



Analisis Pola Konsumsi Energi Listrik Pelanggan Rumah Tangga Menggunakan Alogaritma K-Means Clustering

Hilmi Mubarak^{1*}, Bambang Irawan²

^{1*,2}Program Studi Teknik Informatika, Universitas Muhadi Setiabudi, Indonesia

*Email : mubarokhilmi02@gmail.com

Abstract:

The increase in household electricity consumption is one of the main challenges in national energy management. Diverse electricity usage patterns are influenced by social, economic, and behavioral characteristics of consumers. This study aims to analyze and cluster household electricity consumption patterns using the K-Means Clustering algorithm. The dataset consists of secondary data from 1,200 household customers with attributes including installed power capacity, monthly electricity consumption (kWh), peak usage time, and average daily load. The research stages include data cleaning, normalization using StandardScaler, determination of the optimal number of clusters using the Elbow Method, clustering with K-Means, and evaluation using the Davies-Bouldin Index (DBI). The results indicate that the optimal number of clusters is three, representing low, medium, and high electricity consumption groups. A DBI value of 0.71 indicates good clustering quality. These findings can support electricity providers in designing energy efficiency policies and household load management strategies.

Keywords: Electricity Consumption, Data Mining, K-Means Clustering, Elbow Method, Davies-Bouldin Index

1. PENDAHULUAN

Energi listrik merupakan kebutuhan utama dalam kehidupan modern, terutama di sektor rumah tangga, karena hampir semua aktivitas sehari-hari bergantung pada energi listrik seperti penerangan, memasak, dan penggunaan perangkat elektronik. Dengan meningkatnya jumlah penduduk dan perkembangan teknologi, konsumsi listrik di rumah tangga terus meningkat setiap tahun, yang pada akhirnya menimbulkan tekanan besar terhadap sistem distribusi dan penyediaan energi nasional. Oleh karena itu, upaya efisiensi dan pengelolaan energi berbasis data menjadi sangat penting untuk menjamin ketersediaan listrik yang stabil dan berkelanjutan (Prayoga dkk., 2025)

Peningkatan konsumsi listrik yang terus menerus menunjukkan adanya variasi yang signifikan dalam pola pemakaian energi rumah tangga, yang dipengaruhi oleh berbagai faktor seperti kapasitas daya terpasang, jumlah penghuni rumah, jenis peralatan yang digunakan, serta kebiasaan harian pengguna. Heterogenitas pola konsumsi ini mengindikasikan bahwa setiap rumah tangga memiliki karakteristik tersendiri yang perlu dianalisis secara lebih detail agar distribusi energi dapat disesuaikan dengan kebutuhan nyata di lapangan (Silvia Lestari, 2025)

Oleh karena itu, diperlukan metode analisis yang mampu mengelompokkan pelanggan berdasarkan kemiripan karakteristik konsumsinya, sehingga pola-pola penting

dalam data besar dapat terungkap dengan jelas. Pendekatan klasifikasi seperti clustering memungkinkan pemetaan konsumsi listrik yang lebih akurat, sehingga perencana energi dapat membuat strategi yang sesuai untuk masing-masing segmen pelanggan berdasarkan perilaku konsumsi mereka (Rochfiani dkk., 2025)

Data mining merupakan salah satu pendekatan efektif dalam menemukan pola tersembunyi dalam dataset besar, karena metode ini mampu memproses dan menganalisis jutaan titik data tanpa memerlukan label awal. Teknik ini banyak digunakan di berbagai bidang, termasuk pada analisis perilaku konsumen, deteksi anomali, hingga optimasi proses distribusi energi, sehingga menjadi alat penting dalam pengelolaan sumber daya listrik (Mulyadi dkk., 2024)

Salah satu teknik data mining yang populer dan banyak digunakan adalah clustering, khususnya algoritma K-Means, yang bekerja dengan membagi data ke dalam beberapa kelompok berdasarkan kesamaan karakteristik. Teknik ini sangat berguna saat mengelompokkan data tanpa label atau *unsupervised*, sehingga insight dari distribusi konsumsi listrik yang kompleks dapat diperoleh dengan lebih mudah (Algoritma dkk., 2023)

