

Penerapan Algoritma K-Means untuk Pemetaan Wilayah Penyebaran Stasiun Pengisian Kendaraan Listrik (SPKLU) di Kota Jakarta

Amung Sunarya^{1*}, Irmayansyah²

^{1,2}Program Studi Sistem Informasi, Fakultas Informatika dan Komputer, Universitas Binaniaga Indonesia

¹Email: amungsunarya4793@gmail.com

²Email: irma@unbin.ac.id

*Corresponding Author

ABSTRACT

The adoption of electric vehicles (EVs) has shown significant growth in many countries, including Indonesia. However, this increase has not been fully matched by the availability and optimal distribution of Public Electric Vehicle Charging Stations (SPKLUs). This condition is reflected in the existence of SPKLUs operating at full capacity, while others remain underutilized, indicating inefficiencies in infrastructure planning. This study aims to apply the K-Means Clustering algorithm to analyze and map the distribution of SPKLUs in Jakarta. The dataset was obtained from an SPKLU developer company, covering the period from January 2021 to April 2025, and includes five key attributes: installed power capacity (kW), energy consumption (kWh), number of transactions, number of chargers, and number of connectors. The research methodology follows a Research and Development (R&D) approach with a prototyping model. The initial stage involved normalizing the data using the Robust Scaler method, followed by clustering using the K-Means algorithm. The results show that the distribution of SPKLUs can be categorized into three clusters: Cluster 0 (Low) with 180 SPKLUs (82.2%), Cluster 1 (Medium) with 35 SPKLUs (16%), and Cluster 2 (High) with 4 SPKLUs (1.8%). Validation using the Silhouette Coefficient produced an average score of 0.66, indicating that the clustering structure is at a good level. Furthermore, this study developed a prototype web-based application implementing the K-Means method to visualize the clustering results. In conclusion, the findings demonstrate that the K-Means clustering approach can serve as a decision-making basis for the distribution planning of SPKLUs in Jakarta.

Keywords: K-Means Algorithm, SPKLU Mapping, Prototype, Robust Scaler, Silhouette Coefficient.

ABSTRAK

Adopsi kendaraan listrik (EV) menunjukkan peningkatan yang signifikan di berbagai negara, termasuk Indonesia. Namun demikian, perkembangan jumlah EV tersebut belum sepenuhnya diimbangi dengan ketersediaan serta distribusi Stasiun Pengisian Kendaraan Listrik Umum (SPKLU) yang optimal. Kondisi ini ditunjukkan oleh adanya SPKLU yang beroperasi pada kapasitas penuh, sementara sebagian lainnya justru minim dimanfaatkan, sehingga mencerminkan adanya ketidakefisienan dalam perencanaan infrastruktur. Penelitian ini bertujuan untuk menerapkan algoritma K-Means Clustering dalam menganalisis sekaligus memetakan distribusi SPKLU di wilayah Jakarta. Dataset yang digunakan diperoleh dari perusahaan pengembang SPKLU dengan cakupan periode Januari 2021 hingga April 2025, yang mencakup lima atribut utama, yaitu kapasitas daya terpasang (kW), energi terpakai (kWh), jumlah transaksi, jumlah unit pengisi daya (charger), serta jumlah konektor. Metodologi yang digunakan ialah Research and Development dengan pendekatan prototyping. Tahap penerapan awal diawali dengan data yang di normalisasi menggunakan Robust Scaller, kemudian dilanjutkan dengan proses pemetaan menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa distribusi SPKLU dapat dikategorikan ke dalam tiga klaster, yakni hasil yang didapat pada klaster 0 (Rendah) dengan 180 SPKLU (82,2%), Klaster 1 (Sedang) dengan 35 SPKLU (16%) dan Klaster 2 (Tinggi) dengan 4 SPKLU (1,8%). Validasi hasil klasterisasi menggunakan Silhouette Coefficient menghasilkan skor rata-rata sebesar 0,66, yang mengindikasikan struktur klaster berada pada tingkat baik. Selanjutnya, penelitian ini mengembangkan sebuah prototype aplikasi dengan penerapan metode K-Means berbasis web untuk memvisualisasikan hasil klaster. Dengan demikian, penelitian ini menegaskan bahwa pendekatan dengan metode K-Means clustering dapat digunakan sebagai dasar pengambilan keputusan untuk penyebaran wilayah stasiun pengisian kendaraan listrik umum (SPKLU) di kota Jakarta.

Kata kunci: Algoritma K-Means, Pemetaan SPKLU, Prototipe, Robust Scaller, Silhouette Coefficient.

A. PENDAHULUAN

1. Latar Belakang

Dalam beberapa dekade terakhir, sektor transportasi mengalami percepatan teknologi yang signifikan, khususnya dalam upaya mengurangi dampak lingkungan. Kendaraan berbahan bakar fosil semakin dipandang sebagai permasalahan krusial karena kontribusinya terhadap polusi udara, ketergantungan pada sumber daya tak terbarukan, serta emisi gas rumah kaca yang mempercepat perubahan iklim [1]. Seiring dengan meningkatnya kesadaran global terhadap keberlanjutan lingkungan, kendaraan listrik (EV) muncul sebagai alternatif yang lebih bersih dengan menawarkan efisiensi energi dan emisi yang lebih rendah dibandingkan kendaraan berbasis fosil.

Indonesia juga mengadopsi EV sebagai bagian dari strategi transisi energi. Jumlah EV, khususnya sepeda motor listrik, meningkat secara signifikan dalam beberapa tahun terakhir [2]. Pemerintah turut mendukung tren ini melalui penerbitan regulasi, pemberian insentif bagi produsen, serta subsidi kepada konsumen. Namun, tantangan utama terletak pada ketersediaan infrastruktur yang memadai, terutama Stasiun Pengisian Kendaraan Listrik Umum (SPKLU). Distribusi SPKLU yang belum merata berpotensi menghambat adopsi serta mengurangi efisiensi investasi [3, p. 19]. Meskipun jumlah SPKLU terus mengalami peningkatan (BPS, 2020 dalam Sugianto, 2022, p.1), distribusinya masih belum optimal.

Untuk menjawab permasalahan ini, diperlukan pendekatan berbasis data guna mengoptimalkan pemetaan wilayah penyebaran SPKLU. Salah satu metode yang digunakan ialah algoritma K-Means clustering, yang mampu mengelompokkan SPKLU berdasarkan pola penggunaan dan data operasional [5, p. 156]. merujuk pada penelitian sebelumnya menggunakan algoritma K Means mengenai "Penerapan Algoritma K-Means Untuk Pemetaan Penerimaan Bantuan kesejahteraan Masyarakat di Kota Batam" [6], menunjukkan tingkat keberhasilan penggunaan algoritma K-Means dalam memetakan prioritas penerima bantuan di kota Batam. Dalam mengevaluasi kualitas klusterisasi, digunakan Silhouette Coefficient yang mengukur tingkat kohesi dalam kluster serta pemisahan antar-kluster, di mana nilai yang mendekati 1 menunjukkan struktur kluster yang baik [7, p. 56]. Kombinasi pendekatan ini diharapkan mampu mendukung perencanaan infrastruktur yang lebih efisien sekaligus mempercepat adopsi EV di Indonesia.

