

# Algoritma K-Means Dalam Mengelompokkan Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi

Mhd Gading Sadewo, Anggi Eriza, Agus Perdana Windarto, Dedy Hartama

Prodi Sistem Informasi STIKOM Tunas Bangsa Pematangsiantar, Pematangsiantar, Indonesia

Email: gading.sadewoo@gmail.com<sup>1</sup>

## Abstrak

Lampu jalan atau dikenal juga sebagai Penerangan Jalan Umum (PJU) adalah lampu yang digunakan untuk penerangan jalan dimalam hari sehingga mempermudah pejalan kaki, pesepeda dan pengendara kendaraan dapat melihat dengan lebih jelas jalan/medan yang akan dilalui pada malam hari, sehingga dapat meningkatkan keselamatan lalu lintas dan keamanan dari para pengguna jalan dari kegiatan/aksi kriminal. Penerangan jalan yang baik akan menghalangi pelaku kejahatan yang mengambil manfaat dari kegelapan malam. Penelitian ini membahas tentang Algoritma K-Means Dalam Mengelompokkan Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi Sumber data penelitian ini dikumpulkan berdasarkan dokumen-dokumen keterangan Banyaknya Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama yang dihasilkan oleh Badan Pusat Statistik Nasional. Data yang digunakan dalam penelitian ini adalah data provinsi yang terdiri dari 34 provinsi. Variable yang digunakan ada 4 yaitu Keluarga Pengguna Listrik PLN, Keluarga Pengguna Listrik Non PLN, Sumber Penerangan Jalan Utama Desa Listrik Pemerintah, dan Sumber Penerangan Jalan Utama Desa Listrik Non Pemerintah. Data akan diolah dengan melakukan clustering dalam 3 clushter yaitu clushter tingkat penerangan tinggi, clushter tingkat penerangan sedang dan tingkat penerangan rendah. Hasil yang diperoleh dari proses penilaian berdasarkan indeks Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama dengan 5 provinsi tingkat penerangan tinggi yaitu Aceh, Sumatera Utara, Jawa Barat, Jawa Tengah dan Jawa Timur, 13 Provinsi tingkat penerangan sedang, dan 16 Provinsi lainnya termasuk tingkat penerangan tinggi. Hal ini dapat menjadi masukan kepada pemerintah, provinsi yang menjadi perhatian lebih pada Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama berdasarkan cluster yang telah dilakukan.

**Kata Kunci:** Clustering, K-Means, Data Mining

## 1. PENDAHULUAN

Penerangan Jalan Umum (PJU) digunakan untuk penerangan jalan dimalam hari sehingga mempermudah pejalan kaki, pesepeda pengendara kendaraan dapat melihat dengan lebih jelas jalan/medan yang akan dilalui pada malam hari, sehingga dapat meningkatkan keselamatan lalu lintas dan keamanan dari para pengguna jalan dari kegiatan/aksi criminal. Didalam ilmu komputer banyak cabang ilmu yang dapat menyelesaikan masalah yang kompleks. Hal ini dibuktikan dengan beberapa penelitian dibidang jaringan saraf tiruan [1][2]–[5][6], bidang sistem pendukung keputusan [7][8][9][10]–[13], bidang datamining [14][15][16]–[20]. Berdasarkan penjelasan tersebut peneliti menggunakan datamining dalam menyelesaikan masalah diatas.

Analisis cluster merupakan suatu teknik multivariate dengan tujuan utama mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki. Sekarang ini analisis klaster telah banyak diaplikasikan dalam berbagai bidang yang ditulis dalam berbagai penelitian dan jurnal[16]. Data yang digunakan dalam penelitian ini berdasarkan dokumen-dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>. Dalam hal ini peneliti mengangkat topik mengelompokkan Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Desa dimana proses algoritma yang dilakukan adalah clustering. Hasil dari cluster dapat dijadikan masukan bagi pemerintah agar provinsi yang masuk kedalam clushter rendah mendapat perhatian lebih. Proses clushter dibagi kedalam 3 (tiga) cluster yakni cluster tingkat pengguna listrik tinggi, cluster tingkat pengguna listrik tinggi sedang dan cluster tingkat pengguna listrik tinggi rendah.

