dua nilai MinPts, yaitu 5 dan 10. Setiap kombinasi dievaluasi berdasarkan jumlah cluster yang terbentuk serta jumlah data yang teridentifikasi sebagai noise. Hasil pengujian parameter ditunjukkan pada Tabel 4.

Tabel 4. Hasil Pengujian Parameter DBSCAN

ϵ (eps)	MinPts	Jumlah Cluster	Noise
0,22	5	158	3526
0,22	10	113	7452
0,24	5	128	2738
0,24	10	98	5716
0,26	5	125	2027
0,26	10	92	4263
0,28	5	113	1640
0,28	10	90	3410
0,30	5	104	1285
0,30	10	89	2629
0,32	5	101	1062
0,32	10	84	2164
0,34	5	98	817
0,34	10	85	1712
0,36	5	92	679
0,36	10	86	1387
0,38	5	89	541
0,38	10	85	1132
0,40	5	89	430
0,40	10	85	955

Keterangan Tabel 4. Peningkatan nilai ϵ menyebabkan jumlah cluster dan noise semakin berkurang. Pada $\epsilon = 0,22$ –0,26, DBSCAN menghasilkan lebih dari 120 cluster dengan noise yang tinggi, sedangkan pada $\epsilon = 0,36$ –

0,40 jumlah cluster menjadi lebih stabil (89–92 cluster) dengan noise kurang dari 1%. Selanjutnya, setiap kombinasi parameter dievaluasi menggunakan Silhouette Score untuk menentukan parameter terbaik..

Tabel 5. Evaluasi Parameter Menggunakan Silhouette Score

ϵ (eps)	MinPts	Jumlah Cluster	Noise	Silhouette Score
0,30	5	104	1285	0,0426
0,32	5	101	1062	0,0470
0,34	5	98	817	0,0513
0,36	5	92	679	0,0665
0,38	5	89	541	0,0690
0,40	5	89	430	0,0679

Keterangan Tabel 5. Berdasarkan Tabel 5, kombinasi $\epsilon = 0,38$ dan MinPts = 5 menghasilkan Silhouette Score tertinggi (0,0690) dengan 89 cluster dan 541 data noise. Meskipun $\epsilon = 0,40$ menghasilkan noise yang lebih sedikit, nilai Silhouette Score lebih rendah (0,0679). Oleh karena itu, konfigurasi $\epsilon = 0,38$ dan MinPts = 5 dipilih sebagai parameter terbaik dan digunakan pada proses clustering DBSCAN.

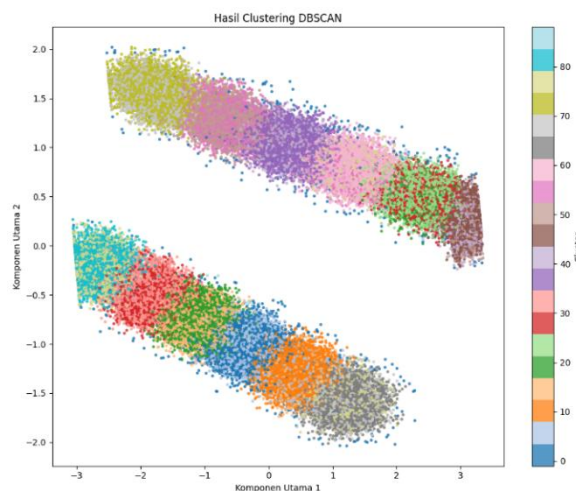
3.2 Hasil Clustering DBSCAN

Setelah diperoleh parameter terbaik $\epsilon = 0,38$ dan MinPts = 5, proses clustering menggunakan DBSCAN menghasilkan **89 cluster** dengan **541 observasi** yang teridentifikasi sebagai noise. Observasi noise merupakan data yang tidak memenuhi kriteria kepadatan minimum sehingga tidak tergabung ke dalam cluster.

