



Clustering Electricity Distribution Data Using the Mean Shift Algorithm

Pengelompokan Data Pendistribusian Listrik Menggunakan Algoritma Mean Shift

Roid Fitrah Utari¹, Fitri Insani^{2*}, Surya Agustian³, Liza Afrianti⁴

^{1,2,3,4}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

E-Mail: ¹12050113980@students.uin-suska.ac.id, ²fitri.insani@uin-suska.ac.id,
³surya.agustian@uin-suska.ac.id, ⁴liza.afrianti@uin-suska.ac.id

Received May 26th 2024; Revised Jun 18th 2024; Accepted Jun 20th 2024
Corresponding Author: Fitri insani

Abstract

This study investigates the regionalization and clustering of electricity distribution data in Indonesia using the Mean Shift algorithm, aimed at enhancing energy distribution efficiency across diverse geographical regions. Electricity plays a crucial role in modern life, yet its distribution remains uneven, especially in remote and rural areas constrained by access and financial limitations. As a major state-owned enterprise in the electricity sector, PLN is responsible for providing electricity nationwide, supporting economic growth through energy provision to industrial, agricultural, and commercial sectors. Utilizing the Mean Shift algorithm, the study divides Indonesia into Sumatra, Java-Bali, Kalimantan-Sulawesi, and Papua based on electricity distribution patterns. It finds that setting an optimal bandwidth of 0.5 results in three clusters per region reflecting similar infrastructure, energy demands, and dominant economic sectors. This underscores Mean Shift's flexibility in handling complex data structures without predetermined cluster numbers, crucial for strategic energy management planning in Indonesia to achieve more efficient and sustainable electricity distribution.

Keyword: Clustering, Data Mining, Electricity Distribution, Mean Shift Algorithm

Abstrak

Penelitian ini mengkaji regionalisasi dan klasterisasi data distribusi listrik di Indonesia menggunakan algoritma Mean Shift, dengan tujuan untuk meningkatkan efisiensi distribusi energi di berbagai wilayah geografis yang beragam. Listrik memiliki peran krusial dalam kehidupan modern namun distribusinya masih belum merata, terutama di daerah terpencil dan pedesaan yang terkendala oleh akses dan keterbatasan dana. Sebagai salah satu Badan Usaha Milik Negera (BUMN) utama di sektor ketenagalistrikan, Perusahaan Listrik Negera (PLN) bertanggung jawab dalam menyediakan listrik di seluruh Indonesia, mendukung pertumbuhan ekonomi melalui penyediaan energi untuk sektor industri, pertanian, dan perdagangan. Dengan menggunakan algoritma Mean Shift, penelitian ini mengelompokkan Indonesia menjadi Sumatra, Jawa-Bali, Kalimantan-Sulawesi, dan Papua berdasarkan pola distribusi listrik, dengan menemukan bahwa pengaturan bandwidth optimal 0.5 menghasilkan tiga klaster per wilayah yang mencerminkan infrastruktur serupa, kebutuhan energi, dan sektor ekonomi dominan. Temuan ini menunjukkan fleksibilitas Mean Shift dalam menangani struktur data yang kompleks tanpa jumlah klaster yang telah ditentukan sebelumnya, yang penting untuk perencanaan strategis dalam pengelolaan energi di Indonesia demi mencapai distribusi listrik yang lebih efisien dan berkelanjutan.

Kata Kunci: Data Mining, Mean Shift, Pendistribusian Listrik, Pengelompokan

1. PENDAHULUAN

Listrik merupakan kebutuhan pokok masyarakat modern. Hampir semua kegiatan setiap hari tidak dapat lepas dari listrik, seperti dari penerangan, memasak, hingga menggunakan peralatan elektronik. Oleh karena itu, pemerataan listrik menjadi hal yang penting untuk dilakukan. Namun, pemerataan listrik di Indonesia belum sepenuhnya merata. Masih banyak wilayah yang belum menikmati listrik, terutama di daerah terpencil dan pedesaan. Ada beberapa faktor yang jadi penyebab untuk hal ini, seperti sulitnya

penjangkauan lokasi dan keterbatasan dana [1]. Maka diperlukannya pengelompokan data dari penggunaan listrik pada setiap wilayah agar pemerataan listrik dapat dilakukan dengan optimal [2].

Perusahaan Listrik Negara (PLN) adalah salah satu Badan Usaha Milik Negara (BUMN) yang memiliki peran penting pada industri ketenagalistrikan Indonesia [3]. PLN bertanggung jawab untuk mengelola dan menyediakan listrik bagi seluruh wilayah Indonesia, termasuk daerah terpencil dan pedesaan. Perusahaan ini memiliki jaringan transmisi dan distribusi listrik yang luas, serta pembangkit listrik yang tersebar di seluruh Indonesia. PLN juga berperan penting untuk mendorong pertumbuhan ekonomi Indonesia. Perusahaan ini menyediakan listrik yang dibutuhkan oleh berbagai sektor ekonomi, seperti industri, pertanian, dan perdagangan. PLN juga membantu meningkatkan kualitas hidup masyarakat, dengan menyediakan listrik untuk penerangan, kebutuhan rumah tangga, dan sarana pendidikan dan kesehatan. dan Pada tahun 2020, 40,6% pembangkit listrik di Indonesia dimiliki oleh perusahaan swasta [4].