2. Permasalahan

Saat ini, penentuan lokasi pembangunan stasiun pengisian kendaraan listrik umum (SPKLU) masih banyak bergantung pada metode konvensional, misalnya kedekatan dengan pusat perbelanjaan atau jalur transportasi utama, tanpa memperhatikan analisis data yang lebih komprehensif mengenai pola penggunaan kendaraan listrik. Kondisi tersebut menyebabkan terjadinya ketimpangan pemanfaatan, di mana beberapa titik SPKLU mengalami antrean panjang dengan beban tinggi, sementara titik lainnya justru kurang digunakan secara optimal. Ketidakerataan ini bukan hanya menimbulkan ketidaknyamanan bagi pengguna, tetapi juga berpotensi memperlambat percepatan adopsi kendaraan listrik di masyarakat. Berdasarkan data penggunaan SPKLU di wilayah Jakarta, terlihat adanya ketimpangan yang cukup signifikan antara stasiun dengan tingkat aktivitas yang sangat tinggi dan masif digunakan dengan stasiun lain yang justru sama sekali tidak dimanfaatkan. SPKLU yang strategis dan aktif digunakan memberi manfaat ekonomi serta mendukung percepatan adopsi kendaraan listrik dan pengurangan emisi. Sebaliknya, SPKLU yang tidak termanfaatkan menimbulkan pemborosan investasi dan biaya operasional, menurunkan efisiensi sumber daya, menimbulkan ketidakpuasan pengguna, serta memperlambat perkembangan ekosistem kendaraan listrik. Penerapan metode K-Means Clustering diperlukan untuk mengelompokkan wilayah berdasarkan pola penggunaan SPKLU. Hasil analisis ini dapat menjadi dasar perencanaan infrastruktur yang lebih efisien, tepat sasaran, dan berkelanjutan. Berdasarkan hal tersebut maka dilakukanlah penelitian ini yang berjudul "Penerapan Algoritma K-means Untuk Pemetaan Wilayah Penyebaran Stasiun Pengisian Kendaraan Listrik (SPKLU) Di Kota Jakarta".

3. Tujuan

Adapun tujuannya dari penelitian ini, yaitu:

- a. Mendapatkan hasil pemetaan lokasi SPKLU yang lebih akurat.
- b. Mendapatkan proses yang lebih efektif pada pemetaan wilayah penyebaran SPKLU di Jakarta.
- c. Mengukur tingkat keakuratan dan keefektifitasan penerapan algoritma K-Means pada pemetaan wilayah penyebaran SPKLU di Jakarta.
- d. Membuat prototype aplikasi penerapan metode K-Means untuk pemetaan wilayah penyebaran SPKLU.

4. Tinjauan Pustaka

a. Data Mining

Data Mining merupakan salah satu cabang ilmu pada bidang statistika dan ilmu komputer yang berfokus pada proses sistematis untuk mengekstraksi pengetahuan dari kumpulan data dalam jumlah besar. Proses ini dilakukan dengan memanfaatkan metode dan algoritma cerdas yang terstruktur secara komprehensif. Data yang telah diolah melalui teknik data mining tidak hanya menjadi sumber informasi baru, tetapi juga dapat dimanfaatkan sebagai dasar dalam menganalisis perkembangan teknologi maupun sebagai bahan pertimbangan dalam pengambilan keputusan di perusahaan maupun instansi lainnya [8, p. 20]. Bagian ini merupakan tahapan di mana data yang telah dikumpulkan kemudian dimasukkan ke dalam aplikasi data mining untuk dianalisis lebih lanjut, sehingga dapat ditemukan pola-pola tersembunyi dari kumpulan data yang besar.

Proses *data mining* dilakukan dengan memanfaatkan beragam teknik statistik, matematis, dan kecerdasan buatan untuk menganalisis data secara sistematis dan otomatis. Melalui proses ini, informasi yang diperoleh dapat dimanfaatkan untuk mendukung pengambilan keputusan, mengenali pola atau tren pasar, meningkatkan efektivitas operasional, hingga merumuskan strategi bisnis yang lebih tepat. Tujuan utama *data mining* adalah menggali pengetahuan atau wawasan yang tidak dapat diperoleh hanya melalui pengamatan langsung terhadap data mentah [9, p. 2].

b. CRISP-DM (Cross-Industry Standard Process for Data Mining)

Metode CRISP-DM (Cross-Industry Standard Process for Data Mining) merupakan sebuah kerangka kerja atau metodologi yang banyak digunakan dalam bidang data mining dan analisis data. Kerangka ini memberikan panduan sistematis bagi para profesional untuk menjalankan proyek data mining mulai dari tahap perencanaan hingga implementasi. CRISP-DM dikenal sebagai metodologi yang terstruktur dan komprehensif, sehingga mampu meningkatkan konsistensi serta efektivitas dalam proses analisis data lintas industri [10, p. 35].

CRISP-DM terdiri dari enam tahap utama, yaitu: [10, p. 36]

- 1) **Pemahaman Bisnis (Business Understanding):** Tahap ini berfokus pada pemahaman yang mendalam mengenai permasalahan yang akan diselesaikan melalui proses data mining. Pada tahap ini dilakukan interaksi dengan pemangku kepentingan guna merumuskan tujuan bisnis, menentukan target yang ingin dicapai, serta mengidentifikasi berbagai kendala yang mungkin memengaruhi jalannya proses analisis.
- 2) **Pemahaman Data (Data Understanding):** Tahap ini menekankan pada proses pengumpulan serta pemahaman data yang relevan untuk mendukung proyek data mining. Kegiatan yang dilakukan mencakup identifikasi dan eksploitasi berbagai sumber data yang tersedia, eksplorasi awal untuk mengetahui kondisi data, serta analisis karakteristik utama dari data tersebut. Melalui tahap ini, peneliti dapat memperoleh gambaran awal mengenai kualitas, kelengkapan, serta potensi permasalahan yang terdapat pada data sebelum masuk ke tahap pengolahan lebih lanjut.
- 3) **Persiapan Data (Data Preparation):** Tahap ini merupakan bagian penting dalam metodologi CRISP-DM yang dikenal sebagai data preparation. Aktivitas pada tahap ini mencakup pembersihan data untuk menghilangkan noise atau data yang tidak relevan, transformasi data agar sesuai dengan format analisis yang diinginkan, serta penggabungan data dari berbagai sumber jika diperlukan. Selain itu, dilakukan pula proses pelabelan atau pengkodean data agar dapat diinterpretasikan dengan baik oleh algoritma data mining. Dengan adanya tahap ini, kualitas data yang digunakan dalam pemodelan dapat terjamin sehingga hasil analisis yang diperoleh menjadi lebih akurat dan reliabel.
- 4) **Modeling:** Pada tahap ini dilakukan pemilihan metode atau algoritma data mining yang paling sesuai dengan tujuan analisis. Proses ini mencakup pembangunan dan pengembangan model analisis data menggunakan teknik tertentu, seperti regresi, klasifikasi, klusterisasi, maupun analisis asosiasi. Pemilihan teknik sangat bergantung pada jenis masalah yang dihadapi serta tujuan akhir dari proyek. Tahap ini juga dapat melibatkan pengaturan parameter dan pengujian awal model agar diperoleh performa yang optimal.
- 5) **Evaluasi (Evaluation):** Setelah model dikembangkan, langkah berikutnya adalah melakukan pengujian dan evaluasi untuk memastikan bahwa model tersebut sesuai dengan tujuan bisnis serta memiliki kinerja yang optimal. Proses evaluasi dilakukan dengan menggunakan teknik statistik maupun metode validasi, seperti pengukuran akurasi, presisi, recall, atau koefisien evaluasi lainnya. Tahap evaluasi ini bertujuan untuk menilai sejauh mana model mampu menjawab permasalahan yang telah dirumuskan pada tahap awal, sekaligus menentukan apakah model tersebut layak untuk dilanjutkan ke tahap implementasi atau masih memerlukan penyempurnaan lebih lanjut.