Tabel 6. Ringkasan Hasil Clustering DBSCAN

Parameter	Nilai
Algoritma	DBSCAN
Epsilon (ϵ)	0,38
MinPts	5
Jumlah Cluster	89
Jumlah Noise	541
Persentase Noise	0,60 %
Silhouette Score	0,0690

Berdasarkan Tabel 6, algoritma DBSCAN menghasilkan 89 cluster dengan 541 data noise atau sekitar 0,60% dari keseluruhan dataset. Persentase noise yang rendah menunjukkan bahwa sebagian besar observasi memiliki tingkat kepadatan yang cukup sehingga dapat dikelompokkan dengan baik. Meskipun Silhouette Score yang diperoleh sebesar 0,0690 masih tergolong rendah, hasil tersebut mencerminkan karakteristik data konsumsi energi yang bersifat kontinu sehingga batas antar cluster tidak terbentuk secara tegas.



Gambar 3. Hasil Clustering DBSCAN

Berdasarkan Gambar 3, algoritma DBSCAN berhasil mengelompokkan data konsumsi energi rumah tangga ke dalam 89 cluster dengan karakteristik kepadatan yang berbeda. Sebagian besar cluster memiliki distribusi anggota yang relatif seimbang, sedangkan hanya lima cluster yang beranggotakan kurang dari 20 observasi. Hasil ini menunjukkan bahwa parameter $\epsilon = 0,38$ dan MinPts = 5 mampu menghasilkan struktur cluster yang stabil. Selain itu, terdapat 541 data noise yang merepresentasikan pola konsumsi energi yang berbeda dari mayoritas observasi. Kemampuan mengidentifikasi noise menjadi salah satu keunggulan DBSCAN dalam mendeteksi data pencilan tanpa memaksanya bergabung ke dalam cluster tertentu. Secara keseluruhan, hasil clustering

menunjukkan bahwa DBSCAN efektif mengelompokkan pola penggunaan energi rumah tangga berdasarkan tingkat kepadatan data.

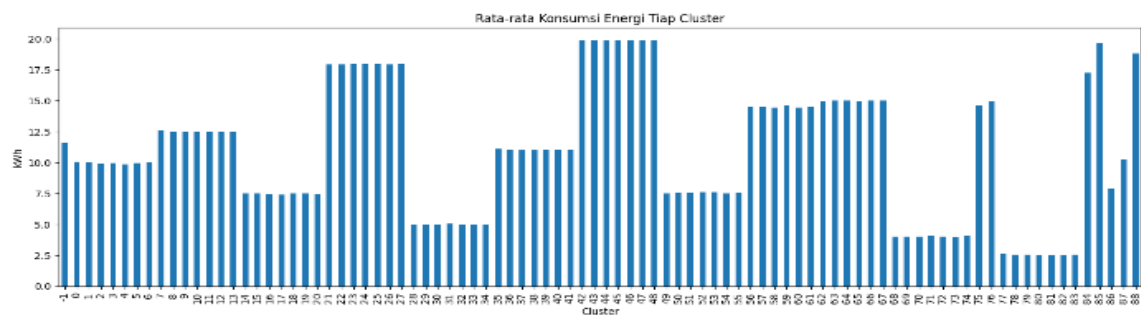
3.3 Analisis Karakteristik Cluster

Setelah proses pembentukan cluster selesai dilakukan, tahap berikutnya adalah menganalisis karakteristik setiap cluster untuk memahami pola penggunaan energi pada lingkungan smart home. Analisis dilakukan berdasarkan nilai rata-rata konsumsi energi, konsumsi energi pada jam puncak (Peak Hours Usage), jumlah penghuni rumah, suhu rata-rata lingkungan, kepemilikan pendingin ruangan (Has_AC), serta distribusi hari pengamatan (Weekday). Analisis ini bertujuan untuk mengidentifikasi perbedaan karakteristik antar kelompok sehingga dapat memberikan gambaran mengenai variasi perilaku konsumsi energi rumah tangga. Ringkasan karakteristik setiap cluster diperoleh melalui proses agregasi statistik menggunakan nilai rata-rata dan jumlah anggota pada masing-masing cluster. Hasil ringkasan statistik ditunjukkan pada Tabel 7.