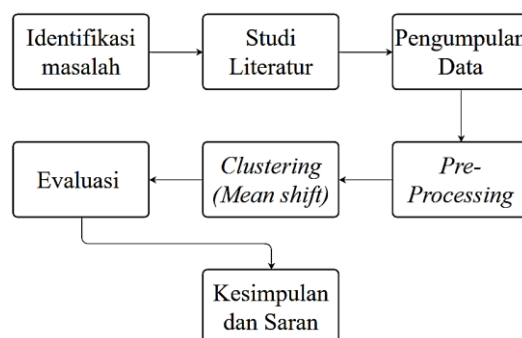
Pada tahun 2023, PLN mencatat bahwa rasio elektrifikasi di Indonesia sudah mencapai 99,74%. Hal ini berarti bahwa hampir seluruh masyarakat Indonesia sudah menikmati listrik. Ditahun 2019 yang semula jumlah pelanggan rumah tangga hanya 69,62 juta, tumbuh menjadi 80,56 juta pada triwulan III tahun 2023 atau meningkat sebesar 15,72 persen [1]. Peningkatan jumlah pelanggan dan kompleksitas pertumbuhan listrik di suatu wilayah menuntut penanganan yang lebih baik agar distribusi listrik dapat berjalan optimal [5]. Oleh karena itu, diperlukan metode yang dapat membantu dalam menentukan pengelompokan pendistribusian listrik di setiap wilayah berdasarkan beberapa sektor seperti industri, rumah tangga, bisnis, gedung kantor pemerintah, sosial, dan jalan umum. Metode ini bisa menerapkan teknik data mining, seperti pengelompokan data, dengan menggunakan algoritma *mean-shift*.

Algoritma *Mean-Shift* adalah metode non-parametrik yang dapat dimanfaatkan untuk pengelompokan data [6][7]. Berbeda dengan metode lain, *mean-shift* tidak memerlukan penetapan jumlah kluster terlebih dahulu. Cara kerjanya yaitu dengan menggunakan jendela geser berbentuk lingkaran yang bergerak mencari area terpadat data. Keunggulan utama *Mean-Shift* terletak pada fleksibilitasnya. Algoritma ini tidak terikat pada jumlah kluster yang pasti, sehingga ideal untuk situasi di mana struktur data tidak diketahui sebelumnya. Selain itu, *Mean-Shift* juga efisien dan mampu menangani data dengan dimensi tinggi [8].

Penelitian terdahulu yang dilakukan oleh Cici Astria, dkk. Mengenai pengelompokan wilayah pendistribusian listrik di Indonesia dengan metode *k-means* menunjukkan bahwa terdapat dua *cluster* wilayah, yaitu kluster dengan tingkat distribusi tinggi dan kluster dengan tingkat distribusi rendah [1]. Penelitian lainnya tentang penerapan algoritma *mean shift* untuk menentukan segmentasi pelanggan pada penjualan toko online melakukan penelitian dengan estimasi nilai *bandwidth* 1.55 dan menghasilkan 3 kluster [9]. Dan juga penelitian tentang aplikasi pengelompokan pelanggan retribusi alat pemadam kebakaran menggunakan metode *Mean Shift Clustering* pada Dinas Pemadam Kebakaran dan Penanggulangan Bencana Kota Palembang menghasilkan kesimpulan algoritma *mean shift clustering* efektif dalam melakukan klusterisasi data [10]. Lalu penerapan algoritma *Mean-Shift* pada klusterisasi penerimaan bantuan pangan non tunai menunjukkan bahwa algoritma ini bisa melakukan klusterisasi data tersebut dengan baik. Kesimpulan ini didasarkan pada hasil pengujian menggunakan *Silhouette Score*. Dalam penelitian ini, algoritma *mean-shift* menghasilkan kluster terbaik dengan nilai *bandwidth* 285 dan *Silhouette Score* 0.95. Nilai *Silhouette Score* sebesar 0.95 menunjukkan bahwa *cluster-cluster* yang dihasilkan sangat terkelompok dengan baik [6]. Dalam konteks penelitian pengelompokan data pendistribusian listrik ini, fokusnya adalah menggunakan algoritma *Mean-Shift* untuk melakukan pengelompokan data pendistribusian listrik yang diregionalisasi berdasarkan wilayah geografis.

2. METODOLOGI PENELITIAN

Metode penelitian merupakan panduan yang memastikan hasil yang dihasilkan sesuai dengan tujuan yaitu menerapkan algoritma *Mean-Shift* untuk melakukan Pengelompokan Data Pendistribusian Listrik [11]. Penjelasan terkait metode penelitian dalam studi ini tercantum dalam Gambar 1.