- 6) Implementasi (Deployment): Tahap terakhir adalah deployment, yaitu mengimplementasikan model ke dalam lingkungan nyata sebagai bagian dari aktivitas organisasi atau perusahaan. Pada tahap ini, model yang telah diuji dan dievaluasi diterapkan agar dapat digunakan secara operasional. Proses implementasi dapat mencakup pengkodean model ke dalam aplikasi, integrasi dengan sistem yang sudah ada, serta pelatihan personel yang akan menggunakan atau mengelola model tersebut. Tujuan utama dari tahap ini adalah memastikan bahwa hasil data mining dapat memberikan nilai nyata bagi organisasi dan mendukung pengambilan keputusan secara efektif.

B. METODE

1. Algoritma K-Means

Algoritma K-Means bekerja dengan membagi data ke dalam beberapa kluster berdasarkan kedekatan atau kesamaan karakteristik. Algoritma ini menggunakan pendekatan iteratif untuk menentukan posisi pusat kluster (centroid) yang optimal. Tujuan dari proses ini adalah meminimalkan variasi di dalam setiap kluster sekaligus memaksimalkan perbedaan antar kluster, sehingga pembagian data menjadi lebih representatif dan bermakna.

Secara definisi, K-Means merupakan algoritma klusterisasi yang mengelompokkan data berdasarkan kedekatannya dengan titik pusat kluster (centroid) terdekat. Dalam konteks pemetaan SPKLU, algoritma K-Means dapat dimanfaatkan untuk mengelompokkan lokasi SPKLU berdasarkan data transaksi maupun data aktual, sehingga mampu mendukung proses pengambilan keputusan yang lebih tepat dalam pengembangan infrastruktur kendaraan listrik. Mekanisme kerja algoritma ini meliputi beberapa tahapan, yaitu: menentukan jumlah kluster (K) yang akan dibentuk, memilih secara acak K titik awal sebagai pusat kluster (centroid), menghitung jarak setiap data terhadap masing-masing centroid kemudian mengelompokkan data ke dalam kluster terdekat, memperbarui posisi centroid berdasarkan rata-rata data pada setiap kluster, serta mengulangi proses perhitungan jarak dan pembaruan centroid hingga posisi centroid tidak lagi berubah atau jumlah iterasi yang ditentukan telah tercapai.

Langkah-langkah utama dalam algoritma K-Means adalah sebagai berikut:

- a. Menentukan jumlah kluster K yang diinginkan.
- b. Memilih K titik sebagai centroid awal secara acak.

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j$$

- c. Menghitung jarak setiap data ke centroid terdekat menggunakan metrik seperti *Euclidean Distance*:

$$d(x, \mu) = \sqrt{\sum_{i=1}^n (x_i - \mu_i)^2}$$

- d. Mengelompokkan setiap data ke kluster dengan centroid terdekat.
- e. Menghitung ulang posisi centroid berdasarkan rata-rata data dalam masing-masing kluster:

$$C_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i$$

di mana C_j adalah centroid baru kluster j , dan S_j adalah himpunan data dalam kluster j .

- f. Mengulangi langkah 3-5 hingga konvergensi tercapai (centroid tidak berubah secara signifikan atau jumlah iterasi maksimum tercapai [9, p. 126].

2. Teknik Analisis Data

Metode pengujian yang diterapkan dalam penelitian ini adalah metode Silhouette Coefficient. Metode ini akan menguji kualitas dan kekuatan suatu kluster, serta seberapa baik setiap objek ditempatkan dalam kluster tersebut. Metode ini dipilih karena mampu mengukur ketepatan pembagian kluster, sehingga sangat sesuai

digunakan sebagai alat evaluasi hasil clustering menggunakan algoritma K-Means. Nilai Silhouette Coefficient yang diperoleh dapat dilihat pada Tabel 1.

Tabel 1 Kategori Nilai Silhouette Coefficient

Nilai Silhouette Coefficient	Keterangan
Dengan rentang $0.71 < SC \leq 1$	Struktur Kuat
Dengan rentang $0.51 < SC \leq 0.7$	Struktur Baik
Dengan rentang $0.25 < SC \leq 0.5$	Struktur Lemah
Dengan rentang $SC \leq 0.25$	Tidak Terstruktur

Sumber: [7, p. 55]

Menurut (Rousseeuw, 1987, p. 55) berikut tahapan perhitungan Silhouette Coefficient;

- menghitung rata-rata jarak dari suatu data dengan permisalan i dengan semua data lain yang berada dalam satu klaster;

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j)$$

dengan j adalah data lain dalam suatu klaster A dan $d(i, j)$ merupakan jarak antara data i dengan j ;

- menghitung rata-rata jarak dari data i tersebut dengan semua data pada klaster lain dan diambil nilai terkecilnya;

$$b(i, C) = \frac{1}{|A|} \sum_{j \in C} d(i, j)$$

dengan $d(i, j)$ merupakan jarak rata-rata data i dengan semua objek pada klaster lain C dimana $A \neq C$;
 $b(i) = \min_{C \neq A} d(i, C)$

- dengan rumus Silhouette Coefficient sebagai berikut;

$$s(i) = \frac{(b(i) - a(i))}{\max(a(i), b(i))}$$

dimana $s(i)$ merupakan semua rata-rata pada semua data.

C. HASIL DAN PEMBAHASAN

1. Hasil

- Menyiapkan Dataset

Data yang digunakan merupakan data milik perusahaan pengembang SPKLU pada bulan Januari 2021 sampai Maret 2025. Dataset terdiri dari 219 baris data, 5 variabel dan disajikan pada Tabel 2.

Tabel 2 Data SPKLU

No	Nama	Kapasitas (Kw)	Jumlah KWH /Daya	Jumlah Transaksi	Jumlah Charger	Jumlah Connector
1	SPKLU PLN UID JAKARTA RAYA	200	810624	32640	4	10
2	SPKLU PLN UP3 Bulungan	200	399048	13780	2	5
3	SPKLU PLN UP3 LENTENG AGUNG	7	64008	2987	3	4

No	Nama	Kapasitas (Kw)	Jumlah KWH /Daya	Jumlah Transaksi	Jumlah Charger	Jumlah Connector
4	SPKLU PLN UP3 CENGKARENG	7	24	5	2	2
5	SPKLU GUDANG PLN UP3 PONDOK GEDE	22	8	2	1	1
6	SPKLU POSKO YANTEK KAPUK	22	76	17	2	2
7	SPKLU Ruko Sentra Niaga Kalimalang	22	84	5	2	2
8	SPKLU TOMANG TOL	22	0	0	1	1
9	SPKLU GEDUNG CYBER-1	22	0	0	3	3
10	SPKLU PT SPE	22	0	0	2	2
...	...	...	...	...	...	...
218	SPKLU CLUSTER DAMAR	22	0	0	3	3
219	SPKLU EVCITY Apartemen Pantai Mutiara	120	0	0	3	5

b. Normalisasi Data

Data yang tercantum dalam Tabel 2 selanjutnya diproses melalui tahap normalisasi menggunakan teknik *Robust Scaler* untuk menyeimbangkan skala antar fitur. Langkah ini penting agar perhitungan jarak antar titik data dalam algoritma K-Means tidak bias terhadap atribut dengan skala yang lebih besar. Setelah normalisasi, data siap digunakan sebagai input untuk proses klusterisasi.

