

Received : 08 May 2026

Revised : 20 May 2026

Published : 10 June 2026

Issued : 30 June 2026



Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Pengelompokan Konsumsi Listrik Rumah Tangga Menggunakan Silhouette Coefficient dan Davies-Bouldin Index

Alief Nur Aisyi Maulidhia^{1,a)*}

¹Politeknik Perkapalan Negeri Surabaya, Jl. Teknik Kimia, Keputih, Kec. Sukolilo, Kota Surabaya, Jawa Timur 60111

*Email: ^{a)}aliefnur@ppns.ac.id

Abstrak

Pengelompokan pola konsumsi listrik rumah tangga merupakan pendekatan penting dalam memahami karakteristik penggunaan energi dan mendukung pengambilan keputusan berbasis data pada sistem kelistrikan. Penelitian ini bertujuan untuk membandingkan kinerja algoritma K-Means dan Hierarchical Clustering dalam mengelompokkan pola konsumsi listrik rumah tangga menggunakan dataset *Household Electric Power Consumption* yang diperoleh dari Kaggle. Data yang digunakan meliputi beberapa parameter kelistrikan, seperti *Global Active Power*, *Voltage*, dan *Current*, yang melalui tahapan *pre-processing* berupa pembersihan data, normalisasi menggunakan *Min-Max Scaling*, dan reduksi dimensi menggunakan *Principal Component Analysis (PCA)*. Evaluasi kualitas kluster dilakukan menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI). Hasil penelitian menunjukkan bahwa algoritma K-Means menghasilkan performa klusterisasi yang lebih baik dibandingkan dengan Hierarchical Clustering, dengan nilai Silhouette Coefficient sebesar 0,67 dan Davies-Bouldin Index sebesar 0,48, sedangkan Hierarchical Clustering memperoleh nilai Silhouette Coefficient sebesar 0,59 dan Davies-Bouldin Index sebesar 0,71. Temuan ini menunjukkan bahwa K-Means lebih efektif dalam mengelompokkan data konsumsi listrik rumah tangga yang terstruktur, sedangkan Hierarchical Clustering lebih sesuai untuk mengidentifikasi hubungan data yang bersifat kompleks. Penelitian ini memberikan kontribusi dalam pemilihan metode klusterisasi yang tepat untuk analisis pola konsumsi listrik rumah tangga serta dapat mendukung pembelajaran analisis data kelistrikan pada pendidikan vokasi.

Kata-kunci: Konsumsi Listrik, K-Means, Hierarchical Clustering, Silhouette Coefficient, Davies-Bouldin Index, Clustering.

Abstract

The clustering of household electricity consumption patterns is an important approach for understanding energy usage characteristics and supporting data-driven decision-making in electrical

systems. This study aims to compare the performance of K-Means and Hierarchical Clustering algorithms in grouping household electricity consumption patterns using the *Household Electric Power Consumption* dataset obtained from Kaggle. The dataset includes electrical parameters such as *Global Active Power*, *Voltage*, and *Current*, which underwent preprocessing stages including data cleaning, normalization using *Min-Max Scaling*, and dimensionality reduction through *Principal Component Analysis (PCA)*. Cluster quality evaluation was conducted using the Silhouette Coefficient and Davies-Bouldin Index (DBI). The results indicate that the K-Means algorithm achieved better clustering performance than Hierarchical Clustering, with a Silhouette Coefficient of 0.67 and a Davies-Bouldin Index of 0.48, while Hierarchical Clustering obtained a Silhouette Coefficient of 0.59 and a Davies-Bouldin Index of 0.71. These findings suggest that K-Means is more effective for clustering structured household electricity consumption data, whereas Hierarchical Clustering is more suitable for identifying complex relationships within the data. This study contributes to selecting appropriate clustering methods for household electricity consumption analysis and supports electrical data analysis learning in vocational education.

Keywords: Electricity Consumption, K-Means, Hierarchical Clustering, Silhouette Coefficient, Davies-Bouldin Index, Clustering.