Gambar 1. Metodologi penelitian

2.1. Identifikasi Masalah

Dalam tahap ini, peneliti melakukan eksplorasi untuk mengidentifikasi dan memahami permasalahan yang ada [12]. Dalam konteks penelitian ini, perhatian difokuskan pada Bagaimana Pengelompokan Data Pendistribusian Listrik menggunakan Algoritma *Mean shift*.

2.2. Studi Literatur

Mengumpulkan referensi penting terkait pengelompokan data distribusi listrik, algoritma *Mean Shift*, dan *silhouette coefficient* untuk mendukung penelitian [13].

2.3. Pengumpulan Data

Pengumpulan data adalah hal pertama yang terpenting pada proses penelitian [14]. Data yang digunakan dalam penelitian ini merupakan data pendistribusian listrik. Data tersebut didapat secara langsung melalui dokumentasi pada situs web resmi PT. PLN yaitu web.pln.co.id [11]. Dalam proses pengumpulan data, data yang akan digunakan dalam bentuk data laporan pertahun dari tahun 2014 - 2022 dan terdapat tujuh atribut yaitu Satuan PLN/Provinsi, Rumah Tangga, Industri, Bisnis, Sosial, Gedung Kantor Pemerintah, dan Jalan Umum. Atribut-atribut ini digunakan sebagai dasar untuk mengelompokkan data menggunakan algoritma *Mean Shift* [15]. Salah satu data distribusi listrik dapat di perhatikan pada Tabel 1.

Tabel 1. Data distribusi listrik tahun 2022.

No	Satuan PLN/Provinsi	Rumah Tangga	Industri	Bisnis	Sosial	Gedung Kantor Pemerintah	Jalan Umum
1	Wilayah Aceh	1,895.07	228.77	558.37	228.92	109.78	133.11
2	Wilayah Sumatera Utara	6,025.69	3,152.63	1,808.45	501.49	152.05	419.50
3	Wilayah Sumatera Barat	1,791.64	827.73	639.01	192.18	82.58	97.30
4	Wilayah Riau	3,007.32	2,960.77	1,236.21	260.79	115.41	111.14
5	Wilayah Kepulauan Riau	499.83	39.76	301.90	40.88	41.20	22.98
6	Wilayah Sumatera Selatan	3,295.08	1,067.48	1,017.77	234.75	116.00	147.23
....							
33	Wilayah Jawa Barat	20,871.52	25,419.11	7,610.34	1,422.88	503.81	398.45
34	Wilayah Banten	6,049.89	16,605.16	3,407.59	395.19	163.58	84.26
35	Dist. Jakarta Raya	14,824.00	4,140.34	12,539.23	1,482.92	1,396.61	195.20

2.4. Clustering

Clustering merupakan suatu teknik untuk mengelompokkan data. Klasterisasi memungkinkan kita untuk melakukan pengelompokan data berdasarkan kesamaan karakteristiknya pada beberapa klaster atau kelompok sedemikian rupa sampai data dalam satu klaster yang punya tingkat kesamaan yang tinggi, sementara data antara klaster mempunyai kesamaan yang rendah. Klasterisasi sering digunakan sebagai metode analisis untuk mengelompokkan data menjadi kelompok dengan keterkaitan yang signifikan [18] [19]. Dalam penelitian kali ini peneliti menggunakan algoritma *Mean Shift* untuk melakukan *clustering* pada tiap variable atau sektor pada data distribusi listrik.

2.5. Algoritma Mean Shift

Algoritma *Mean-Shift* merupakan suatu teknik pengelompokan data yang memfokuskan pada penemuan titik pusat atau "nodes" melalui sebaran data yang menggunakan metode pergeseran jendela untuk mengidentifikasi wilayah-wilayah dengan tingkat kepadatan data yang tinggi. Algoritma ini didasarkan pada konsep centroid, dimana tujuannya yaitu mendapati titik pusat pada tiap kelompok atau kelas dengan cara memperbarui kandidat titik pusat menggunakan rata-rata titik-titik di dalam jendela pergeseran. Proses pasca-pemrosesan digunakan untuk menyaring kandidat jendela tersebut sehingga menghilangkan kemungkinan titik-titik pusat ganda yang nyaris serupa, Menyusun himpunan akhir dari titik-titik pusat beserta kelompok-kelompok yang terkait. [20]. Algoritma *Mean shift* bekerja dengan cara berikut:

1. Tentukan jenis kernel yang akan digunakan serta lebar bandwidth (h) yang akan mengatur jangkauan pencarian tetangga. Fungsi kernel.

$$(K(x)) = \left(\frac{1}{h}\right) \varphi\left(\frac{x}{h}\right) \quad (1)$$

di mana h merupakan bandwidth yang mempengaruhi jarak atau radius dari pusat cluster saat menghitung perpindahan rata-rata. Fungsi Gaussian atau Epanechnikov (φ) memungkinkan penentuan bobot pada setiap titik data Sesuai dengan jaraknya dari pusat cluster.