Dengan penerapan algoritma K-Means, sistem mampu mengelompokkan SPKLU ke dalam beberapa klaster berdasarkan kesamaan pola penggunaan. Hasil klusterisasi ini selanjutnya menjadi dasar bagi pengambilan keputusan strategis, termasuk penentuan prioritas peningkatan layanan, penambahan unit SPKLU, maupun pemerataan distribusi infrastruktur.

Dalam pemetaan SPKLU, perusahaan pengembang menggunakan sejumlah variabel sebagai indikator utama untuk menilai performa, pemanfaatan, serta kebutuhan peningkatan layanan setiap stasiun. Variabel-variabel ini dipilih dengan mempertimbangkan ketersediaan data aktual dan relevansinya terhadap tujuan penelitian, yaitu untuk mengelompokkan SPKLU berdasarkan pola penggunaan dan kemampuan teknisnya. Tabel 3 memberikan penjelasan mengenai definisi dan makna masing-masing variabel yang digunakan dalam proses pemetaan serta klusterisasi SPKLU.

Tabel 3 Variabel yang Digunakan Beserta Definisinya

No	Variabel	Definisi
1	Kapasitas (KWH)	Menyatakan kapasitas maksimum daya yang tersedia pada masing-masing SPKLU dalam satuan kilowatt-hour (kWh).
2	Jumlah KWH/Daya	Menyatakan jumlah energi listrik (kWh) yang telah digunakan atau dikeluarkan oleh SPKLU selama periode tertentu.
3	Jumlah Transaksi	Menyatakan total jumlah transaksi pengisian daya oleh pengguna kendaraan listrik selama periode tertentu.
4	Jumlah Charger	Menyatakan jumlah unit charger fisik yang tersedia pada SPKLU

No	Variabel	Definisi
5	Jumlah Konektor	Menyatakan jumlah total konektor yang tersedia pada charger

Sebelum diterapkan pada algoritma K-Means, kelima variabel tersebut terlebih dahulu dinormalisasi untuk mencegah dominasi salah satu variabel akibat perbedaan skala atau satuan. Setelah normalisasi, variabel-variabel ini menjadi dasar utama dalam pengelompokan SPKLU ke dalam kategori kluster seperti “Rendah”, “Sedang”, dan “Tinggi” berdasarkan kesamaan karakteristik penggunaannya.

Proses normalisasi bertujuan menyamakan skala nilai di setiap atribut agar tidak ada atribut yang mendominasi saat klusterisasi. Dalam penelitian ini, digunakan metode Robust Scaler yang memanfaatkan median dan rentang interkuartil (IQR) untuk meminimalkan pengaruh outlier (nilai ekstrem) selama proses penskalaan.

$$x' = \frac{x - \text{median}(x)}{Q3 - Q1}$$

Di mana:

x = nilai asli dari (variabel) yang akan dinormalisasi

$\text{median}(x)$ = nilai median (nilai tengah) dari seluruh variabel x

$Q1$ = kuartil pertama (persentil ke-25) dari variabel x

$Q3$ = kuartil ketiga (persentil ke-75) dari variabel x

$Q3-Q1$ = *Interquartile Range* (IQR), yaitu rentang di antara kuartil

Langkah pertama yaitu cari nilai Q_1 , Q_2 (Median), Q_3 dan IQR setiap variable dari data baris ke-1 sampai ke-219.

Tabel 4 Nilai Median, Q1, Q2 dan IQR Setiap Variabel

Variabel	Median	Q1	Q3	IQR
kapasitas_kw	22	22	30	8
jumlah_kwh	1484	161	17973	17812
jumlah_transaksi	76	11	686	675
jumlah_charger	2	1	2	1
jumlah_konektor	2	1	3	2

Setelah memperoleh nilai median, Q1, Q2 dan IQR maka lanjut ke langkah perhitungan normalisasi pada setiap variabel dengan metode robust scaler. Berikut persamaannya;

$$x' = \frac{x - \text{median}(x)}{IQR}$$

Perhitungan pada variabel Kapasitas (Kwh) adalah sebagai berikut:

Data ke-1:

$$x' = \frac{200 - 22}{8} = 22,25$$

Data ke-2:

$$x' = \frac{200 - 22}{8} = 22,25$$

Data ke-3:

$$x' = \frac{7 - 22}{8} = -1,88$$

... ..

Data ke-218:

$$x' = \frac{22 - 22}{8} = 0$$

Data ke-219:

$$X' = \frac{120 - 22}{8} = 12,25$$

Perhitungan pada variabel Jumlah KWH/Daya adalah sebagai berikut:

Data ke-1:

$$X' = \frac{810624 - 1484}{17812} = 45,43$$

... ..

Data ke-2:

$$X' = \frac{399048 - 1484}{17812} = 22,32$$

Data ke-218:

$$X' = \frac{0 - 1484}{17812} = -0,08$$

Data ke-219:

$$X' = \frac{0 - 1484}{17812} = -0,08$$

Data ke-3:

$$X' = \frac{64008 - 1484}{17812} = 3,51$$

Perhitungan pada variabel Jumlah Transaksi adalah sebagai berikut:

Data ke-1:

$$X' = \frac{32640 - 76}{675} = 48,24$$

... ..

Data ke-218:

$$X' = \frac{0 - 6}{675} = -0,11$$

Data ke-2:

$$X' = \frac{13780 - 76}{675} = 20,30$$

Data ke-219:

$$X' = \frac{0 - 76}{675} = -0,11$$

Data ke-3:

$$X' = \frac{2987 - 76}{675} = 4,31$$

Perhitungan pada variabel Jumlah Charger adalah sebagai berikut:

Data ke-1:

$$X' = \frac{4 - 2}{1} = 2$$

... ..

Data ke-218:

$$X' = \frac{3 - 2}{1} = 1$$

Data ke-2:

$$X' = \frac{2 - 2}{1} = 0$$

Data ke-219:

$$X' = \frac{3 - 2}{1} = 1$$

Data ke-3:

$$X' = \frac{3 - 2}{1} = 1$$

Perhitungan pada variabel Jumlah Connector adalah sebagai berikut:

Data ke-1:

$$X' = \frac{10 - 2}{2} = 4$$

$$X' = \frac{4 - 2}{2} = 1$$

Data ke-2:

... ..

Data ke-218:

$$X' = \frac{5 - 2}{2} = 1,5$$

$$X' = \frac{3 - 2}{2} = 0,5$$

Data ke-3:

Data ke-219:

$$X' = \frac{5 - 2}{2} = 1,5$$

Bisa dilihat pada Tabel 5 merupakan hasil dari perhitungan *robust scaller* untuk normalisasi data setiap variabel sebagai berikut;

Tabel 5 Hasil Normalisasi Data

Data Ke-	Kapasitas (Kw)	Jumlah /Daya	Jumlah Transaksi	Jumlah Charger	Jumlah Connector
1	22,25	45,43	48,24	2,00	4,00
2	22,25	22,32	20,30	0,00	1,50
3	-1,88	3,51	4,31	1,00	1,00
4	-1,88	-0,08	-0,11	0,00	0,00
5	0,00	-0,08	-0,11	-1,00	-0,50
6	0,00	-0,08	-0,09	0,00	0,00
7	0,00	-0,08	-0,11	0,00	0,00
8	0,00	-0,08	-0,11	-1,00	-0,50
9	0,00	-0,08	-0,11	1,00	0,50
10	0,00	-0,08	-0,11	0,00	0,00
...	...	...	...	...	...
218	0,00	-0,08	-0,11	1,00	0,50
219	12,25	-0,08	-0,11	1,00	1,50

Data yang digunakan merupakan data milik perusahaan pengembang SPKLU pada bulan Januari 2021 sampai Maret 2025. Data tersebut telah melalui tahap normalisasi menggunakan metode Robust Scaller sehingga mengurangi pengaruh outlier (nilai ekstrem) saat melakukan scaling. Dataset yang telah dinormalisasi ini terdiri dari 219 baris data, 5 variabel dan disajikan pada Tabel 6.