Pendahuluan

Konsumsi energi listrik rumah tangga memiliki peran penting dalam keberlanjutan sistem energi dan efisiensi penggunaan sumber daya listrik. Peningkatan kebutuhan energi yang dipengaruhi oleh pertumbuhan populasi, perkembangan teknologi, serta peningkatan penggunaan perangkat elektronik rumah tangga menyebabkan pola konsumsi listrik menjadi semakin kompleks. Oleh karena itu, pemahaman terhadap pola konsumsi listrik rumah tangga menjadi aspek penting dalam mendukung efisiensi energi (Anagra, 2020)(Hendrawan, 2024), pengelolaan beban listrik, serta pengambilan keputusan berbasis data pada sistem kelistrikan rumah tangga. Salah satu pendekatan yang dapat digunakan untuk memahami karakteristik konsumsi energi adalah teknik *clustering*, yaitu metode *unsupervised learning* yang mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik tanpa memerlukan label sebelumnya (Impron, 2025)

Penerapan *clustering* pada data konsumsi listrik rumah tangga dapat memberikan manfaat dalam segmentasi pola penggunaan energi, identifikasi konsumsi abnormal, pengelompokan karakteristik pelanggan, serta penyusunan strategi efisiensi energi yang lebih tepat (Jamilatun et al., 2025). Dalam konteks pendidikan vokasi kelistrikan, analisis pola konsumsi listrik juga dapat dimanfaatkan sebagai media pembelajaran berbasis data (*data-driven learning*) untuk meningkatkan pemahaman mahasiswa terhadap karakteristik sistem tenaga listrik dan perilaku beban listrik rumah tangga (Toba et al., 2023).

Berbagai algoritma *clustering* telah digunakan dalam penelitian sebelumnya, di antaranya K-Means dan Hierarchical Clustering (Yanuar, 2022)(Brina & Rizky, 2025) Algoritma K-Means merupakan metode berbasis partisi yang mengelompokkan data berdasarkan kedekatan terhadap pusat klaster (*centroid*) sehingga memiliki keunggulan dalam efisiensi komputasi dan performa pada data yang terstruktur (Rahmati et al., 2021). Namun, metode ini memiliki keterbatasan ketika menangani data dengan distribusi tidak teratur, sensitivitas terhadap *outlier*, serta memerlukan penentuan jumlah klaster di awal proses. Sebaliknya, Hierarchical Clustering mampu membentuk struktur hubungan antar data secara hierarkis melalui *dendrogram* dan lebih fleksibel dalam menangani pola distribusi data yang kompleks tanpa harus menentukan jumlah klaster secara eksplisit (Alamtaha et al., 2023). Meskipun demikian, metode ini cenderung membutuhkan waktu komputasi yang lebih tinggi, terutama pada dataset berskala besar.

Beberapa penelitian terdahulu telah menerapkan algoritma *clustering* untuk berbagai permasalahan klasifikasi dan segmentasi data, termasuk pada bidang energi dan konsumsi listrik. Namun, sebagian besar penelitian masih berfokus pada penggunaan satu algoritma tertentu atau belum memberikan evaluasi komparatif yang komprehensif terhadap performa algoritma *clustering* pada data konsumsi

listrik rumah tangga. Selain itu, penelitian yang membandingkan K-Means dan Hierarchical Clustering dengan pendekatan evaluasi kuantitatif menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI) pada dataset konsumsi listrik rumah tangga masih terbatas (Rohmah et al., 2021) (Gianyar, 2024). Kondisi ini menunjukkan adanya research gap, yaitu belum tersedianya rekomendasi algoritma *clustering* yang paling sesuai untuk karakteristik data konsumsi listrik rumah tangga berdasarkan evaluasi performa yang terukur.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk membandingkan performa algoritma K-Means dan Hierarchical Clustering dalam pengelompokan pola konsumsi listrik rumah tangga menggunakan metrik evaluasi Silhouette Coefficient dan Davies-Bouldin Index. Hasil penelitian diharapkan dapat memberikan rekomendasi metode *clustering* yang lebih efektif untuk analisis pola konsumsi listrik rumah tangga serta mendukung pengembangan pembelajaran analisis data kelistrikan pada pendidikan vokasi.

Metode Penelitian

Dataset penelitian ini menggunakan dataset Individual Household Electric Power Consumption Dataset yang tersedia pada platform Kaggle dan berasal dari *UCI Machine Learning Repository*. Dataset ini berisi data konsumsi listrik rumah tangga yang dikumpulkan selama periode Desember 2006 hingga November 2010 dengan interval pencatatan setiap satu menit, sehingga menghasilkan sekitar 2.075.259 records. Dataset mencakup beberapa parameter kelistrikan yang relevan, seperti *Global Active Power*, *Global Reactive Power*, *Voltage*, *Global Intensity (Current)*, serta variabel lainnya yang berkaitan dengan pola konsumsi listrik rumah tangga. Penggunaan dataset ini dipilih karena memiliki karakteristik data time-series yang kompleks dan representatif untuk analisis pola konsumsi energi rumah tangga. Dataset dapat diakses melalui platform Kaggle dan digunakan sebagai sumber utama dalam proses analisis clustering. Kaggle Household Electric Power Consumption Dataset