2. Lakukan inialisasi pusat-pusat massa awal yang akan menjadi centroid
3. Pada setiap centroid, hitung ulang titik pusat massa baru dengan rumus Mean-Shift.

$$m(x) = \frac{\sum (K(x - x_i)x_i)}{\sum K(x - x_i)} \quad (2)$$

Perhitungan pusat massa baru ($m(x)$) dihitung berdasarkan total dari fungsi kernel ($K(x - x_i)$) dibagi dengan total dari nilai fungsi kernel ($K(x - x_i)$) di dalam radius kernel yang diberikan, di mana x adalah pusat massa saat ini dan x_i adalah titik data dalam radius kernel.

4. Ulangi langkah perhitungan pusat massa baru untuk setiap centroid sampai centroid tidak berubah lagi atau memenuhi kriteria berhenti yang telah ditetapkan, seperti mencapai batas jumlah iterasi yang ditentukan.
5. Setelah proses konvergensi, tiap centroid akan merepresentasikan mode atau pusat dari kelompok data yang diidentifikasi.

2.6. Evaluasi

Evaluasi adalah tahap pengujian yang bertujuan untuk menentukan jumlah *bandwidth* yang optimal dalam algoritma *Mean Shift* dengan memanfaatkan *silhouette coefficient*. yang mana nilai siluet yang optimal adalah nilai yang mendekati satu, menunjukkan bahwa objek dalam suatu kluster saling berdekatan dengan baik dan terpisah dengan jelas dari kluster lainnya, berikut adalah rumus dari *silhouette coefficient* [6][21].

$$S = \frac{b - a}{\max(a, b)} \quad (3)$$

Keterangan :

- S : *Silhouette coefficient*
 a : Rata-rata jarak antara data point dan semua data point lainnya dalam satu cluster yang sama.
 b : Rata-rata jarak antara data point dan semua data point dalam cluster terdekat (cluster lain yang paling dekat)

3. HASIL DAN PEMBAHASAN

3.1 Pengolahan Data

Preprocessing bertujuan untuk mengubah data yang tidak terstruktur menjadi terstruktur dengan membersihkan, normalisasi, dan mempersiapkan data agar bisa diproses semakin efisien oleh algoritma yang di gunakan [16][17]. Dalam penelitian ini, data yang diperoleh telah terstruktur dan memiliki sedikit nilai null. Oleh karena itu, penulis telah menggabungkan data distribusi listrik dari tahun 2014 hingga 2022 ke dalam satu tabel yang ditampilkan dalam Tabel 2.

Tabel 2. Akumulasi data distribusi listrik tahun 2014 - 2022.

No	Satuan PLN/ Provinsi	Rumah Tangga	Industri	Bisnis	Sosial	Gedung Kantor Pemerintah	Jalan Umum
1	Wilayah Aceh	14,718.97	1,302.87	3,866.09	1,565.62	855.05	1,050.67
2	Wilayah Sumatera Utara	46,551.45	23,630.17	13,906.16	3,375.31	1,150.99	3,698.17
3	Wilayah Sumatera Barat	14,938.30	8,056.82	4,514.05	1,311.53	621.73	839.32
4	Wilayah Riau	22,881.68	6,931.58	9,077.21	1,690.11	881.13	1,228.73
5	Wilayah Kepulauan Riau	3,954.30	311.98	2,151.17	274.38	304.17	193.92
6	Wilayah Sumatera Selatan	27,097.24	8,066.36	7,913.83	1,617.35	916.32	1,177.85
....							
33	Wilayah Jawa Barat	166,559.08	205,651.98	50,909.48	9,347.15	3722	3,105.9
34	Wilayah Banten	40,207.45	111,446.98	218,91.22	2,311.47	1,161.61	673.2
35	Dist. Jakarta Raya	125,995.80	51,211.27	108,466.03	11,719.59	11,799.27	2,684.04

Langkah selanjutnya dalam pengolahan data, yaitu membagi keseluruhan provinsi berdasarkan geografisnya ke dalam beberapa daerah regional, yaitu regional Sumatra, Jawa-Bali, Kalimantan-Sulawesi,

dan Papua. Penulis juga memberikan label pada nama satuan PLN/Provinsi. Data yang telah diregionalisasikan beserta labelnya dapat dilihat pada Tabel 3 berikut ini.