Tabel 7. Statistik Karakteristik Cluster

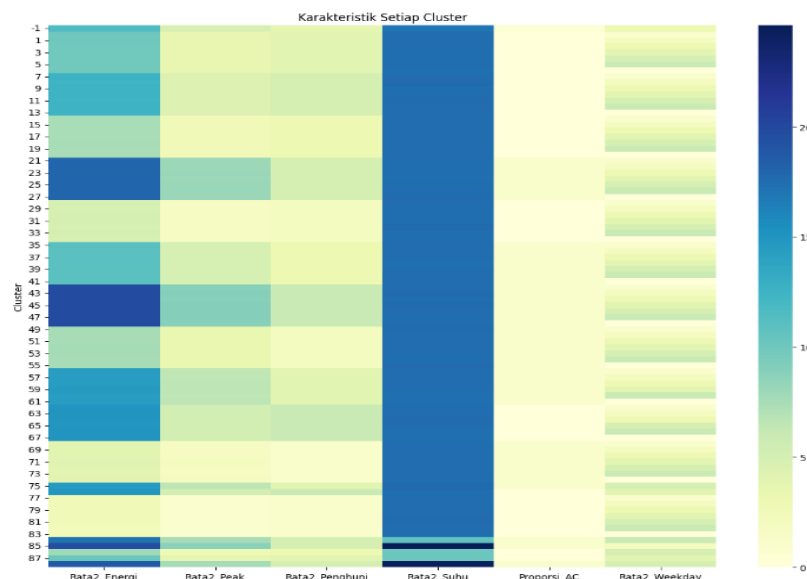
Cluster	Jumlah Data	Rata-rata Energi (kWh)	Peak Hours (kWh)	Penghuni	Suhu (°C)	AC	Weekday
-1	541	11.62	4.81	3.67	17.06	0.55	3.06
0	1053	10.05	3.52	4.00	17.52	0.00	1.00
1	1051	10.02	3.52	4.00	17.57	0.00	2.00

Tabel 7 menyajikan ringkasan karakteristik beberapa cluster yang paling representatif. Statistik lengkap seluruh 89 cluster digunakan pada proses analisis, namun tidak ditampilkan seluruhnya untuk menjaga keterbacaan artikel..



Gambar 4. Rata-rata Konsumsi Energi Tiap Cluster

Berdasarkan Gambar 4, rata-rata konsumsi energi berbeda pada setiap cluster, menunjukkan adanya variasi pola penggunaan energi rumah tangga. Perbedaan ini dipengaruhi oleh faktor seperti konsumsi pada jam puncak, jumlah penghuni, dan penggunaan pendingin ruangan. Hasil tersebut menunjukkan bahwa **DBSCAN** mampu mengelompokkan rumah tangga berdasarkan karakteristik konsumsi energi secara efektif sebagai dasar pengelolaan dan efisiensi energi pada smart home.



Gambar 5. Heatmap Karakteristik Setiap Cluster

Berdasarkan Gambar 5, heatmap menunjukkan variasi karakteristik pada setiap cluster berdasarkan rata-rata konsumsi energi, konsumsi pada jam puncak, jumlah penghuni, suhu, kepemilikan pendingin ruangan (Has_AC), dan weekday. Beberapa cluster memiliki konsumsi energi dan penggunaan energi pada jam puncak yang tinggi, sedangkan cluster lainnya menunjukkan pola konsumsi yang lebih rendah. Variasi pada variabel Has_AC mengindikasikan bahwa penggunaan pendingin ruangan berkontribusi terhadap peningkatan konsumsi energi, sementara perbedaan nilai weekday menunjukkan adanya pengaruh waktu penggunaan terhadap pola konsumsi. Secara keseluruhan, hasil ini menunjukkan bahwa algoritma DBSCAN mampu mengidentifikasi kelompok rumah tangga dengan karakteristik penggunaan energi yang berbeda, sehingga dapat dimanfaatkan sebagai dasar dalam pengembangan sistem smart energy management dan strategi efisiensi energi pada lingkungan smart home.

3.4 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) mampu mengelompokkan pola penggunaan energi rumah tangga berdasarkan karakteristik kepadatan data. Dengan parameter terbaik $\epsilon = 0,38$ dan $\text{MinPts} = 5$, algoritma menghasilkan 89 cluster, 541 data noise (0,60%), serta Silhouette Score sebesar 0,0690. Hasil tersebut menunjukkan bahwa DBSCAN mampu membentuk kelompok data berdasarkan kemiripan karakteristik konsumsi energi tanpa memerlukan penentuan jumlah cluster di awal proses analisis.