Algoritma K-Means Clustering dipilih karena keunggulannya yang sederhana, efisien, dan mudah diimplementasikan dalam berbagai platform analisis data, sehingga banyak peneliti dan praktisi bidang energi menggunakannya untuk menganalisis pola konsumsi listrik rumah tangga. Algoritma ini dapat membantu mengidentifikasi segmen pelanggan dengan karakteristik konsumsi yang mirip, sehingga penyusun kebijakan energi dapat menargetkan strategi penghematan sesuai dengan kebutuhan masing-masing segmen. Selain mampu mengelompokkan konsumen, penerapan K-Means juga memungkinkan evaluasi kualitas hasil pengelompokan, misalnya dengan menggunakan indeks validasi internal seperti Davies-Bouldin Index (DBI). Indeks ini membantu menilai seberapa baik cluster terbentuk berdasarkan jarak antar pusat cluster dan jarak di dalam cluster, sehingga hasil analisis menjadi lebih dapat dipercaya dan objektif (Berliana & Irwansyah, 2025)

Melalui pendekatan clustering, pihak penyedia listrik seperti PLN maupun pemerintah dapat mengidentifikasi pola yang sulit terlihat hanya dari statistik deskriptif, seperti rumah tangga dengan konsumsi tinggi di malam hari atau selama musim tertentu. Informasi ini kemudian dapat dimanfaatkan untuk merancang program efisiensi energi, subsidi, atau tarif berbasis waktu guna mengurangi beban puncak dan meningkatkan stabilitas pasokan listrik.

Dengan demikian, penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering dalam menganalisis pola konsumsi energi listrik rumah tangga serta mengevaluasi hasil pengelompokan menggunakan Davies-Bouldin Index, sehingga dapat memberikan kontribusi empiris terhadap literatur analisis energi berbasis data dan mendukung pengambilan keputusan yang lebih efektif di bidang manajemen energi.

2. METODOLOGI PENELITIAN

2.1 *Data Set*

Dataset yang digunakan dalam penelitian ini merupakan data konsumsi energi listrik rumah tangga yang mencakup 1.200 pelanggan, dimana setiap entri mencatat empat atribut utama: daya terpasang (VA) yang menunjukkan kapasitas maksimal listrik yang tersedia bagi pelanggan, konsumsi energi bulanan (kWh) sebagai total energi yang terpakai setiap bulan, rata-rata beban harian (kWh/hari) yang menggambarkan pola konsumsi harian, serta persentase pemakaian pada jam beban puncak (%) yang merefleksikan intensitas penggunaan ketika permintaan maksimum terjadi; penggunaan atribut-atribut ini untuk karakterisasi konsumsi rumah tangga sejalan dengan praktik dalam literatur kepakaran energi, dimana parameter seperti konsumsi bulanan dan beban puncak diperlukan untuk analisis pola pemakaian dan peramalan beban listrik dengan akurasi tinggi dalam konteks model pembelajaran mesin atau statistik modern, karena variabel-variabel tersebut secara langsung mencerminkan perilaku konsumsi dan kapasitas jaringan distribusi listrik di tingkat mikro (mis. pelanggan individu) serta mendukung pemodelan beban harian dan puncak yang penting dalam perencanaan dan pengelolaan daya, termasuk perbandingan antara beban aktual dan kapasitas terpasang pada jam-jam kritis penggunaan energi (global energy databases dan dataset energi rumah tangga menunjukkan perlunya kombinasi data konsumsi dengan parameter listrik lainnya untuk analisis terperinci) (Tavarov dkk., 2025)

2.2 *Pre-Processing*

Tahapan *Pre-Processing*, bertujuan untuk mempersiapkan data mentah agar siap dianalisis. Proses ini meliputi dua tahap utama.