Tahap awal penelitian dilakukan melalui proses *pre-processing* data untuk memastikan kualitas data sebelum diterapkan algoritma clustering. Proses ini meliputi pembersihan data (*data cleaning*), normalisasi data, dan reduksi dimensi. Pada tahap pembersihan data, dilakukan identifikasi terhadap data hilang (*missing values*), data duplikat, maupun data anomali (*outlier*). Penanganan data kosong dilakukan menggunakan metode imputasi berbasis rata-rata (*mean imputation*) atau interpolasi berdasarkan nilai terdekat agar tidak memengaruhi kualitas analisis.

Selanjutnya, dilakukan proses normalisasi data menggunakan metode Min-Max Scaling agar seluruh fitur memiliki rentang nilai yang seragam dan tidak menyebabkan dominasi variabel tertentu dalam perhitungan jarak antar data. Proses normalisasi dilakukan menggunakan Persamaan (1):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

di mana x' merupakan nilai hasil normalisasi, x adalah nilai asli data, x_{min} adalah nilai minimum, dan x_{max} adalah nilai maksimum pada fitur tertentu.

Untuk mengurangi kompleksitas data dengan jumlah variabel yang besar tanpa kehilangan informasi penting, penelitian ini menerapkan Principal Component Analysis (PCA). Metode PCA digunakan untuk mengekstraksi komponen utama yang mampu menjelaskan proporsi variasi terbesar pada dataset sehingga meningkatkan efisiensi komputasi selama proses clustering.

Pada tahap clustering, penelitian ini membandingkan dua algoritma, yaitu K-Means Clustering dan Hierarchical Clustering. Algoritma K-Means bekerja dengan membagi data ke dalam sejumlah klaster berdasarkan kedekatan terhadap pusat klaster (*centroid*). Perhitungan jarak antar data dilakukan menggunakan Euclidean Distance sebagaimana ditunjukkan pada Persamaan (2):

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

Penentuan jumlah kluster optimal dilakukan menggunakan Elbow Method dengan mengevaluasi perubahan nilai *inertia* pada setiap jumlah kluster (k). Berdasarkan hasil pengujian, terjadi penurunan nilai *inertia* yang signifikan hingga $k = 4$, sedangkan setelah nilai tersebut penurunan mulai melandai (*elbow point*). Oleh karena itu, jumlah kluster optimal ditetapkan sebanyak empat kluster karena memberikan keseimbangan terbaik antara kompleksitas model dan kualitas pengelompokan. Untuk memperkuat justifikasi, nilai *inertia* pada setiap pengujian jumlah kluster dapat ditampilkan pada tabel atau grafik hasil Elbow Method.

Selain K-Means, penelitian ini juga menerapkan Hierarchical Clustering menggunakan pendekatan Agglomerative Clustering dengan metode Single Linkage untuk menghitung jarak antarkluster. Perhitungan jarak minimum antar kluster dilakukan menggunakan Persamaan (3):

$$D(A, B) = \min_{x_i \in A, x_j \in B} d(x_i, x_j) \quad (3)$$

di mana A dan B merupakan kluster yang dibandingkan, sedangkan $d(x_i, x_j)$ adalah jarak Euclidean antar titik data.

Kualitas hasil clustering dievaluasi menggunakan dua metrik validasi, yaitu Silhouette Coefficient dan Davies-Bouldin Index (DBI). Silhouette Coefficient digunakan untuk mengukur tingkat kesamaan suatu data terhadap kluster yang dibandingkan dengan kluster lain. Nilai Silhouette berada pada rentang -1 hingga 1 , di mana nilai yang lebih tinggi menunjukkan kualitas pemisahan kluster yang lebih baik. Persamaan Silhouette Coefficient ditunjukkan pada Persamaan (4):

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4)$$

di mana $a(i)$ adalah rata-rata jarak data terhadap anggota dalam kluster yang sama dan $b(i)$ adalah rata-rata jarak data terhadap kluster terdekat.