Tabel 6 Data Hasil Normalisasi

Data Ke-	Nama	Kapasitas (Kw)	Jumlah /Daya	Jumlah Transaksi	Jumlah Charger	Jumlah Connector
1	SPKLU PLN UID JAKARTA RAYA	22,25	45,43	48,24	2,00	4,00
2	SPKLU PLN UP3 Bulungan	22,25	22,32	20,30	0,00	1,50
3	SPKLU PLN UP3 LENTENG AGUNG	-1,88	3,51	4,31	1,00	1,00
4	SPKLU PLN UP3 CENGKARENG	-1,88	-0,08	-0,11	0,00	0,00
5	SPKLU GUDANG PLN UP3 PONDOK GEDE	0,00	-0,08	-0,11	-1,00	-0,50
6	SPKLU POSKO	0,00	-0,08	-0,09	0,00	0,00

Data Ke-	Nama	Kapasitas (Kw)	Jumlah /Daya	Jumlah Transaksi	Jumlah Charger	Jumlah Connector
	YANTEK KAPUK					
7	SPKLU Ruko Sentra Niaga Kalimantan	0,00	-0,08	-0,11	0,00	0,00
8	SPKLU TOMANG TOL	0,00	-0,08	-0,11	-1,00	-0,50
9	SPKLU GEDUNG CYBER-1	0,00	-0,08	-0,11	1,00	0,50
10	SPKLU PT SPE	0,00	-0,08	-0,11	0,00	0,00
...	...					
218	SPKLU CLUSTER DAMAR	0,00	-0,08	-0,11	1,00	0,50
219	SPKLU EVCITY Apartemen Pantai Mutiara	12,25	-0,08	-0,11	1,00	1,50

c. Menentukan Jumlah Cluster

Dari dataset perusahaan pengembang SPKLU yang terdapat pada Tabel 4.7, proses klasterisasi dilakukan dengan membentuk tiga klaster, sehingga nilai K dalam algoritma K-Means ditetapkan menjadi 3. Penentuan jumlah klaster ini didasarkan pada kebutuhan untuk mengelompokkan unit SPKLU ke dalam tiga kategori utama.

Berikut adalah interpretasi dari masing-masing klaster:

1. Klaster 0 yaitu SPKLU dengan tingkat penggunaan dan kapasitas yang rendah,
2. Klaster 1 yaitu SPKLU dengan tingkat pemanfaatan dan fasilitas yang sedang,
3. Klaster 2 yaitu SPKLU dengan tingkat pemakaian yang tinggi.

d. Menentukan Titik Centroid

Dalam proses K-Means Clustering, langkah pertama yang sangat penting adalah menentukan titik centroid awal untuk setiap klaster. Karena kita telah menetapkan jumlah klaster menjadi 3, maka kita memerlukan 3 titik centroid awal, yaitu untuk klaster 0, klaster 1, dan klaster 2.

Titik centroid awal ini ditentukan secara acak oleh algoritma, yang mengambil nilai dari data yang sebelumnya telah dinormalisasi menggunakan robust scaler. Titik-titik centroid ini akan menjadi acuan awal dalam mengelompokkan setiap data SPKLU berdasarkan jarak Euclidean terdekat. Anda dapat melihat nilai-nilai centroid awal tersebut pada Tabel 7.

Tabel 7 Nilai Centroid Awal

Kalster 0	Kalster 1	Klaster 2
-1,88	1,54	22,25
-0,08	1,54	45,43
-0,11	1,49	48,24
-1,00	-0,22	3,00
-0,50	0,16	4,00

e. Hitung Jarak Data Dengan Centroid

Perhitungan jarak antara setiap data SPKLU dan titik centroid dilakukan untuk menentukan seberapa dekat atau mirip data tersebut dengan masing-masing klaster. Metode yang digunakan adalah Jarak Euclidean, yang mengukur jarak lurus dalam ruang multidimensi antara dua titik. Rumus Jarak Euclidean antara titik data dan titik centroid c dalam n dimensi adalah sebagai berikut:

$$D_e = \sqrt{(x_1 - c_1)^2 + (x_2 - c_2)^2 + \dots + (x_n - c_n)^2}$$

Dimana:

D_e = Jarak Euclidean.

$x_1, x_2, \dots, x_n$ = Nilai atribut dari satu titik data.

$c_1, c_2, \dots, c_n$ = Nilai atribut dari centroid kluster yang dibandingkan

n = Jumlah variabel yang digunakan dalam proses klusterisasi

$(x_i - c_i)^2$ = Kuadrat dari selisih nilai variabel ke i data dengan centroid ke i

Iterasi Pertama

Perhitungan jarak dengan kluster 0:

Data ke-1

$$D(X_1, C_0)$$

$$D_e = \sqrt{(22,25 - (-1,88))^2 + (45,43 - (-0,08))^2 + (48,24 - (-0,11))^2 + (2,00 - (-1,00))^2 + (4,00 - (-0,50))^2} = 70,86$$

Data ke-2

$$D(X_2, C_0)$$

$$D_e = \sqrt{(22,25 - (-1,88))^2 + (22,32 - (-0,08))^2 + (20,30 - (-0,11))^2 + (0,00 - (-1,00))^2 + (1,50 - (-0,50))^2} = 38,80$$

Data ke-3

$$D(X_3, C_0)$$

$$D_e = \sqrt{(-1,88 - 1,88)^2 + (3,51 - (-0,08))^2 + (4,31 - (-0,11))^2 + (1,00 - (-1,00))^2 + (1,00 - (-0,50))^2} = 6,22$$

... ..

Data ke-219

$$D(X_{219}, C_0)$$

$$D_e = \sqrt{(12,25 - (-1,88))^2 + (-0,08 - (-0,08))^2 + (-0,11 - (-0,11))^2 + (1,00 - (-1,00))^2 + (1,50 - (-0,50))^2} = 14,41$$

Perhitungan jarak dengan kluster 1:

Data ke-1

$$D(X_1, C_1)$$

$$D_e = \sqrt{(22,25 - 1,54)^2 + (45,43 - 1,54)^2 + (48,24 - 1,49)^2 + (2,00 - (-0,22))^2 + (4,00 - 0,16)^2} = 67,53$$

Data ke-2

$$D(X_2, C_1)$$

$$D_e = \sqrt{(22,25 - 1,54)^2 + (22,32 - 1,54)^2 + (20,30 - 1,49)^2 + (0,00 - (-0,22))^2 + (1,50 - 0,16)^2} = 34,88$$

Data ke-3

$$D(X_3, C_1)$$

$$D_e = \sqrt{\frac{(-1,88 - 1,54)^2 + (3,51 - 1,54)^2 + (4,31 - 1,49)^2}{+(1,00 - (-0,22))^2 + (1,00 - 0,16)^2}} = 5,07$$

.....