Berbeda dengan algoritma K-Means yang mengharuskan penentuan jumlah cluster sebelum proses pengelompokan dilakukan, DBSCAN membentuk cluster berdasarkan distribusi kepadatan data. Karakteristik tersebut menjadikan DBSCAN lebih fleksibel untuk menganalisis data konsumsi energi rumah tangga yang memiliki distribusi tidak seragam dan berpotensi mengandung data pencilan (outlier). Pada penelitian ini, hanya sekitar 0,60% data yang dikategorikan sebagai noise, sedangkan lebih dari 99% observasi berhasil dikelompokkan ke dalam cluster tertentu. Hal tersebut menunjukkan bahwa sebagian besar data memiliki pola konsumsi energi yang konsisten sehingga dapat diidentifikasi sebagai kelompok dengan karakteristik serupa.

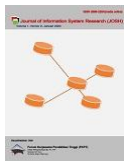
Nilai Silhouette Score sebesar 0,0690 menunjukkan bahwa pemisahan antar cluster masih relatif rendah. Nilai tersebut tidak serta-merta menunjukkan bahwa hasil clustering buruk, melainkan menggambarkan karakteristik data konsumsi energi rumah tangga yang bersifat kontinu sehingga batas antar kelompok tidak terbentuk secara tegas. Dataset yang digunakan juga memiliki jumlah observasi yang besar (sekitar 90.000 data) dengan variasi konsumsi energi yang tinggi. Kondisi tersebut menyebabkan sebagian cluster saling berdekatan sehingga menurunkan nilai Silhouette Score meskipun struktur kepadatan lokal masih dapat dikenali oleh DBSCAN.

Analisis karakteristik cluster menunjukkan bahwa konsumsi energi harian dan konsumsi energi pada jam puncak (Peak_Hours_Usage_kWh) menjadi variabel yang paling membedakan antar kelompok rumah tangga. Cluster dengan konsumsi energi rata-rata tinggi umumnya juga memiliki konsumsi energi yang tinggi pada jam puncak, sedangkan cluster dengan konsumsi energi rendah menunjukkan pola penggunaan energi yang lebih stabil. Selain itu, jumlah penghuni rumah, suhu rata-rata lingkungan, dan kepemilikan pendingin ruangan (Has_AC) turut memberikan kontribusi terhadap pembentukan cluster. Variabel-variabel tersebut menggambarkan bahwa pola penggunaan energi tidak hanya dipengaruhi oleh besarnya konsumsi listrik, tetapi juga oleh karakteristik rumah tangga dan kondisi lingkungan.

Hasil visualisasi menggunakan Principal Component Analysis (PCA) memperlihatkan bahwa sebagian besar cluster membentuk kelompok dengan distribusi yang relatif rapat, sementara sejumlah kecil observasi berada di luar kelompok utama dan diklasifikasikan sebagai noise. Keberadaan data noise tersebut mengindikasikan adanya pola penggunaan energi yang berbeda dari mayoritas rumah tangga, misalnya akibat penggunaan energi yang sangat tinggi atau sangat rendah dibandingkan kelompok lainnya. Kemampuan DBSCAN dalam mengidentifikasi data pencilan menjadi salah satu keunggulan algoritma ini karena observasi yang memiliki karakteristik ekstrem tidak dipaksa masuk ke dalam cluster tertentu.

Jika dibandingkan dengan penelitian-penelitian sebelumnya yang menggunakan metode clustering berbasis centroid, seperti K-Means, pendekatan DBSCAN memberikan beberapa kelebihan. Pertama, jumlah cluster tidak perlu ditentukan sebelum proses analisis dilakukan. Kedua, DBSCAN mampu mendeteksi data pencilan secara otomatis sehingga hasil segmentasi lebih representatif terhadap distribusi data sebenarnya. Ketiga, algoritma ini mampu membentuk cluster dengan bentuk yang tidak beraturan (arbitrary shape), yang sulit dicapai oleh metode clustering yang mengasumsikan bentuk cluster berbentuk sferis. Karakteristik tersebut menjadikan DBSCAN sesuai untuk menganalisis data konsumsi energi pada lingkungan smart home yang memiliki variasi perilaku penggunaan energi antar rumah tangga.