2. 2.1 *Data Cleaning*

Tahap data cleaning dalam penelitian ini dilakukan untuk menghapus data duplikat dan menangani nilai kosong (*missing value*) agar kualitas dataset yang digunakan optimal untuk analisis, karena data yang tidak lengkap atau redundan dapat menyebabkan bias, inkonsistensi, serta menurunkan keakuratan hasil analisis; dari 1.200 data awal, proses pembersihan mengidentifikasi 48 entri yang tidak lengkap dan kemudian menghapusnya, sehingga tersisa 1.152 data yang valid untuk dianalisis, sesuai dengan prosedur pembersihan data yang umum diterapkan dalam penelitian ilmiah di Indonesia saat ini yang mencakup deteksi dan penghapusan missing value serta duplikasi untuk memastikan dataset bebas dari data yang tidak relevan dan layak diproses lebih lanjut, seperti yang dijelaskan dalam studi-studi jurnal nasional terkini tentang data cleaning dan preprocessing data untuk analisis statistik dan pemodelan data (Suherman dkk., 2024)

2. 2.2 *Normalisasi Data*

Normalisasi dilakukan menggunakan metode *StandardScaler* untuk menyeragamkan skala data. Rumus normalisasi ditunjukkan pada Persamaan (1):

$$X' = \frac{(X - \mu)}{\sigma}$$

Dimana:

X' : Nilai data hasil normalisasi

X : Nilai data asli

μ : Nilai rata-rata

σ : Standar deviasi

Proses ini penting dalam pra-pemrosesan karena perbedaan skala antar atribut dapat memengaruhi kinerja model analitis atau *machine learning* yang sensitif terhadap variasi skala data, sehingga dengan *StandardScaler* model menjadi lebih stabil dan efisien dalam mempelajari pola dari data yang telah dinormalisasi, sebagaimana dijelaskan dalam literatur jurnal nasional yang menerapkan *StandardScaler* untuk memastikan fitur berada dalam skala konsisten demi meningkatkan akurasi dan efektivitas model dalam berbagai konteks aplikasi data mining dan klasifikasi (Studi Informatika dkk., 2024)

2.3 Penentuan Jumlah Cluster

Penentuan jumlah cluster optimal dalam clustering dengan algoritma K-Means dilakukan menggunakan *Elbow Method*, yakni dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk setiap kandidat jumlah cluster mulai dari $k = 1$ sampai $k = 6$, kemudian memplotkan nilai WCSS terhadap k untuk melihat pada titik mana penurunan WCSS mulai melambat secara signifikan; titik "*elbow*" tersebut menunjukkan jumlah cluster optimal karena penambahan cluster selanjutnya tidak memberikan pengurangan varian intra-cluster yang berarti, sehingga jumlah cluster yang dipilih bisa merepresentasikan struktur data dengan seimbang antara kompleksitas model dan kompaksi cluster, seperti yang diterapkan dalam berbagai penelitian nasional terkini yang menggunakan *Elbow Method* untuk mengevaluasi nilai WCSS sebagai indikator utama dalam menentukan k optimal sebelum melanjutkan ke proses pengelompokan lebih lanjut dalam konteks data mining dan pengolahan data (Vania & Nurina Sari, 2023) Rumus WCSS ditunjukkan pada Persamaan (2)

$$WCSS = \sum \sum ||x - c_i||^2$$

Dimana:

\sum (penjumlahan pertama) : Penjumlahan untuk seluruh cluster yang terbentuk.

\sum (penjumlahan kedua) : Penjumlahan untuk seluruh data x yang berada di dalam masing-masing cluster.

x : Vektor data atau objek ke- j yang menjadi anggota suatu cluster.

c_i : Pusat (centroid) dari cluster ke- i , yang diperoleh dari rata-rata seluruh data dalam cluster tersebut.

$\|x - c_i\|^2$: Jarak Euclidean kuadrat antara data x dan centroid c_i , yang menunjukkan seberapa jauh data tersebut dari pusat clusternya.