Selanjutnya, Davies-Bouldin Index (DBI) digunakan untuk mengevaluasi tingkat pemisahan antarkluster dengan mempertimbangkan kepadatan kluster dan jarak antarpusat kluster. Nilai DBI yang lebih kecil menunjukkan kualitas clustering yang lebih baik. Persamaan DBI ditunjukkan pada Persamaan (5):

$$DBI = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \left(\frac{S_i + S_j}{M_{ij}} \right) \quad (5)$$

dengan S_i dan S_j sebagai ukuran dispersi kluster, sedangkan M_{ij} merupakan jarak antar pusat kluster.

Hasil dan Pembahasan

A. Persiapan Data

Dataset konsumsi listrik rumah tangga yang digunakan pada penelitian ini ditampilkan pada Gambar 1. Dataset terdiri dari beberapa parameter kelistrikan utama, seperti *Global Active Power*, *Voltage*, dan *Current*, yang digunakan untuk merepresentasikan pola konsumsi energi rumah tangga. Tahap persiapan data dilakukan untuk memastikan bahwa data yang digunakan telah sesuai dengan kebutuhan analisis clustering. Dataset awal menunjukkan adanya variasi nilai konsumsi listrik antarwaktu pengamatan, yang mengindikasikan keberadaan pola konsumsi listrik yang berbeda pada setiap rumah tangga. Variasi ini menjadi dasar penting dalam proses pengelompokan menggunakan metode clustering karena memungkinkan identifikasi kelompok pelanggan berdasarkan karakteristik penggunaan energi.

Date/Time	Active Power (kW)	Voltage (V)	Current (A)
1/1/2023 0:00	2.185430535	201.5714593	14.63047469
1/1/2023 1:00	4.778214379	231.8205206	6.262099475
1/1/2023 2:00	3.793972738	215.7177991	7.424430711
1/1/2023 3:00	3.193963179	225.4285346	18.47831283
1/1/2023 4:00	1.202083882	245.3783237	14.09643589
1/1/2023 5:00	1.201975342	212.4646115	5.137955774
1/1/2023 6:00	0.761376255	220.5191462	6.522073143
1/1/2023 7:00	4.397792656	237.7775569	14.95252654
1/1/2023 8:00	3.205017553	211.4399083	5.075923758
1/1/2023 9:00	3.6863266	203.8489955	7.412120771
1/1/2023 10:00	0.592630224	214.4875726	13.23100684
1/1/2023 11:00	4.864594335	208.0610644	15.37842797
1/1/2023 12:00	4.245991884	246.4848826	14.77941889
1/1/2023 13:00	1.455525998	240.406019	8.364039642
1/1/2023 14:00	1.318212352	231.6701878	15.68268832
1/1/2023 15:00	1.325320294	243.5730295	8.558736312
1/1/2023 16:00	1.869090093	240.1836038	9.880995472
1/1/2023 17:00	2.861403942	209.3285029	16.19737108
1/1/2023 18:00	2.443752584	244.6279499	14.74449349
1/1/2023 19:00	1.810531131	226.9671121	17.73835116
1/1/2023 20:00	3.253338026	240.3720078	14.86419338
1/1/2023 21:00	1.127722373	244.804565	13.52462905
1/1/2023 22:00	1.814650918	215.9001737	6.405121517
1/1/2023 23:00	2.148628295	205.5025962	10.51573705

Gambar 1 Dataset Konsumsi Listrik Rumah Tangga

B. Data Pre-Processing

Tahap *pre-processing* dilakukan melalui proses normalisasi untuk menyamakan rentang nilai antar fitur agar tidak terjadi dominasi variabel tertentu dalam proses pengukuran jarak. Normalisasi menggunakan metode *Min-Max Scaling* diterapkan pada seluruh fitur numerik sehingga data memiliki rentang nilai antara 0 hingga 1.

Hasil normalisasi menunjukkan bahwa distribusi data menjadi lebih seragam sehingga meningkatkan kualitas proses clustering. Tahapan ini sangat penting karena algoritma berbasis jarak, seperti K-Means dan Hierarchical Clustering, sangat sensitif terhadap perbedaan skala data. Tanpa normalisasi, fitur dengan rentang besar dapat mendominasi pembentukan kluster.