Meskipun demikian, penelitian ini juga memiliki beberapa keterbatasan. Nilai parameter ϵ dan MinPts sangat memengaruhi hasil clustering sehingga proses penentuan parameter memerlukan tahapan pengujian yang cukup intensif. Selain itu, dataset yang digunakan berasal dari satu jenis sumber data sehingga karakteristik pola penggunaan energi yang diperoleh belum tentu dapat digeneralisasikan pada kondisi rumah tangga dengan karakteristik wilayah atau musim yang berbeda. Penelitian selanjutnya dapat mengembangkan pendekatan ini dengan menggunakan dataset yang mencakup periode waktu yang lebih panjang, menambahkan variabel sensor



Internet of Things (IoT), atau membandingkan DBSCAN dengan algoritma clustering lain seperti OPTICS, HDBSCAN, atau Spectral Clustering untuk memperoleh hasil segmentasi yang lebih komprehensif.

Secara keseluruhan, hasil penelitian menunjukkan bahwa algoritma DBSCAN mampu mengidentifikasi pola penggunaan energi rumah tangga berdasarkan karakteristik kepadatan data serta mendeteksi data pencilan secara efektif. Informasi yang dihasilkan dari proses clustering dapat dimanfaatkan sebagai dasar dalam pengembangan sistem smart energy management, penyusunan strategi efisiensi energi, maupun perancangan layanan personalisasi konsumsi energi pada lingkungan smart home. Segmentasi pengguna berdasarkan pola konsumsi energi juga berpotensi mendukung pengambilan keputusan dalam penerapan program penghematan energi dan pengelolaan beban listrik secara lebih adaptif.

4. KESIMPULAN

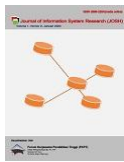
Penelitian ini berhasil mengidentifikasi pola penggunaan energi pada lingkungan smart home menggunakan algoritma Density-Based Spatial Clustering of Applications with Noise (DBSCAN) berbasis data time series sensor. Berdasarkan proses penentuan parameter menggunakan K-Distance Graph dan evaluasi Silhouette Score, diperoleh parameter terbaik $\epsilon = 0,38$ dan $\text{MinPts} = 5$, yang menghasilkan 89 cluster, 541 data noise (0,60% dari total data), serta Silhouette Score sebesar 0,0690. Hasil tersebut menunjukkan bahwa DBSCAN mampu mengelompokkan sebagian besar data berdasarkan tingkat kepadatan sekaligus mendeteksi data pencilan tanpa memerlukan penentuan jumlah cluster di awal. Analisis karakteristik cluster menunjukkan bahwa konsumsi energi, penggunaan energi pada jam puncak, jumlah penghuni, suhu rata-rata, dan kepemilikan pendingin ruangan merupakan faktor utama yang membedakan pola penggunaan energi antar kelompok rumah tangga. Meskipun nilai Silhouette Score relatif rendah, hasil clustering tetap mampu memberikan segmentasi yang representatif karena karakteristik data konsumsi energi bersifat kontinu dan memiliki batas antar kelompok yang tidak tegas. Dengan demikian, tujuan penelitian untuk mengidentifikasi pola penggunaan energi rumah tangga menggunakan DBSCAN telah tercapai. Temuan penelitian ini memberikan dasar bagi pengembangan smart energy management, deteksi anomali konsumsi energi, serta penelitian lanjutan yang membandingkan DBSCAN dengan algoritma density-based clustering lainnya pada dataset IoT yang lebih beragam.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Direktorat Penelitian dan Pengabdian kepada Masyarakat (DPPM) serta Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa atas dukungan, fasilitasi, dan kesempatan yang diberikan sehingga penelitian ini dapat terlaksana dengan baik. Penulis juga menyampaikan apresiasi kepada penyedia Household Energy Consumption Dataset melalui platform Kaggle yang telah menyediakan data penelitian sebagai sumber utama dalam proses analisis. Selain itu, penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan, masukan, dan kontribusi selama proses penyusunan artikel ini.