2.4 Proses *K-Means Clustering*

Algoritma K-Means diterapkan dengan jumlah cluster $k = 3$, dimana jarak antar setiap data dan pusat cluster (*centroid*) dihitung menggunakan metrik jarak Euclidean sebagaimana dinyatakan pada Persamaan (3)

$$d(X_i, C_k) = \sqrt{\sum (X_{ij} - C_{kj})^2}$$

Dimana:

$d(X_i, C_k)$: Jarak antara data X_i dan centroid C_k

X_{ij} : Nilai fitur ke- j dari data i

C_{kj} : Nilai fitur ke- j dari centroid k

Karena ukuran jarak ini mengukur kesamaan berdasarkan kuadrat selisih antara setiap fitur data dan koordinat centroid sehingga data yang paling dekat secara geometrik akan dikelompokkan bersama; perhitungan jarak *Euclidean* dalam K-Means adalah langkah kritis yang menentukan penugasan anggota cluster pada iterasi setiap langkah pembaruan centroid, dan pendekatan ini telah banyak digunakan dalam penelitian jurnal nasional terkini yang menerapkan algoritma K-Means untuk berbagai konteks aplikasi seperti segmentasi data, pengelompokan pola, dan analisis data empiris dengan model unsupervised learning untuk memastikan bahwa setiap objek berada di cluster dengan centroid yang paling representatif berdasarkan jarak Euclidean yang dihitung.

2.5 Evaluasi *Cluster*

Evaluasi kualitas cluster dilakukan menggunakan *Davies-Bouldin Index* (DBI) yang dirumuskan pada Persamaan (4)

$$DBI = \frac{1}{k} \sum \max \left(\frac{S_i + S_j}{d_{ij}} \right)$$

Dimana:

S_i dan S_j : Ukuran jarak rata-rata data terhadap centroid dalam satu *cluster*

d_{ij} : Jarak antara *centroid cluster* i dan j

DBI merupakan metrik evaluasi internal yang menilai kualitas hasil pengelompokan berdasarkan keseimbangan antara *intra-cluster cohesion* (kepadatan anggota dalam cluster) dan *inter-cluster separation* (pemisahan antar cluster), sehingga nilai DBI yang lebih kecil menunjukkan cluster yang lebih kompak dan jelas terpisah satu sama lain, yang berarti pengelompokan lebih baik dan valid dalam konteks analisis data tanpa label eksternal sebagaimana dibahas dalam berbagai penelitian jurnal nasional tentang evaluasi clustering menggunakan DBI (Kadafi & Finandhita, 2024)

3. HASIL DAN PEMBAHASAN

3.1. Data Set

Data set yang digunakan dalam penelitian ini merupakan data sekunder konsumsi energi listrik pelanggan rumah tangga yang berjumlah 1.200 pelanggan. Data ini merepresentasikan karakteristik penggunaan energi listrik rumah tangga dalam periode satu tahun. Setiap data terdiri dari empat atribut utama yang digunakan sebagai dasar proses clustering.

Tabel 1. Struktur Dataset Konsumsi Energi Listrik Rumah Tangga

| No | Nama Atribut | Simbol | Satuan | Keterangan |
|----|----------------------------|--------|----------|--|
| 1 | Daya Terpasang | X1 | VA | Besarnya kapasitas daya listrik yang terpasang pada pelanggan rumah tangga |
| 2 | Konsumsi Energi Bulanan | X2 | kWh | Total energi listrik yang dikonsumsi pelanggan dalam satu bulan |
| 3 | Beban Harian Rata-Rata | X3 | kWh/hari | Rata-rata konsumsi energi listrik pelanggan per hari |
| 4 | Pemakaian Jam Beban Puncak | X4 | % | Persentase penggunaan listrik pada jam beban puncak (WBP) |

Keempat atribut tersebut dipilih karena dinilai mampu merepresentasikan pola konsumsi energi listrik rumah tangga secara menyeluruh, baik dari sisi kapasitas, intensitas pemakaian, maupun perilaku penggunaan listrik pada waktu tertentu. Seluruh atribut bersifat numerik sehingga sesuai untuk dianalisis menggunakan algoritma K-Means Clustering.