Tabel 3. Data Satuan PLN/ Propinsi dikelompokkan ke dalam Region

Region	No	Satuan PLN/Provinsi	Label
Sumatera	1	Wilayah Aceh	Aceh
	2	Wilayah Sumatera Utara	Sumut
	3	Wilayah Sumatera Barat	Sumbar
	4	Wilayah Riau	Riau
	5	PT PLN Batam	Batam
	6	Wilayah Kepulauan Riau	Kepri
	7	Wilayah Jambi	Jambi
	8	Wilayah Bengkulu	Bengkulu
	9	Wilayah Sumatera Selatan	Sumsel
	10	Wilayah Bangka Belitung	Babel
	11	Distribusi Lampung	Lampung
Jawa-Bali	1	Wilayah Banten	Banten
	2	Wilayah Jawa Barat	Jabar
	3	Dist. Jakarta Raya	Jakarta
	4	Wilayah Jawa Tengah	Jateng
	5	Wilayah D.I. Yogyakarta	Jogja
	6	Dist. Jawa Timur	Jatim
	7	Distribusi Bali	Bali
	8	Wilayah Nusa Tenggara Barat	NTB
	9	Wilayah Nusa Tenggara Timur	NTT
Kalimantan-Sulawesi	1	Wilayah Kalimantan Barat	Kalbar
	2	Wilayah Kalimantan Tengah	Kalteng
	3	Wilayah Kalimantan Selatan	Kalsel
	4	Wilayah Kalimantan Timur dan Utara	Kaltim dan Kaltara
	5	PT PLN Tarakan	Tarakan
	6	Wilayah Sulawesi Barat	Sulbar
	7	Wilayah Sulawesi Selatan	Sulsel
	8	Wilayah Sulawesi Tengah	Sulteng
	9	Wilayah Sulawesi Tenggara	Sultra
	10	Wilayah Gorontalo	Gorontalo
	11	Wilayah Sulawesi Utara	Sulut
	12	Wilayah Maluku	Maluku
	13	Wilayah Maluku Utara	Malut
Papua	1	Wilayah Papua Barat	Pabar
	2	Wilayah Papua	Papua

Dari data yang telah diuraikan dapat dilihat bahwa diregional papua hanya terdapat dua wilayah pendistribusian listrik dan tidak perlu di sertakan dalam proses *clustering*, karena jumlah propinsi tidak mencukupi untuk membentuk klaster.

3.2. Pencarian Model Cluster Optimal

Untuk membentuk cluster yang paling optimal, parameter yang dilakukan penyesuaian dalam metode *mean shift* adalah *bandwidth*. Nilai *bandwidth* dipilih secara empiris dengan mencoba-coba angka dari 0.1 sampai 1.0. Namun dari semua percobaan, didapatkan hasil cluster yang beragam. Dalam penelitian ini, untuk memilih hasil cluster yang dianggap optimal, maka hasil cluster dievaluasi menggunakan *silhouette coefficient*.

Hasil klaster yang dianggap terbaik, dengan *silhouette score* tertinggi ditampilkan dalam Tabel 4-6 berikut ini. Dari Tabel 4 terlihat bahwa untuk *bandwidth* 0.3, jumlah klaster yang dihasilkan bervariasi antara 4-6 untuk ketiga regional yang dievaluasi. Sedangkan bila memilih *bandwidth* 0.4, hasil clusternya untuk Sumatera dan Kalimantan-Sulawesi menurun menjadi 3 klaster, sedangkan untuk Jawa-Bali masih tetap di angka 6. Dengan pengujian *bandwidth* 0.5, diperoleh ketiga region memiliki 3 cluster. Ditinjau dari nilai *silhouette score* masing-masing, maka nilai *silhouette* tertinggi berada pada *bandwidth* 0.5. Untuk selanjutnya, dipilih *bandwidth* 0.5 untuk membentuk klaster dari penelitian ini.

Table 4. Pengujian dengan nilai *bandwidth* 0.3

No	Regional	Jumlah kluster	<i>Silhouette Score</i>
1	Sumatra	4	0.44
2	Jawa-Bali	6	0.39
3	Kalimantan - Sulawesi	5	0.56

Table 5. Pengujian dengan nilai *bandwidth* 0.4

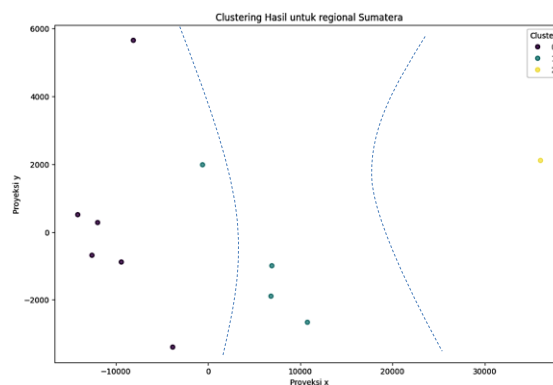
No	Regional	Jumlah kluster	<i>Silhouette Score</i>
1	Sumatra	3	0.50
2	Jawa-Bali	6	0.39
3	Kalimantan - Sulawesi	3	0.61

Table 6. Pengujian dengan nilai *bandwidth* 0.5

No	Regional	Jumlah kluster	<i>Silhouette Score</i>
1	Sumatra	3	0.50
2	Jawa-Bali	3	0.51
3	Kalimantan - Sulawesi	3	0.61

3.3. Hasil Clustering

Dalam tahap klasterisasi, penelitian ini menerapkan Algoritma *Mean Shift* menggunakan *Google Colab*. Setelah menemukan nilai *bandwidth* yang optimal pada pengujian sebelumnya maka pada penerapan algoritma *Mean Shift* peneliti menggunakan nilai *bandwidth* 0.5. Dengan menggunakan nilai *bandwidth* yang telah ditentukan, penelitian ini mengoptimalkan kinerja algoritma *Mean Shift* untuk mendapatkan hasil kluster yang akurat. Hasil kluster dan *scatter plot* dari tiap regional kemudian disajikan dalam Tabel 7-9 dan Gambar 2-4.