REFERENCES

- [1] A. M. Alrasyid, Rivai, N. R. Diasri, D. Ulandari, and R. P. Laksana, "Pengaruh Teknologi Internet of Things (IoT) terhadap Efisiensi Energi di Smart Home," *Journal of Information Systems Management and Digital Business*, vol. 2, no. 3, pp. 223–230, Apr. 2025, doi: 10.70248/jismdb.v2i3.2209.
- [2] A. M. Norouzzadeh, S. P. Toufighi, J. Vang, and A. Edalatpour, "Adoption of internet of things in residential smart homes: A structural equation modeling approach," *Sustainable Futures*, vol. 9, p. 100665, Jun. 2025, doi: 10.1016/j.sfr.2025.100665.
- [3] T. L. Narayana et al., "Advances in real time smart monitoring of environmental parameters using IoT and sensors," *Heliyon*, vol. 10, no. 7, p. e28195, Apr. 2024, doi: 10.1016/j.heliyon.2024.e28195.
- [4] A. Eltaleb et al., "Smart Home Sensor Systems: Advancements and Applications," *International Journal of Electrical Engineering and Sustainability*, vol. 1, no. 3, pp. 120–228, Aug. 2023, doi: 10.65998/ijeess.v1i3.55.
- [5] A. Hafiz, F. Amanah, R. Damurti, and P. Harliana, "Mengidentifikasi Pola Konsumsi Energi Rumah Tangga Menggunakan Algoritma Graf Berbasis C++," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 6, pp. 12057–12061, Nov. 2024, doi: 10.36040/jati.v8i6.11788.
- [6] A. Pai H, K. K. Mishra, M. T. R, J. V. M. L. Jeyan, and A. Sayal, "Enhanced household energy consumption forecasting using multivariate long short-term memory (LSTM) networks with weather data integration," *Results in Engineering*, vol. 27, p. 106512, Sep. 2025, doi: 10.1016/j.rineng.2025.106512.
- [7] Muhammad Refly Adrian, "Analisis Pengaruh Teknologi IoT terhadap Optimalisasi Efisiensi Smart City," *Merkurius : Jurnal Riset Sistem Informasi dan Teknik Informatika*, vol. 3, no. 2, pp. 26–35, Feb. 2025, doi: 10.61132/mercurius.v3i2.702.
- [8] R. Ramadani, "Potensi Internet of Things (IoT) sebagai Sumber Official Statistics Bidang Pertanian," *Seminar Nasional Official Statistics*, vol. 2023, no. 1, pp. 161–166, Oct. 2023, doi: 10.34123/semnasoffstat.v2023i1.1900.
- [9] J. Zhao, F. Chu, L. Xie, Y. Che, Y. Wu, and A. F. Burke, "A survey of transformer networks for time series forecasting," *Comput. Sci. Rev.*, vol. 60, p. 100883, May 2026, doi: 10.1016/j.cosrev.2025.100883.
- [10] E. S. Ortigossa, F. F. Dias, D. C. Nascimento, and L. G. Nonato, "Time Series Information Visualization - A Review of Approaches and Tools," *IEEE Access*, vol. 13, pp. 161653–161684, 2025, doi: 10.1109/ACCESS.2025.3609404.