3.2. Pre-Processing Data

Tahap *pre-processing* data dilakukan untuk memastikan kualitas data sebelum diterapkan algoritma *K-Means Clustering*. Proses ini sangat penting karena K-Means sensitif terhadap perbedaan skala data dan keberadaan data yang tidak valid. Tahapan *pre-processing* dalam penelitian ini meliputi data cleaning dan normalisasi data.

Data cleaning bertujuan untuk menghilangkan data yang tidak valid, seperti data duplikat dan data yang memiliki nilai kosong (*missing value*). Dari total 1.200 data awal, ditemukan sebanyak 48 data yang memiliki nilai tidak lengkap pada salah satu atribut, khususnya pada atribut konsumsi energi bulanan dan beban harian. Data-data tersebut dihapus agar tidak mempengaruhi proses perhitungan jarak pada algoritma K-Means. Setelah proses data cleaning, jumlah data yang digunakan dalam penelitian ini menjadi 1.152 data.

Tabel 2. Contoh Data Sebelum Data Cleaning

| No | Daya Terpasang (VA) | Konsumsi Bulanan (kWh) | Beban Harian (kWh/hari) | Jam Puncak (%) |
|----|---------------------|------------------------|-------------------------|----------------|
| 1 | 1300 | 145 | 4.8 | 35 |
| 2 | 900 | - | 3.2 | 28 |
| 3 | 2200 | 310 | 10.5 | - |
| 4 | 1300 | 120 | 4.0 | 30 |

Data pada baris ke-2 dan ke-3 mengandung nilai kosong, sehingga tidak digunakan dalam proses analisis.

Setelah data dibersihkan, langkah berikutnya adalah normalisasi data menggunakan metode *StandardScaler*. Normalisasi ini bertujuan untuk menyeragamkan skala antar atribut agar tidak terjadi dominasi atribut tertentu dalam perhitungan jarak Euclidean. Metode *StandardScaler* mengubah data sehingga memiliki nilai rata-rata 0 dan standar deviasi 1.

Rumus normalisasi ditunjukkan pada Persamaan (1):

$$X' = \frac{(X - \mu)}{\sigma}$$

Dilakukan normalisasi pada atribut konsumsi energi bulanan. Berdasarkan data setelah cleaning, diperoleh nilai rata-rata (μ) sebesar 180 kWh dan standar deviasi (σ) sebesar 65 kWh. Jika salah satu pelanggan memiliki konsumsi energi bulanan sebesar 310 kWh, maka perhitungannya adalah sebagai berikut:

$$X' = \frac{(X - \mu)}{\sigma}$$

$$X' = \frac{(310 - 180)}{65}$$

$$X' = 2,00$$

Hasil tersebut menunjukkan bahwa konsumsi energi pelanggan tersebut berada dua standar deviasi di atas rata-rata populasi.

Tabel 3. Data Sebelum dan Sesudah Normalisasi

| No | Atribut | Nilai Asli | Nilai Normalisasi |
|----|-------------------------------|------------|-------------------|
| 1 | Daya Terpasang (VA) | 1300 | -0,45 |
| 2 | Konsumsi Energi Bulanan (kWh) | 310 | 2,00 |
| 3 | Beban Harian (kWh/hari) | 10.5 | 1,85 |
| 4 | Jam Puncak | 42 | 1,12 |

Hasil normalisasi pada Tabel 3 menunjukkan bahwa seluruh atribut telah berada pada skala yang sama, sehingga siap digunakan dalam proses clustering menggunakan algoritma K-Means.

3.3. Cluster

Penentuan jumlah cluster optimal dilakukan menggunakan Elbow Method dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS) untuk k = 1 hingga k = 6. Rumus WCSS ditunjukkan pada Persamaan (2):

$$WCSS = \sum \sum ||x - c_i||^2$$

Hasil perhitungan nilai WCSS untuk setiap nilai k disajikan pada Tabel 4.