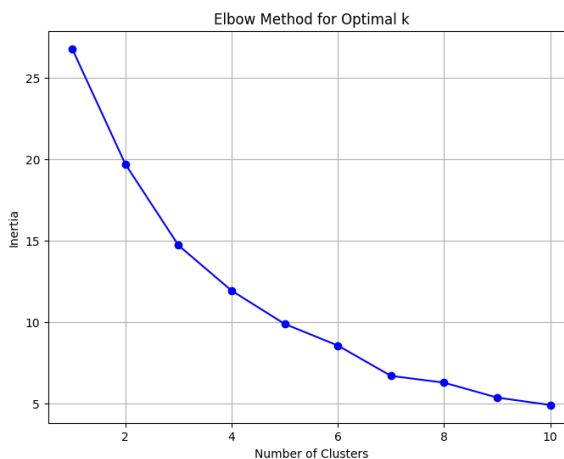
Date/Time	Active Power (kW)	Voltage (V)	Current (A)
1/1/2023 0:00	0.376025301	0.025009806	0.646675191
1/1/2023 1:00	0.963140486	0.643158638	0.080283251
1/1/2023 2:00	0.740266831	0.314094592	0.158952649
1/1/2023 3:00	0.60439946	0.512536446	0.907106213
1/1/2023 4:00	0.153354309	0.920216497	0.61053015
1/1/2023 5:00	0.15332973	0.247614707	0.004198477
1/1/2023 6:00	0.053559588	0.41221159	0.097878899
1/1/2023 7:00	0.876997027	0.764892501	0.668472442
1/1/2023 8:00	0.606902635	0.226674584	0
1/1/2023 9:00	0.715891223	0.071551956	0.158119483
1/1/2023 10:00	0.015348397	0.288954538	0.551955811
1/1/2023 11:00	0.982700537	0.157626873	0.697298484
1/1/2023 12:00	0.842622955	0.942829368	0.656756096
1/1/2023 13:00	0.210744251	0.818605928	0.222547662
1/1/2023 14:00	0.179650673	0.640086542	0.717891563
1/1/2023 15:00	0.181260209	0.883324758	0.235725205
1/1/2023 16:00	0.304392536	0.814060806	0.325218913
1/1/2023 17:00	0.529094078	0.183527369	0.752726541
1/1/2023 18:00	0.434520265	0.90488238	0.65439226
1/1/2023 19:00	0.291132327	0.543977749	0.857023812
1/1/2023 20:00	0.61784442	0.817910898	0.66249384
1/1/2023 21:00	0.136515739	0.908491563	0.571828875
1/1/2023 22:00	0.29206522	0.317821475	0.08996333
1/1/2023 23:00	0.367691728	0.105343793	0.368179763

Gambar 2 Dataset Setelah Normalisasi

C. Penerapan Algoritma K-Means

Algoritma K-Means diterapkan untuk mengelompokkan data konsumsi listrik rumah tangga berdasarkan kemiripan karakteristik penggunaan energi. Penentuan jumlah kluster dilakukan menggunakan Elbow Method dengan mengevaluasi perubahan nilai *inertia* terhadap jumlah kluster.

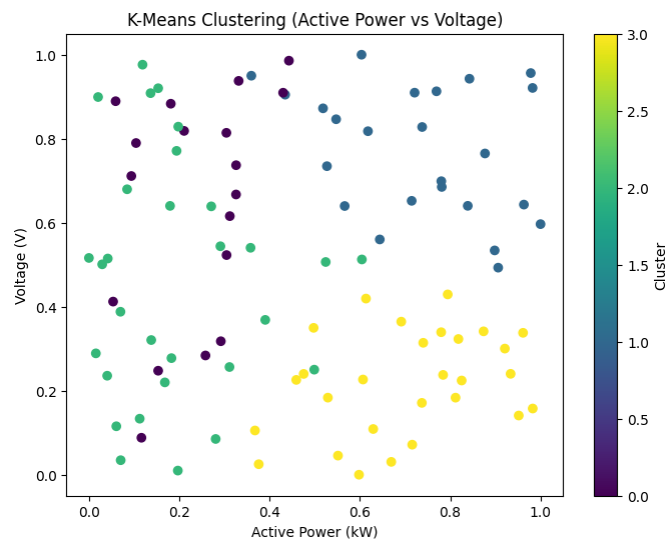
Berdasarkan grafik Elbow Method, terjadi penurunan nilai *inertia* yang signifikan dari jumlah kluster kecil hingga $k = 4$. Setelah titik tersebut, penurunan *inertia* cenderung melandai, yang menunjukkan bahwa penambahan jumlah kluster tidak lagi memberikan peningkatan kualitas model secara signifikan. Oleh karena itu, jumlah kluster optimal ditetapkan sebanyak empat kluster.