Data ke-219

$$D(X_{219}, C_1)$$

$$D_e = \sqrt{\frac{(12,25 - 1,54)^2 + (-0,08 - 1,54)^2 + (-0,11 - 1,49)^2}{+(1,00 - (-0,22))^2 + (1,50 - 0,16)^2}} = 11,10$$

Perhitungan jarak dengan kluster 2:

Data ke-1

$$D(X_1, C_2)$$

$$D_e = \sqrt{\frac{(22,25 - 22,25)^2 + (45,43 - 45,43)^2 + (48,24 - 48,24)^2}{+(2,00 - 3,00)^2 + (4,00 - 4,00)^2}} = 1,00$$

Data ke-2

$$D(X_2, C_2)$$

$$D_e = \sqrt{\frac{(22,25 - 22,25)^2 + (22,32 - 45,43)^2 + (20,30 - 48,24)^2}{+(0,00 - 3,00)^2 + (1,50 - 4,00)^2}} = 36,47$$

Data ke-3

$$D(X_3, C_2)$$

$$D_e = \sqrt{\frac{(-1,88 - 22,25)^2 + (3,51 - 45,43)^2 + (4,31 - 48,24)^2}{+(1,00 - 3,00)^2 + (1,00 - 4,00)^2}} = 65,44$$

.....

Data ke-219

$$D(X_{219}, C_2)$$

$$D_e = \sqrt{\frac{(0,00 - 22,25)^2 + (-0,08 - 45,43)^2 + (-0,11 - 48,24)^2}{+(1,00 - 3,00)^2 + (0,50 - 4,00)^2}} = 70,15$$

Tabel 8 Hasil Pengelompokan Iterasi Pertama

Data Ke-	Nama	ED0	ED1	ED2	Jarak Terdekat	kelas
1	SPKLU PLN UID JAKARTA RAYA	70,8568	67,5325	1,0000	1,0000	2
2	SPKLU PLN UP3 Bulungan	38,8032	34,8801	36,4672	34,8801	1

Data Ke-	Nama	ED0	ED1	ED2	Jarak Terdekat	kelas
3	SPKLU PLN UP3 LENTENG AGUNG	6,2246	5,0684	65,4361	5,0684	1
4	SPKLU PLN UP3 CENGKARENG	1,1181	4,1090	70,8209	1,1181	0
5	SPKLU GUDANG PLN UP3 PONDOK GEDE	1,8750	2,9300	70,2881	1,8750	0
6	SPKLU POSKO YANTEK KAPUK	2,1832	2,7467	70,1902	2,1832	0
7	SPKLU Ruko Sentra Niaga Kalimalang	2,1830	2,7567	70,2022	2,1830	0
9	SPKLU TOMANG TOL	1,8750	2,9319	70,2904	1,8750	0
10	SPKLU GEDUNG CYBER-1	2,9182	3,0299	70,1480	2,9182	0
11	SPKLU PT SPE	2,1830	2,7637	70,2103	2,1830	0
...	...	...	...	...	...	2
218	SPKLU CLUSTER DAMAR	2,9182	11,1046	67,2285	2,9182	0
219	SPKLU EVCITY Apartemen Pantai Mutiara	14,4054	2,6507	70,0788	2,6507	1

Tabel 8 merupakan hasil perhitungan dengan rumus *Euclidean Distance* antara tiap variable dengan *centroid* kluster 0, kluster 1 dan kluster 2, sehingga didapat jarak terdekat dari masing-masing atribut (data N), kemudian untuk ditentukan masuk ke kluster masing-masing. Setelah proses pengelompokan data lalu dilakukan penentuan titik *centroid* baru. Nilai *centroid* baru didapatkan dengan menggunakan rumus berikut:

$$\mu_k = \left(\frac{1}{n_k} \right) \sum_{i=1}^{n_k} x_{ik}$$

Dimana:

- μ_k = Titik centroid dari kluster ke-k
- n_k = Banyaknya data pada kluster ke-k
- x_{ik} = Data ke-i kluster ke-k

Mencari nilai centroid kluster 0 baru:

$$\mu_0 = \frac{-24,98}{130} = -0,19$$

$$\mu_1 = \frac{-1,33}{130} = -0,01$$

$$\mu_2 = \frac{3,08}{130} = 0,02$$

$$\mu_3 = \frac{-54}{130} = -0,42$$

$$\mu_4 = \frac{-24}{130} = -0,18$$

Mencari nilai centroid kluster 1 baru:

$$\mu_0 = \frac{335,5}{87} = 3,86$$

$$\mu_1 = \frac{262,79}{87} = 3,02$$

$$\mu_2 = \frac{250,61}{87} = 2,88$$

$$\mu_3 = \frac{1}{87} = 0,01$$

$$\mu_4 = \frac{51,5}{87} = 0,59$$

Mencari nilai centroid klaster 2 baru:

$$\mu_0 = \frac{25,75}{2} = 12,88$$

$$\mu_1 = \frac{76,06}{2} = 38,03$$

$$\mu_2 = \frac{72,53}{2} = 36,26$$

$$\mu_3 = \frac{4}{2} = 2$$

$$\mu_4 = \frac{6,5}{2} = 3,25$$

Sehingga di dapatkan hasil centroid baru dari iterasi pertama untuk klaster 0, klaster 1 dan klaster 2 bisa dilihat pada tabel 9.

Tabel 9 Centroid Baru Hasil Iterasi Pertama

Kalster 0	Kalster 1	Klaster 2
-0,1921	3,8563	12,8750
-0,0102	3,0206	38,0324
0,0237	2,8806	36,2637
-0,4154	0,0115	2,0000
-0,1846	0,5920	3,2500

Ulangi perhitungan jarak data dengan centroid sampai nilai dari titik centroid tidak lagi berubah. Dari perhitungan yang dilakukan pada penelitian ini terjadi iterasi sebanyak 5 kali titik pusat cluster tidak ada lagi perubahan dan data pertama hingga data yang terakhir tidak ada lagi yang bergeser dari satu cluster ke cluster yang lainnya. Adapun hasil perhitungan pada iterasi terakhir dapat dilihat pada tabel 10, 11 dan 12 berikut:

Tabel 10 Hasil Pemetaan SPKLU Pada Klaster 0

No	nama	kapasitas_kwh	jumlah_kwh	jumlah_transaksi	jumlah_charger	jumlah_connector	kluster
1	SPKLU PLN UP3 LENTENG AGUNG	7	64008	2987	3	4	0
2	SPKLU PLN UP3 Jatinegara	7	16645	1762	2	2	0
3	SPKLU PLN UP3 TANJUNG PRIOK	7	43680	2129	4	5	0
4	SPKLU PT POS FATMAWATI	25	39299	2544	2	3	0
5	SPKLU PLN UP3 PONDOK KOPI	7	45196	2057	3	5	0
6	SPKLU MENARA BNI	25	37838	1693	2	3	0
7	SPKLU GRHA BNI	25	48608	2354	2	3	0
8	SPKLU KFC JALAN PANJANG	22	17462	1642	1	1	0
9	SPKLU Brilian 1	25	48410	2076	2	3	0

No	nama	kapasitas_kwh	jumlah_kwh	jumlah_transaksi	jumlah_charger	jumlah_connector	kluster
10	SPKLU TMII (Taman Mini Indonesia Indah)	60	11009	630	1	3	0
...	...	...	...	...	...	...	...
179	SPKLU RS EMC KEDOYA	22	0	0	2	2	0
180	SPKLU CLUSTER DAMAR	22	0	0	3	3	0