Tabel 4. Hasil Perhitungan WCSS

| Cluster (k) | Nilai WCSS |
|-------------|------------|
| 1 | 4612.35 |
| 2 | 2894.17 |
| 3 | 1726.42 |
| 4 | 1602.88 |
| 5 | 1521.64 |
| 6 | 1480.27 |

Berdasarkan Tabel 4, terlihat penurunan WCSS yang signifikan dari $k = 1$ hingga $k = 3$, kemudian mulai melandai setelahnya. Hal ini menunjukkan bahwa jumlah cluster optimal berada pada $k = 3$.

3.4. Proses K-Means Clustering

Algoritma K-Means diterapkan dengan jumlah cluster $k = 3$, dimana jarak antar setiap data dan pusat cluster (*centroid*) dihitung menggunakan metrik jarak *Euclidean* sebagaimana dinyatakan pada Persamaan (3)

$$d(X_i, C_k) = \sqrt{\sum (X_{ij} - C_{kj})^2}$$

Proses iterasi dilakukan hingga posisi centroid konvergen. Hasil akhir berupa nilai centroid dari masing-masing cluster disajikan pada Tabel 5.

Tabel 5. Hasil Perhitungan Centroid Akhir

| Cluster | Daya Terpasang (VA) | Konsumsi Bulanan (kWh) | Beban Harian (kWh) | Jam Puncak (%) |
|-------------|---------------------|------------------------|--------------------|----------------|
| C1 (Rendah) | -0,82 | -0,91 | -0,88 | -0,79 |
| C2 (Sedang) | 0,12 | 0,08 | 0,15 | 0,19 |
| C3 (Tinggi) | 1,05 | 1,12 | 1,03 | 1,97 |

Nilai centroid pada Tabel 5 merupakan hasil normalisasi menggunakan *Stabdar Scaler* sehingga menunjukkan posisi relative masing-masing cluster terhadap rata-rata populasi data.

3.5. Evaluasi Cluster

Untuk mengevaluasi kualitas hasil clustering, digunakan *Davies-Bouldin Index* (DBI). Hasil perhitungan DBI menghasilkan nilai sebesar 0,71. Nilai ini menunjukkan bahwa cluster yang terbentuk memiliki tingkat kekompakan internal yang baik serta pemisahan antar cluster yang cukup jelas.

Semakin kecil nilai DBI, semakin baik kualitas clustering, dan nilai di bawah 1 menunjukkan bahwa struktur cluster dapat dikategorikan baik. Dengan demikian, hasil evaluasi ini mengonfirmasi bahwa algoritma K-Means mampu mengelompokkan pola konsumsi energi listrik pelanggan rumah tangga secara efektif dan konsisten dengan karakteristik data.

4. KESIMPULAN

Berdasarkan hasil perhitungan dan analisis yang telah dilakukan, dapat disimpulkan bahwa algoritma K-Means Clustering mampu mengelompokkan pola konsumsi energi

listrik pelanggan rumah tangga secara efektif ke dalam tiga cluster utama, yaitu konsumsi rendah, sedang, dan tinggi. Tahap pre-processing yang meliputi data cleaning dan normalisasi menggunakan StandardScaler terbukti berperan penting dalam meningkatkan kualitas data sehingga perhitungan jarak Euclidean dapat dilakukan secara optimal. Penentuan jumlah cluster menggunakan Elbow Method menunjukkan nilai optimal pada $k = 3$, yang selanjutnya diperkuat oleh hasil evaluasi menggunakan Davies-Bouldin Index (DBI) dengan nilai sebesar 0,71. Nilai DBI tersebut menunjukkan bahwa cluster yang terbentuk memiliki tingkat kekompakan internal yang baik serta pemisahan antar cluster yang cukup jelas, sehingga hasil pengelompokan dapat merepresentasikan karakteristik konsumsi energi listrik rumah tangga secara akurat dan berbasis data. Kesimpulan harus mengindikasikan secara jelas hasil-hasil yang diperoleh, kelebihan dan kekurangannya, serta kemungkinan pengembangan selanjutnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dalam pelaksanaan penelitian ini. Ucapan terima kasih disampaikan kepada pihak institusi pendidikan yang telah menyediakan fasilitas dan lingkungan akademik yang mendukung, serta kepada dosen pembimbing yang telah memberikan arahan, bimbingan, dan masukan yang konstruktif selama proses penelitian berlangsung. Selain itu, penulis juga mengapresiasi pihak penyedia data dan referensi ilmiah yang digunakan dalam penelitian ini. Dukungan dan kontribusi dari berbagai pihak tersebut sangat berperan dalam tersusunnya penelitian ini hingga dapat diselesaikan dengan baik.