Gambar 3 Elbow Method for Optimal K

Setelah jumlah kluster ditentukan, algoritma K-Means digunakan untuk mengelompokkan data berdasarkan kedekatan terhadap pusat kluster (*centroid*).

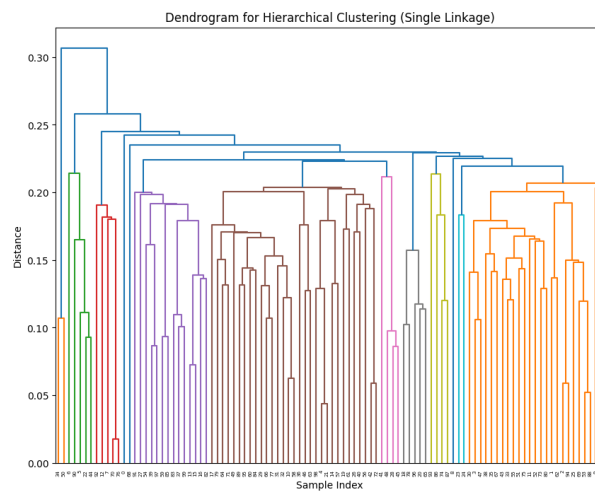
Visualisasi hasil K-Means menunjukkan bahwa data berhasil terpisah ke dalam empat kelompok dengan karakteristik konsumsi listrik yang berbeda. Sebagian besar data terlihat terkonsentrasi di area tertentu, yang mengindikasikan adanya pola konsumsi dominan. Hasil ini menunjukkan bahwa K-Means cukup efektif dalam mengidentifikasi pola konsumsi listrik pada data yang memiliki struktur relatif jelas. Temuan ini sejalan dengan penelitian yang dilakukan oleh (Rahmati et al., 2021) yang menunjukkan bahwa algoritma K-Means memiliki performa baik pada data terstruktur karena efisiensi komputasinya dan kemampuan membentuk kelompok data yang relatif homogen. Selain itu, penelitian (Herdiaman et al., 2024) juga menunjukkan bahwa K-Means efektif dalam segmentasi data berbasis karakteristik numerik. Dalam penelitian ini, performa K-Means diperkuat oleh nilai Silhouette Coefficient sebesar 0,67 dan Davies-Bouldin Index sebesar 0,48, yang menunjukkan kualitas pemisahan kluster yang baik.



Gambar 4 K-Means Clustering

D. Penerapan Algoritma Hierarchical Clustering

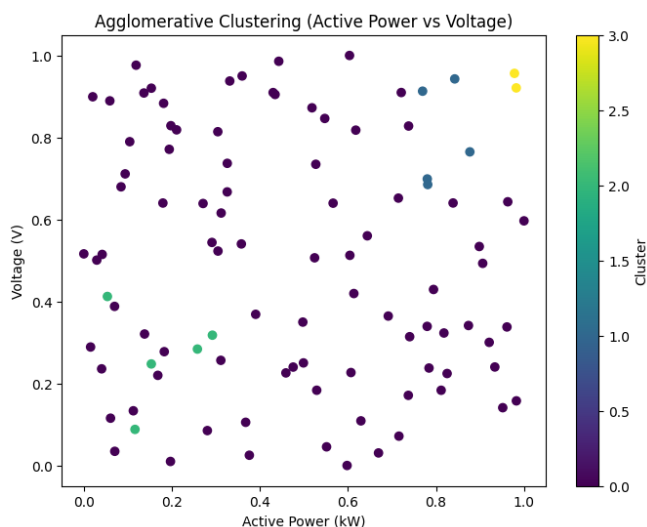
Selain K-Means, penelitian ini menerapkan Hierarchical Clustering menggunakan pendekatan Agglomerative Clustering dengan metode *Single Linkage*. Berbeda dengan K-Means, metode ini membentuk hubungan antardata secara bertahap berdasarkan tingkat kemiripan antartitik data.



Gambar 5 Dendrogram for Hierarchical Clustering

Dendrogram pada hasil Hierarchical Clustering menunjukkan proses penggabungan data ke dalam kelompok berdasarkan tingkat kemiripan. Struktur hierarki pada dendrogram memungkinkan identifikasi hubungan antarkluster secara lebih mendalam dan dapat digunakan untuk memahami karakteristik data secara bertingkat.

Selanjutnya, proses *Agglomerative Clustering* menghasilkan pengelompokan data berdasarkan hubungan hierarkis yang terbentuk.