- [11] J. Parhusip, N. I. Sriyanto, T. P. -, Z. C. Mavanudin, and D. S. Augustin, “Analisis Pola Konsumsi Energi Bangunan Berdasarkan Data Sensor Lingkungan Berbasis Algoritma Apriori,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 14, no. 1, Jan. 2026, doi: 10.23960/jitet.v14i1.8524.
- [12] L. Lei, B. Wu, X. Fang, L. Chen, H. Wu, and W. Liu, “A dynamic anomaly detection method of building energy consumption based on data mining technology,” *Energy*, vol. 263, p. 125575, Jan. 2023, doi: 10.1016/j.energy.2022.125575.
- [13] Y. Andi Rozzi, D. Dwi Andiska, and F. Putra Winarta, “Rancang Bangun Sistem Smart Home Berbasis IoT untuk Kendali Lampu dan Pemantauan Suhu,” *Jurnal Surya Energy*, pp. 24–30, Sep. 2025, doi: 10.32502/jse.v10i1.939.
- [14] R. El Abed, A. Hammoud, M. El-Gohary, and B. Taher, “Recent optimization methods and techniques in residential home energy management systems,” *Results in Engineering*, vol. 29, p. 108974, Mar. 2026, doi: 10.1016/j.rineng.2026.108974.
- [15] R. I. Pramudya, T. A. Kurniawan, M. H. Candra, C. W. Onn, and K. K. Dissanayake, “Advancing unsupervised clustering: A systematic review of hybrid K-means and metaheuristic optimization algorithms in data mining,” *Comput. Sci. Rev.*, vol. 61, p. 100972, Aug. 2026, doi: 10.1016/j.cosrev.2026.100972.
- [16] S. M. Miraftebadeh, C. G. Colombo, M. Longo, and F. Foiadelli, “K-Means and Alternative Clustering Methods in Modern Power Systems,” *IEEE Access*, vol. 11, pp. 119596–119633, 2023, doi: 10.1109/ACCESS.2023.3327640.
- [17] M. G. H. Omran, A. P. Engelbrecht, and A. Salman, “An overview of clustering methods,” *Intelligent Data Analysis*, vol. 11, no. 6, pp. 583–605, Nov. 2007, doi: 10.3233/IDA-2007-11602.
- [18] I. B. G. Sarasvananda, R. Wardoyo, and A. K. Sari, “The K-Means Clustering Algorithm With Semantic Similarity To Estimate The Cost of Hospitalization,” *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 13, no. 4, p. 313, Oct. 2019, doi: 10.22146/ijccs.45093.
- [19] L. F. Naz, R. Qamar, R. Asif, S. Hina, M. Imran, and S. Ahmed, “Intelligent energy management in IoT-enabled smart homes: Anomaly detection and consumption prediction for energy-efficient usage,” *Mehran University Research Journal of Engineering and Technology*, vol. 44, no. 1, p. 113, Jan. 2025, doi: 10.22581/muet1982.3291.
- [20] A. Baset and M. Jradi, “Data-Driven Decision Support for Smart and Efficient Building Energy Retrofits: A Review,” *Applied System Innovation*, vol. 8, no. 1, p. 5, Dec. 2024, doi: 10.3390/asi8010005.
- [21] A. Apriani, O. K. Sulaiman, and A. Antoni, “Implementasi Algoritma Density Based Spatial Clustering of Applications with Noise (DBSCAN) pada Aplikasi Penerima Bantuan Langsung Tunai (BLT) Online di Desa Bahtera Makmur Rokan Hilir - Riau,” *Hello World Jurnal Ilmu Komputer*, vol. 4, no. 1, pp. 1–11, Mar. 2025, doi: 10.56211/helloworld.v4i1.613.
- [22] A. Apriani, O. K. Sulaiman, and A. Antoni, “Implementasi Algoritma Density Based Spatial Clustering of Applications with Noise (DBSCAN) pada Aplikasi Penerima Bantuan Langsung Tunai (BLT) Online di Desa Bahtera Makmur Rokan Hilir - Riau,” *Hello World Jurnal Ilmu Komputer*, vol. 4, no. 1, pp. 1–11, Mar. 2025, doi: 10.56211/helloworld.v4i1.613.
- [23] E. asfarina Fadlilah, “Identifikasi Anomali Data Akademik Menggunakan Dbscan Outlier Detection,” *Prosiding Sains Nasional dan Teknologi*, vol. 12, no. 1, pp. 336–342, Nov. 2022, doi: 10.36499/psnst.v12i1.7012.
- [24] Hanna Arini Parhusip et al., “Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Principal Component Analysis (PCA) for Particulate Matter (PM) Anomaly Detection,” *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, vol. 15, no. 02, pp. 75–86, Oct. 2025, doi: 10.24843/LKJITI.2024.v15.i02.p01.
- [25] J. Chen et al., “Improved DBSCAN-Based Electricity Theft Detection Using Spatiotemporal Fusion Features,” *Applied Sciences*, vol. 15, no. 22, p. 12028, Nov. 2025, doi: 10.3390/app152212028.
- [26] Q. Li, Y. Ma, and Y. Wu, “Utilize DBN and DBSCAN to detect selective forwarding attacks in event-driven wireless sensors networks,” *Eng. Appl. Artif. Intell.*, vol. 126, p. 107122, Nov. 2023, doi: 10.1016/j.engappai.2023.107122.