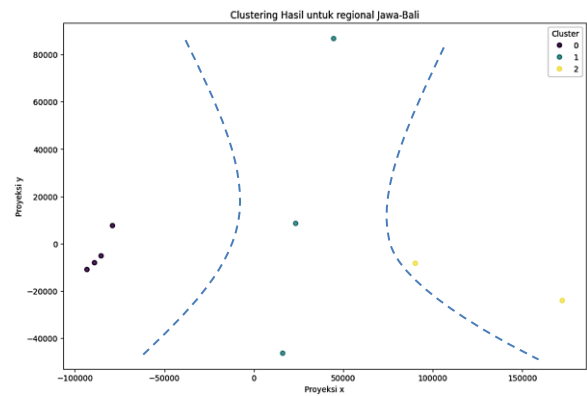
**Gambar 2.** *scatter plot* Klasterisasi Regional Sumatera.**Table 7.** Hasil klasterisasi Regional Sumatera.

C1	C2	C3
Kepri	Sumbar	Sumut
Bengkulu	Riau	
bangka	Lampung	
Batam	Sumsel	
Jambi		
Aceh		

Dari Tabel 7 dan Gambar 2 bisa dilihat wilayah-wilayah di kluster 1 cenderung memiliki karakteristik distribusi listrik yang mirip, karena kesamaan dalam infrastruktur listrik, populasi, dan tingkat urbanisasi. Pada kluster 2 wilayah seperti Sumbar, Riau, Lampung, dan Sumsel menunjukkan kesamaan dalam hal kebutuhan energi yang didorong oleh sektor agrikultur dan perkebunan. Sumut memiliki karakteristik unik dalam konsumsi distribusi listrik, tingginya aktivitas industri dan populasi yang besar. Hasil kluster ini memiliki kemiripan dengan metode DBSCAN [22], di mana untuk regional sumatera, Sumut terpisah dalam kluster sendiri, sedangkan untuk kluster tengah, ada perbedaan bahwa Aceh berada di dalam kluster tengah pada metode DBSCAN, pada metode *mean shift* Aceh masih satu kluster dengan Batam dan Kepri.

Sedangkan bila dibandingkan dengan metode *Mini Batch K-means* [23], analisisnya sama dengan perbandingan terhadap DBSCAN, karena hasil kluster keduanya sama. Oleh karena itu, hasil kluster ini

masih perlu diselidiki lebih lanjut dengan melibatkan data yang lebih banyak dan variatif dibandingkan data pemakaian listrik yang ada saat ini.



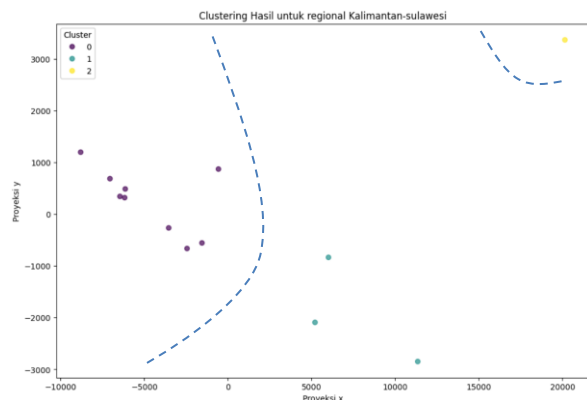
Gambar 3. scatter plot Klasterisasi Regional Jawa-Bali.

Table 8. Hasil klasterisasi Regional Jawa-Bali.

C0	C1	C2
NTT	Banten	Jakarta
NTB	Jateng	Jabar
Jogja	Jatim	
Bali		

Dari tabel 8 dan Gambar 3 dapat diperhatikan wilayah pada kluster 1 ini menunjukkan kesamaan dalam pola konsumsi listrik yang dipengaruhi oleh sektor pariwisata dan pendidikan, terutama di Jogja dan Bali, serta ekonomi agraris di NTT dan NTB. Pada kluster 2 yaitu wilayah Banten, Jateng, dan Jatim memiliki kebutuhan listrik yang signifikan akibat kombinasi antara industri berat dan kawasan permukiman. Di kluster 3 Jakarta dan Jabar memiliki kebutuhan energi yang tinggi dan kompleks karena merupakan pusat ekonomi dan industri utama di Indonesia.