Gambar 6 Agglomerative Clustering

Visualisasi hasil clustering menunjukkan bahwa Hierarchical Clustering mampu mengidentifikasi hubungan antar data yang lebih kompleks dibandingkan K-Means. Namun, pemisahan antarkelompok terlihat kurang tegas sehingga menghasilkan tingkat overlap yang lebih tinggi pada beberapa area data.

Hasil penelitian ini menunjukkan kesesuaian dengan penelitian (Alamataha et al., 2023) yang menyatakan bahwa Hierarchical Clustering lebih fleksibel dalam menangani pola distribusi data yang tidak teratur. Namun, penelitian ini menemukan bahwa kualitas klaster yang dihasilkan masih berada di bawah K-Means berdasarkan nilai Silhouette Coefficient sebesar 0,59 dan Davies-Bouldin Index sebesar 0,71. Hal ini menunjukkan bahwa meskipun Hierarchical Clustering mampu menangkap hubungan data yang kompleks, kualitas pemisahan klasternya kurang optimal pada dataset konsumsi listrik rumah tangga yang digunakan.

- E. Evaluasi kuantitatif menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI). Evaluasi kualitas clustering dilakukan menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI) untuk menentukan algoritma yang memberikan performa terbaik. Nilai Silhouette yang lebih tinggi menunjukkan kualitas pengelompokan yang lebih baik, sedangkan nilai DBI yang lebih rendah menunjukkan tingkat pemisahan klaster yang lebih optimal.

Berdasarkan Tabel 1, algoritma K-Means memperoleh nilai Silhouette Coefficient sebesar 0,67 dan Davies-Bouldin Index sebesar 0,48, sedangkan Hierarchical Clustering memperoleh nilai Silhouette sebesar 0,59 dan DBI sebesar 0,71. Hasil ini menunjukkan bahwa K-Means memiliki performa yang lebih baik dalam mengelompokkan pola konsumsi listrik rumah tangga pada dataset yang digunakan.

Temuan ini konsisten dengan penelitian Rahmati et al. (2021) yang menyatakan bahwa K-Means lebih unggul pada data dengan struktur yang jelas dan homogen. Namun demikian, Hierarchical Clustering tetap memiliki keunggulan dalam memahami hubungan antar data secara lebih mendalam melalui struktur hierarkisnya.

Tabel 1 Hasil Evaluasi K-Means dan Hierarchical Clustering

Algoritma	Jumlah Cluster	Silhouette Coefficient	Davies-Bouldin Index
K-Means Clustering	4	0,67	0,48

Hierarchical Clustering	4	0,59	0,71
-------------------------	---	------	------

F. Implikasi Praktis Penelitian

Hasil penelitian ini memiliki beberapa implikasi praktis. Bagi penyedia energi atau perusahaan listrik, hasil clustering dapat dimanfaatkan untuk melakukan segmentasi pelanggan berdasarkan pola konsumsi listrik sehingga mendukung pengembangan tarif listrik yang lebih adaptif dan identifikasi konsumsi abnormal. Bagi pembuat kebijakan, hasil ini dapat menjadi dasar dalam merancang program efisiensi energi rumah tangga berbasis karakteristik konsumsi pelanggan. Sementara itu, dalam konteks pendidikan vokasi kelistrikan, penelitian ini dapat digunakan sebagai media pembelajaran berbasis data (*data-driven learning*) untuk meningkatkan pemahaman mahasiswa mengenai analisis konsumsi energi, pengolahan data kelistrikan, serta penerapan teknik *machine learning* pada sistem tenaga listrik

Kesimpulan

Bagian penelitian ini bertujuan untuk membandingkan performa algoritma K-Means dan Hierarchical Clustering dalam mengelompokkan pola konsumsi listrik rumah tangga berdasarkan parameter kelistrikan seperti *Global Active Power*, *Voltage*, dan *Current*. Berdasarkan hasil evaluasi kuantitatif menggunakan Silhouette Coefficient dan Davies-Bouldin Index (DBI), algoritma K-Means menunjukkan performa yang lebih baik dibandingkan Hierarchical Clustering, dengan nilai Silhouette Coefficient sebesar 0,67 dan Davies-Bouldin Index sebesar 0,48, sedangkan Hierarchical Clustering memperoleh nilai Silhouette Coefficient sebesar 0,59 dan Davies-Bouldin Index sebesar 0,71. Hasil tersebut menunjukkan bahwa K-Means lebih efektif dalam mengelompokkan data konsumsi listrik rumah tangga yang memiliki struktur relatif homogen, ditunjukkan oleh kualitas pemisahan kluster yang lebih baik dan tingkat overlap antarkluster yang lebih rendah.