REFERENSI

- Algoritma, I., Pada, K.-M., Listrik, B., Tangga, R., Nur, A., Maulidhia, A., Widyastuti, I. I., Sukarno, F. I., Basya, R., Tsany, S., Brian, T., Teknik, P., Kapal, K., Perkapalan, P., Surabaya, N., Kapal, B., Bisnis, P. M., & Perpipa-an, P. T. (n.d.). *Implementation of the K-Means Algorithm on Household Electric Load*. <https://doi.org/10.32503/jtecs.v5i1.6739>
- Berliana, Y. L., & Irwansyah, I. (2025). Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Konsumsi Listrik Menurut Provinsi di Indonesia. *Progresif: Jurnal Ilmiah Komputer*, 21(2), 599. <https://doi.org/10.35889/progresif.v21i2.2876>
- Kadafi, M. N., & Finandhita, A. (2024). *PENERAPAN DATA MINING DALAM PENGELOMPOKAN PENYAKIT UNTUK REKOMENDASI MATERI PENYULUHAN KESEHATAN DI KLINIK KELUARGA CIANJUR*. 13(1).
- Mulyadi, S., Insani, F., Agustian, S., & Afriyanti, L. (2024). Pengelompokan Data Pendistribusian Listrik Menggunakan Algoritma Mini Batch K-Means Clustering.

-
- MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 1051–1062. <https://doi.org/10.57152/malcom.v4i3.1425>
- Prayoga, F., Aulia Meliala, N., Aulia, D., Fi, S., Mar'ah, M., Bima, H., Alam, P., & Furqan, M. (2025). *Journal of Artificial Intelligence and Engineering Applications Utilization of Data Mining in Finding Household Electricity Consumption Patterns in Indonesia* (Vol. 4, Number 3). <https://ioinformatic.org/>
- Rochfiani, P., Chulkamdi, M. T., & Mawaddah, U. (2025). Comparative Study Of Monthly Electricity Consumption Clusterization Using K-Means and DBSCAN. *JOSAR*, 10(2), 2503–1155. <https://doi.org/10.35457/josar.v10i2.5232>
- Silvia Lestari. (2025). Penggunaan Metode K-Means dalam Data Mining untuk Identifikasi Pola Konsumsi Listrik pada Rumah Tangga. *Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer*, 5(2), 88–103. <https://doi.org/10.55606/juisik.v5i2.1165>
- Studi Informatika, P., Matematika dan Ilmu Pengetahuan Alam, F., Raya Kampus UNUD, J., Jimbaran, B., & Selatan, K. (2024). Klasifikasi Genre Musik Menggunakan Metode Support Vector Machine Dengan Multi-Kernel I Gusti Agung Istri Agrivina Shyta Devi a1 , I Made Widiartha a2. *JNATIA*, 3(1). <https://www.kaggle.com>
- Suherman, Arsita, A. D., & Afriantoro, I. (2024). Analisis Pola Penjualan di Toko Retail Menggunakan Algoritma Apriori dengan Rapid Minner. *Jurnal Teknologi Informatika Dan Komputer*, 10(2), 734–753. <https://doi.org/10.37012/jtik.v10i2.2407>
- Tavarov, S., Sidorov, A., & Glotova, N. (2025). Forecasting Average Daily and Peak Electrical Load Based on Average Monthly Electricity Consumption Data. *Electricity*, 6(2). <https://doi.org/10.3390/electricity6020026>
- Vania, P., & Nurina Sari, B. (2023). Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Klaster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means. *Jurnal Ilmiah Wahana Pendidikan*, 9(21), 547–558. <https://doi.org/10.5281/zenodo.10081332>