Bila dibandingkan dengan metode DBSCAN, hasil kluster pada metode ini juga memiliki perbedaan. Dalam metode ini, Jakarta dan Jawa Barat berada dalam 1 kluster. Sedangkan pada kluster tengah terdapat Jawa Tengah, Jawa Timur dan Banten. Sisanya sama dengan metode DBSCAN [22]. Metode *mean shift* memasukkan Banten ke cluster tengah, padahal dari segi jumlah penduduk, kepadatan area industri dan gedung-gedung penting yang memakai listrik lebih banyak, kondisi menurut Tabel 2, Banten relatif agak jauh berbeda.



Gambar 4. scatter plot Klasterisasi Regional Kalimantan-Sulawesi.

Table 9. Hasil klasterisasi Regional Kalimantan-Sulawesi.

C1	C2	C3
Tarakan	Kalbar	Sulsel
Sulbar	Kalsel	
Malut	Kaltim dan Kaltara	
Maluku		

C1	C2	C3
Gorontalo		
Sultra		
Sulteng		
Kalteng		
Sulut		

Pada Tabel 9 dan Gambar 4 bisa diperhatikan wilayah klaster 1 menunjukkan karakteristik distribusi listrik yang dipengaruhi oleh kondisi geografis dan infrastruktur yang berkembang, dengan ketergantungan pada sektor agrikultur dan perikanan pada klaster 2 Kalbar, Kalsel, Kaltim, dan Kaltera menunjukkan kebutuhan listrik yang besar terkait dengan sektor pertambangan dan energi dan di klaster 3 Sulsel memiliki kebutuhan energi yang tinggi, dengan tingkat industrialisasi yang lebih maju dibandingkan wilayah lain di regional Kalimantan-Sulawesi.

Dari hasil dari penerapan algoritma DBSCAN [22] yang memiliki 3 klaster, untuk area Kalimantan-Sulawesi, terdapat perbedaan pada anggota klaster kanan. Metode ini hanya menempatkan Sulsel, pada klaster kanan, sedangkan DBSCAN menghasilkan 2 propinsi yaitu Kaltim/Kaltara dan Sulsel. Klaster tengah dan kiri sama, bila tanpa Kaltim/Kaltara dipindah ke klaster kanan. Sedangkan pada algoritma *Mini batch k-means* [23] hanya menghasilkan 2 cluster.

4. KESIMPULAN

Mengacu pada tujuan penelitian dan hasil analisis yang sudah dilakukan, bisa disimpulkan bahwa Penelitian ini berhasil melakukan regionalisasi dan klasterisasi data pendistribusian listrik di Indonesia menggunakan algoritma Mean Shift, menghasilkan klaster yang optimal dengan bandwidth 0.5. Evaluasi menunjukkan bahwa peningkatan nilai bandwidth menghasilkan kualitas klasterisasi yang lebih baik, dengan wilayah dalam klaster yang sama menunjukkan karakteristik distribusi listrik yang serupa. Sumatra, Jawa-Bali, dan Kalimantan-Sulawesi masing-masing memiliki tiga klaster yang mencerminkan kesamaan dalam infrastruktur, kebutuhan energi, dan sektor ekonomi dominan. Penggunaan algoritma ini memberikan wawasan berharga untuk perencanaan distribusi energi yang lebih efisien. Kekuatan penelitian ini terletak pada analisis mendalam dan penggunaan metode yang tangguh, walaupun terdapat keterbatasan dalam cakupan regional Papua dan data yang mungkin mempengaruhi hasil klasterisasi.

Meskipun penelitian ini memberikan pengetahuan yang sekiranya berguna, tentu masih ada keterbatasan dalam hal cakupan data dan metode analisis. Oleh karena itu, diperlukan penelitian tambahan untuk memperluas cakupan data dan menerapkan metode analisis yang lebih kompleks guna mendapatkan pemahaman yang lebih komprehensif tentang distribusi listrik di Indonesia. Dengan demikian, langkah-langkah strategis yang lebih efektif dapat dirancang untuk meningkatkan pemerataan distribusi listrik dan mendukung pembangunan berkelanjutan di seluruh negeri.