Di sisi lain, Hierarchical Clustering tetap memiliki keunggulan dalam membentuk hubungan antardata secara hierarkis melalui visualisasi *dendrogram*, sehingga lebih sesuai digunakan ketika tujuan analisis adalah memahami keterkaitan antarkelompok data secara lebih mendalam atau ketika jumlah kluster belum diketahui secara pasti. Secara keseluruhan, pemilihan algoritma clustering sangat dipengaruhi oleh karakteristik data dan tujuan analisis yang diinginkan. Penelitian ini memberikan kontribusi dalam pemilihan metode clustering yang tepat untuk analisis konsumsi listrik rumah tangga serta dapat mendukung pengembangan pembelajaran analisis data kelistrikan berbasis *data-driven learning* pada pendidikan vokasi.

References

- Alamtaha, Z., Djakaria, I., & Yahya, N. I. (2023). *Implementasi Algoritma Hierarchical Clustering dan Non-Hierarchical Clustering untuk Pengelompokan Pengguna Media Sosial*. 4(1), 33–43. <https://doi.org/10.20956/ejsa.vi.24830>
- Anagra, F. (2020). Audit Energi dan Analisis Peluang Penghematan Konsumsi Energi Listrik di Unit 1 PLTU Banten 3 Lontar. *Jurnal Teknologi Elektro*, 11(1), 32. <https://doi.org/10.22441/jte.2020.v11i1.005>
- Brina, S., & Rizky, W. (2025). *PERBANDINGAN ALGORITMA K-MEANS DAN HIERARCHICAL CLUSTERING DALAM SEGMENTASI KABUPATEN / KOTA DI JAWA TIMUR BERDASARKAN DATA PERJALANAN DAN PERGERAKAN WISATAWAN*. 9(5), 8499–8506.
- Gianyar, D. (2024). *PENERAPAN ALGORITMA K-MEANS CLUSTERING*. 8(2), 1543–1547.
- Hendrawan, M. (2024). *SISTEM MONITORING KONSUMSI BEBAN LISTRIK RUMAH TANGGA REALTIME BERBASIS WEBSITE*. 15(2), 1–16.
- Herdiaman, E. A., Sudiarjo, A., Hikmatyar, M., Informatika, T., Perjuangan, U., & Barat, J. (2024).

RSUD MENGGUNAKAN METODE K-MEANS. 12(3).

- Impron, A. (2025). *Analisis Pola Konsumsi Energi Listrik Rumah Tangga Berbasis Simulasi IoT Menggunakan Model Hybrid LSTM-Attention Analysis of Household Electricity Consumption Patterns Based on IoT Simulation Using Hybrid LSTM-Attention Model. 1*, 361–369. <https://doi.org/10.26798/jiko.v9i1.1922>
- Jamilatun, S., Rhomadoni, F. R., Astuti, E., Setya, B., Idris, M., & Auliasari, P. A. (2025). *Peran Manajemen Energi terhadap Efisiensi Konsumsi Listrik Rumah Tangga di Indonesia*. 1–12.
- Rahmati, R., Wijayanto, A. W., Studi, P., Statistik, K., & Sains, P. (2021). *Rizqina Rahmati 1, Arie Wahyu Wijayanto 2 Program Studi Komputasi Statistik Peminatan Sains Data, Politeknik Statistika STIS. 5(2)*, 73–80.
- Rohmah, A., Sembiring, F., & Erfina, A. (2021). *IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING ANALYSIS UNTUK MENENTUKAN HAMBATAN PEMBELAJARAN DARING (STUDI KASUS : SMK YASPIM GEGERBITUNG)*. 290–298.
- Toba, F., Albert, V., Stevy, H., Indra, H., Fisika, S., Matematika, F., Ilmu, D., Alam, P., Sam, U., & Manado, R. (2023). *Analisis Perbandingan Daya Listrik saat Sebelum dan Sesudah Variasi Kapasitor pada Beban listrik Rumah Tangga. 13(1)*, 11–17.
- Yanuar, M. U. (2022). *Metode K-Means Clustering untuk Pengelompokan Kabupaten / Kota dalam Upaya Pengendalian Tingkat Inflasi di Pulau Jawa dan Sumatera K-Means Clustering Method for District / City Grouping in Effort to Control Inflation Rates in Java. 1*, 29–42.