Tabel 11 Hasil Pemetaan SPKLU Pada Klaster 1

No	nama	kapasitas_kwh	jumlah_kwh	jumlah_transaksi	jumlah_charger	jumlah_connector	kluster
1	SPKLU INDOMOBIL	25	71888	4123	2	3	1
2	SPKLU Kantor DJK Kementrian ESDM	30	88107	4561	3	5	1
3	SPKLU BBG Plaza Mandiri	25	71732	3717	2	3	1
4	SPKLU PLN UP3 CEMPAKA PUTIH	7	218886	9573	4	6	1
5	SPKLU AEON MALL TANJUNG BARAT	50	98799	4787	1	3	1
6	SPKLU PLN UP3 KEBUN JERUK	60	71874	3224	4	6	1
7	SPKLU PLN UP3 BANDENGAN	60	205076	8849	2	4	1
8	SPKLU AEON JGC	60	124234	5306	1	3	1
9	SPKLU PLN UIP JBB	30	95419	4620	3	5	1
10	SPKLU KFC Artha Gading	25	93404	4738	2	3	1
...	...	...	...	...	...	...	...
34	SPKLU (ARISTA POWER) BYD ARISTA PLUIT	180	2232	3	3	6	1
35	SPKLU EVCITY Apartemen Pantai Mutiara	120	0	0	3	5	1

Tabel 12 Hasil Pemetaan SPKLU Pada Klaster 2

No	nama	kapasitas_kwh	jumlah_kwh	jumlah_transaksi	jumlah_charger	jumlah_connector	kluster
1	SPKLU PLN UID JAKARTA RAYA	200	810624	32640	4	10	2
2	SPKLU PLN UP3 Bulungan	200	399048	13780	2	5	2
3	SPKLU PLN Kantor Pusat	50	547209	16468	4	7	2
4	SPKLU POOL DAMRI KEMAYORAN	150	439238	12275	1	2	2

2. Pembahasan

Berdasarkan hasil pengembangan pada metode yang diterapkan yaitu K-Means clustering. Hasil yang diuji dengan menggunakan *Silhouette Coefficient*. Metode ini yaitu untuk menguji kualitas dari setiap kluster yang telah dihasilkan. Hasil perhitungan akhir dengan menggunakan metode *K-Means Clustering* sebagai hasil dari kluster. Adapun proses dari perhitungan *Silhouette Coefficient* dengan Langkah-langkah sebagai berikut:

1. Menghitung suatu data dari rata-rata jarak setiap kluster, Misal data ke-1 dengan data lain yang berada pada kluster yang sama:

Klaster 0:

$$\alpha(i)_0 = \frac{\sqrt{(-1.88 - -1.88)^2 + (3.51 - 0.85)^2 + (4.31 - 2.50)^2 + (1.00 - 0.00)^2 + \dots + (1.00 - 0.00)^2}}{179} = 6.06$$

$$\alpha(i)_0 = \frac{\sqrt{(-1.88 - -1.88)^2 + (0.85 - 3.51)^2 + (2.50 - 4.31)^2 + (0.00 - 1.00)^2 + \dots + (0.00 - 1.00)^2}}{179} = 3.57$$

Klaster 1:

$$\alpha(i)_1 = \frac{\sqrt{(0.38 - 1.00)^2 + (3.95 - 4.86)^2 + (6.00 - 6.64)^2 + (0.00 - 1.00)^2 + \dots + (0.50 - 1.50)^2}}{34} = 8.47$$

$$\alpha(i)_1 = \frac{\sqrt{(1.00 - 0.38)^2 + (4.86 - 3.95)^2 + (6.64 - 6.00)^2 + (1.00 - 0.00)^2 + \dots + (1.50 - 0.50)^2}}{34} = 8.34$$

Klaster 2:

$$\alpha(i)_2 = \frac{\sqrt{(22.25 - 22.25)^2 + (45.43 - 22.32)^2 + (48.24 - 20.30)^2 + (2.00 - 0.00)^2 + \dots + (4.00 - 1.50)^2}}{3} = 35.93$$

$$\alpha(i)_2 = \frac{\sqrt{(22.25 - 22.25)^2 + (22.32 - 45.43)^2 + (20.30 - 48.24)^2 + (0.00 - 2.00)^2 + \dots + (1.50 - 4.00)^2}}{3} = 21.55$$

Tabel 13 Rata-Rata Jarak Pada Kluster

Data Ke	a(i)
p1	35,93
p2	21,55
p3	6,06
p4	3,57
p5	5,16
p6	8,47
p7	4,31
p8	23,53
p9	8,34
p10	8,39
...	...
p218	2,16
p219	11,01

Tabel 13 ialah hasil perhitungan jarak rata-rata dari setiap data yang berada pada kluster yang sama.

- Menghitung rata-rata jarak dari data ke-1 dengan semua data pada kluster lain dan diambil nilai terkecilnya.
 Kluster 2 ke Kluster 0

$$b(P1)^{(2 \rightarrow 0)} = \frac{\sqrt{(22.25 - -1.88)^2 + (45.43 - 3.51)^2 + (48.24 - 4.31)^2 + (2.00 - 1.00)^2 + \dots + (4.00 - 1.00)^2}}{180} = 69.63$$

Kluster 2 ke Kluster 1

$$b(P1)^{(2 \rightarrow 1)} = \frac{\sqrt{(22.25 - 0.38)^2 + (45.43 - 3.95)^2 + (48.24 - 6.00)^2 + (2.00 - 0.00)^2 + \dots + (4.00 - 0.50)^2}}{35} = 62.04$$

Kluster 0 ke Kluster 1

$$b(P3)^{(0 \rightarrow 1)} = \frac{\sqrt{(-1.88 - 0.38)^2 + (3.51 - 3.95)^2 + (4.31 - 6.00)^2 + (1.00 - 0.00)^2 + \dots + (1.00 - 0.50)^2}}{35} = 10.07$$

Kluster 0 ke Kluster 2

$$b(P3)^{(0 \rightarrow 2)} = \sqrt{\frac{(-1.88 - 22.25)^2 + (3.51 - 45.43)^2 + (4.31 - 48.24)^2 + (1.00 - 2.00)^2 + \dots + (1.00 - 4.00)^2}{4}} = 41.26$$

Klaster 1 ke Klaster 0

$$b(P6)^{(1 \rightarrow 0)} = \sqrt{\frac{(0.38 - -1.88)^2 + (3.95 - 3.51)^2 + (6.00 - 4.31)^2 + (0.00 - 1.00)^2 + \dots + (0.50 - 1.00)^2}{180}} = 7.04$$

Klaster 1 ke Klaster 2

$$b(P6)^{(1 \rightarrow 2)} = \sqrt{\frac{(0.38 - 22.25)^2 + (3.95 - 45.43)^2 + (6.00 - 48.24)^2 + (0.00 - 2.00)^2 + \dots + (0.50 - 4.00)^2}{4}} = 39.10$$

Tabel 14 Nilai Rata-Rata Jarak Pada Klatser Lain

Data Ke	d (i, C)		
	d(i,0)	d(i,1)	d(i,2)
p1	69,63	62,04	0
p2	37,11	28,96	0
p3	0	10,07	41,26
p4	0	11,01	44,08
p5	0	10,67	42,74
p6	7,04	0	39,1
p7	0	8,69	41,63
p8	39,04	33,27	0
p9	8,3	0	37,85
p10	6,56	0	39,43
...	...	...	...
p218	0	11,07	45,46
p219	12,27	0	42,58

Tabel 14 merupakan hasil perhitungan jarak rata-rata dari data i dengan semua data pada klaster lain. Jika jarak rata-rata dari data I dengan semua data pada kluster lain sudah dilakukan, maka diambil nilai terkecilnya. Maka diperoleh nilai terkecil pada jarak rata-rata dari data I dengan semua data pada klaster lain dapat dilihat pada tabel 15.