REFERENSI

- [1] Astria, Cici, et al. "Metode K-Means Pada Pengelompokan Wilayah Pendistribusian Listrik." *Seminar Nasional Sains dan Teknologi Informasi (SENSASI)*. Vol. 2. No. 1. 2019.
- [2] Herlambang, Amanda Austin, Muhammad Ary Murti, and Casi Setianingsih. "Pengelompokan Data Penggunaan Energi Listrik Menggunakan Algoritma Mini Batch K-means Clustering." *eProceedings of Engineering* 9.5 (2022).
- [3] A. K. Poluakan, R. F. Runtuwene, and S. A. P. Sambul, "Pengaruh Kompensasi Terhadap Kinerja Pegawai PT. PLN (Persero) UP3 Manado," *J. Adm. Bisnis*, vol. 9, no. 2, p. 70, 2019, doi: 10.35797/jab.9.2.2019.25114.70-77.
- [4] H. Hendrocahyo and L. Kurniawati, "Understanding the Financial Performance of PT PLN (Persero): A Narrative on State-Owned Enterprise (SOE) with a Mandate of Electricity in Indonesia," *Binus Bus. Rev.*, vol. 13, no. 3, pp. 241–258, 2022, doi: 10.21512/bbr.v13i3.7883.
- [5] Wang, T., Ren, C., Luo, Y., & Tian, J. (2019). NS-DBSCAN: A density-based clustering algorithm in network space. *ISPRS International Journal of Geo-Information*, 8(5), 218.
- [6] Rizuan, Rizuan, et al. "Penerapan Algoritma Mean-Shift Pada Clustering Penerimaan Bantuan Pangan Non Tunai." *Journal of Computer System and Informatics (JoSYC)* 4.4 (2023): 1019-1027.457
- [7] R. Yamasaki dan T. Tanaka, "Properties of Mean Shift," *IEEE Trans Pattern Anal Mach Intell*, vol. 42, no. 9, hlm. 2273–2286, Sep 2020, doi: 10.1109/TPAMI.2019.2913640.
- [8] J. Chen, J. Yang, J. Huang, dan Y. Liu, "Robust truth discovery scheme based on mean shift clustering algorithm," *Journal of Internet Technology*, vol. 22, no. 4, hlm. 835–842, 2021, doi: 10.53106/160792642021072204011.
- [9] Reliovani, Ryan, et al. "Mean Shift Algorithm to Determine Customer Segmentation in Online Store Sales." *Gunung Djati Conference Series*. Vol. 3. 2021.
- [10] Caesar, Rafli Agil, Desi Apriyanty, and Andre Mariza Putra. "Implementasi Mean Shift Clustering

- Dalam Mengelompokkan Pelanggan Retribusi Alat Pemadam Kebakaran Pada Dinas Pemadam Kebakaran dan Penanggulangan Bencana Kota Palembang." *Jurnal Sistem Informasi (JASISFO)* 4.2 (2023).
- [11] Simanjuntak, Krisman Pratama, and Ulfa Khaira. "Pengelompokkan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering: Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 1.1 (2021): 7-16.
 - [12] Prasetya, Indra. *Metodologi Penelitian Pendekatan Teori dan Praktik*. umsu press, 2022.
 - [13] Wijaya, C., Irsyad, H., & Widhiarso, W. (2020). Klasifikasi Pneumonia Menggunakan Metode K-Nearest Neighbor Dengan Ekstraksi GLCM. *Jurnal Algoritme*, 1(1), 33-44.
 - [14] Marcelina, Dona, Annisa Kurnia, and Terttiaavini Terttiaavini. "Analisis Klaster Kinerja Usaha Kecil dan Menengah Menggunakan Algoritma K-Means Clustering: Cluster Analysis of Small Medium Enterprise Performance with K-Means Clustering Algorithm." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 3.2 (2023): 293-301.
 - [15] Luchia, Nanda Try, et al. "Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia: Comparison of K-Means and K-Medoids on Poor Data Clustering in Indonesia." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 2.2 (2022): 35-41
 - [16] Isnain, Auliya Rahman, et al. "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm." *Jurnal Data Mining Dan Sistem Informasi* 2.1 (2021): 31-37.
 - [17] Putra, Febrianda, et al. "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia: Application of K-Nearest Neighbor Algorithm Using Wrapper as Preprocessing for Determination of Human Weight Information." *MALCOM: Indonesian Journal of Machine Learning and Computer Science* 4.1 (2024): 273-281.
 - [18] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, "Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Center," *Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Asessment Cent. Untuk Clust. Progr. Sdp*, vol. 3, no. 1, pp. 87–93, 2018.
 - [19] T. Alfina and B. Santosa, "Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," *Anal. PerbandinganMetode Hierarchical Clust. K-means dan Gabungan Keduanya dalam Clust. Data*, vol. 1, no. 1, pp. 1–5, 2012.
 - [20] Cinderatama, Toga Aldila, Rinanza Zulmy Alhamri, and Yoppy Yunhasnawa. "Implementasi Metode K-Means, Dbscan, dan Meanshift Untuk Analisis Jenis Ancaman Jaringan Pada Intrusion Detection System." *INOVTEK Polbeng-Seri Informatika* 7.1 (2022): 169-184.
 - [21] Nugraha, M. F., Martano, M., & Hayati, U. (2024). Clustering Data Indonesian Food Delivery Menggunakan Metode K-Means Pada Gofood Product List. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 3484-3492.
 - [22] M. Farid, F. Insani, S. Agustian, and L. Afriyanti, "Pengelompokkan Data Pendistribusian Listrik Menggunakan Algoritma Density Based Spatial Clustering Of Application With Noise (DBSCAN)," 2024, vol. 3.
 - [23] S. Mulyadi, F. Insani, S. Agustian, and L. Afriyanti, "Pengelompokkan Data Pendistribusian Listrik Menggunakan Algoritma Mini Batch K-Means," vol. 3, 2024.