Tabel 15 Nilai Terkecil Pada Jarak Rata-Rata Klaster Lain

Data Ke	b(i)
p1	62,04
p2	28,96
p3	10,07
p4	11,01
p5	10,67
p6	7,04

Data Ke	b(i)
p7	8,69
p8	33,27
p9	8,30
p10	6,56
...	...
p218	11,07
p219	12,27

3. Setelah diperoleh nilai a(i) dan b(i), hitung nilai silhouette coefficient sebagai berikut:

Data ke-1

$$s(P1) = \frac{b(P1) - a(P1)}{\max(a(P1), b(P1))} = \frac{62.04 - 35.93}{62.04} = 0.42$$

Data ke-2

$$s(P2) = \frac{b(P2) - a(P2)}{\max(a(P2), b(P2))} = \frac{28.96 - 21.55}{28.96} = 0.26$$

Data ke-3

$$s(P3) = \frac{b(P3) - a(P3)}{\max(a(P3), b(P3))} = \frac{10.07 - 6.06}{10.07} = 0.40$$

...

Data ke-219

$$s(P219) = \frac{b(P219) - a(P219)}{\max(a(P219), b(P219))} = \frac{12.27 - 11.01}{12.27} = 0.10$$

Tabel 16 Nilai Silhouette Coefficient Semua Data

Data Ke	s(i)
p1	0,421
p2	0,256
p3	0,398
p4	0,676
p5	0,517
p6	-0,169
p7	0,504
p8	0,293
p9	-0,004
p10	-0,219
...	...
p218	0,805
p219	0,103

Tabel 16 merupakan hasil dari perhitungan silhouette coefficient pada semua data. Sehingga didapatkan untuk nilai silhouette coefficientnya secara keseluruhan yaitu menghitung rata-rata nilai silhouette coefficient pada semua data sebagai berikut:

$$s_{(i)} = \frac{0,421+0,256+0,398+0,676+0,517+(-0,169)}{219} = \frac{+0,504+0,293+(-0,004)+(-0,219)+...+0,805+0,103}{219}$$

$$s_{(i)} = 0,66$$

Berdasarkan tabel 1 dan perolehan hasil nilai *silhouette coefficient* dengan nilai 0,66 berada pada kategori “Struktur Baik”, hal tersebut menunjukkan bahwa pengelompokan data SPKLU yang telah dilakukan menggunakan metode K-Means memiliki struktur klaster dengan kualitas dan kekuatan yang baik.

D. KESIMPULAN DAN SARAN

1. Kesimpulan

Berdasarkan hasil penelitian dan pengembangan penerapan metode K-Means untuk “Pemetaan Wilayah Penyebaran Stasiun Pengisian Kendaraan Listrik Umum (SPKLU) di Kota Jakarta”, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Penerapan algoritma K-Means pada penelitian ini mampu menghasilkan pemetaan lokasi SPKLU yang lebih akurat, sehingga distribusi wilayah penyebaran dapat terlihat dengan jelas berdasarkan karakteristik data.
2. Proses pemetaan wilayah penyebaran SPKLU menjadi lebih efektif, karena data dapat dikelompokkan ke dalam tiga cluster yaitu;
 - a. Klaster 0 (rendah) sebanyak 180 SPKLU (82,2%),
 - b. Klaster 1 (sedang) sebanyak 35 SPKLU (16%) dan
 - c. Klaster 2 (tinggi) sebanyak 4 SPKLU (1,8%)
3. Tingkat keakuratan dan keefektifan algoritma K-Means dapat diukur melalui evaluasi menggunakan nilai Silhouette Coefficient yang memperoleh nilai 0,66, yang menunjukkan kualitas hasil clustering pada data SPKLU di Jakarta masuk kedalam “Struktur Baik”.
4. Penelitian ini berhasil membuat prototype aplikasi berbasis K-Means yang dapat digunakan sebagai alat bantu dalam pengambilan keputusan terkait pemetaan lokasi SPKLU. Pengujian terhadap sistem menggunakan Post-Study System Usability Questionnaire (PSSUQ) menghasilkan skor 93,75 % yang mengindikasikan tingkat kepuasan pengguna dalam kategori yang “Sangat Layak”.

2. Saran

Berdasarkan temuan penelitian dan keterbatasan yang ditemukan selama proses pengembangan, berikut adalah beberapa saran untuk pengembangan lebih lanjut:

- 1) Peneliti mendatang dapat mempertimbangkan penggunaan algoritma clustering lain seperti DBSCAN, Hierarchical Clustering, atau Gaussian Mixture Models untuk membandingkan efektivitas pengelompokan.
- 2) Pengembangan sistem dapat ditingkatkan dengan menambahkan fitur optimasi otomatis untuk menentukan jumlah cluster optimal menggunakan metode seperti Elbow Method atau Gap Statistic.
- 3) Pengembangan sistem dapat diperluas dengan mengintegrasikan data dari berbagai sumber seperti data demografis, pola mobilitas penduduk, dan informasi infrastruktur transportasi.

E. DAFTAR PUSTAKA

- [1] L. Ketahanan and N. Ri, “Peningkatan penggunaan kendaraan listrik nasional guna mendukung ekonomi hijau,” pp. 1–105, 2023.
- [2] Databoks, “Riset Deloitte dan Foundry: Penggunaan Motor Listrik di Indonesia Naik 13 Kali Lipat dalam Dua Tahun,” *Katadata*, p. 1, 2023, [Online]. Available: <https://databoks.katadata.co.id/datapublish/2023/09/15/riset-deloitte-dan-foundry-penggunaan-motor-listrik-di-indonesia-naik-13-kali-lipat-dalam-dua-tahun>
- [3] M. I. Asaad, “Road Map Pengembangan Infrastruktur Kendaraan Listrik 2020 -2024,” *Roadmap infrastruktur kendaraan Elektr.*, vol. 1, no. September, p. 26, 2020, [Online]. Available: shorturl.at/jxHLM

- [4] A. M. Sugianto, “Penentuan Lokasi Stasiun Pengisian Kendaraan Listrik Umum Yang Optimum Dengan Menggunakan Analisis Spasial,” 2022, [Online]. Available: <https://repository.its.ac.id/93907/>
- [5] R. Oktavia *et al.*, “Penerapan Metode Algoritma K-means Dalam Pengelompokan Angka Harapan Hidup Saat Lahir Menurut Provinsi,” *J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 1, no. 4, pp. 154–161, 2020, [Online]. Available: <https://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/41>
- [6] N. Azwanti and N. E. Putria, “Penerapan Algoritma K-Means Untuk Pemetaan Penerimaan Bantuan Kesejahteraan Masyarakat di Kota Batam,” *Pros. Semin. Nas. Ilmu Sos. dan Teknol.*, vol. 5, no. September, pp. 250–256, Sep. 2023, doi: 10.33884/psnistek.v5i.8092.
- [7] P. J. Rousseeuw, “Siluet : bantuan grafis untuk interpretasi dan validasi analisis klaster,” vol. 1, 1987.
- [8] S. Suhirman, *Data Warehouse dan Data Mining*, no. January. 2023. [Online]. Available: <https://digitalpress.gaes-edu.com/index.php/gaespace/article/view/122>
- [9] P. Rahayu *et al.*, *Buku Ajar Data Mining*, vol. 1, no. January 2024. 2018.
- [10] F. SLN, *Buku Dasar Data Mining from A to Z*, no. January. 2023.