## 2. METODOLOGI PENELITIAN

### 2.1 Data Mining

Data Mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran computer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis[16]Defenisi lain diantaranya adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining ini juga dikenal dengan istilah pattern recognition[21]. Data mining disebut juga dengan Knowledge Discovery in Database (KDD) ataupun pattern recognition. Istilah KDD atau disebut penemuan pengetahuan data karena tujuan utama data mining adalah untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. Sedangkan istilah pattern recognition atau disebut pengenalan pola mempunyai tujuan pengetahuan yang akan digali dari dalam bongkahan data yang sedang dihadapi[22].

### 2.2 Clustering

Analisis Pengelompokan / Clushtering merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok

lain. Potensi clustering adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola[23]

Teknik cluster mempunyai dua metode dalam pengelompokannya yaitu hierarchical clustering dan non-hierarchical clustering. hierarchical clustering merupakan suatu metode pengelompokan data yang cara kerjanya dengan mengelompokkan dua data atau lebih yang mempunyai kesamaan atau kemiripan, kemudian proses dilanjutkan ke objek lain yang memiliki kedekatan dua, proses ini terus berlangsung hingga cluster membentuk semacam tree dimana ada hirarki atau tingkatan yang jelas antar objek dari yang paling mirip hingga yang paling tidak mirip. Namun secara logika semua objek pada akhirnya hanya akan membentuk sebuah cluster[23].

### 2.3 K-Means

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok[24].

Algoritma K-Means dimulai dengan pemilihan secara acak K, K disini merupakan banyaknya cluster yang ingin dibentuk. Kemudian tetapkan nilai-nilai K secara random, untuk sementara nilai tersebut menjadi pusat dari cluster atau biasa disebut dengan centroid secara random, mean atau “means”. Hitung jarak setiap data yang ada terhadap masing-masing centroid menggunakan rumus Euclidean hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid. Klasifikasikan setiap data berdasarkan kedekatannya dengan centroid. Lakukan langkah tersebut hingga nilai centroid tidak berubah [25].

Jarak Euclidean yang dirumuskan sebagai berikut:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (1)$$

dimana:

De adalah Euclidean Distance

i adalah banyaknya objek,

(x,y) merupakan koordinat object dan

(s,t) merupakan koordinat centroid.

### 2.4 Metodologi

Kerangka kerja dalam penelitian ini dapat dijelaskan sebagai berikut :

1. Mengumpulkan Data  
Pada tahap pengumpulan data ini, penulis memperoleh data dari Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>. Data yang digunakan dalam penelitian ini adalah data Banyaknya Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Desa. <sup>1</sup> Desa pada tabel ini termasuk nagari, Unit Permukiman Transmigrasi (UPT), dan Satuan Permukiman Transmigrasi (SPT) yang masih dibina oleh kementerian terkait.
2. Mempelajari Literatur.  
Penulis melakukan pengumpulan pengetahuan dari berbagai macam sumber literatur berupa buku-buku, jurnal dan karya ilmiah yang berkaitan dengan topik yang penulis angkat.
3. Menganalisa Data.  
Pada tahap ini penulis melakukan analisa data menggunakan tahapan KDD (Knowledge Discovery in Database). Di mana tahapan KDD ini meliputi : Seleksi Data, Preprocessing dan Pembersihan Data, Transformasi Data, Data Mining, Evaluasi/Interpretation.
4. Menyeleksi Data (Data Selection)  
Pada proses ini dilakukan pemilihan himpunan data, menciptakan himpunan data target, atau memfokuskan pada subset variable (sampel data) di mana penemuan (Discovery) akan dilakukan. Hasil seleksi disimpan dalam suatu berkas yang terpisah dari basis data operasional Preprocessing.
5. Melakukan Preprocessing dan Pembersihan Data (Cleaning Data)  
Pre processing dan Cleaning Data dilakukan dengan membuang data yang tidak konsisten dan noise, duplikasi data, memperbaiki kesalahan data dan bisa diperkaya dengan data eksternal yang relevan.
6. Metransformasi Data (Data Transformation)  
Proses ini mentransformasikan atau menggabungkan data ke dalam yang lebih tepat untuk melakukan proses Mining dengan cara melakukan peringkasan (Agregasi).
7. Merancang Struktur K Means Clustering  
Proses Data Mining yaitu proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik, metode atau algoritma tertentu sesuai dengan tujuan dari proses KDD secara keseluruhan. Algoritma yang digunakan dalam penelitian ini menggunakan K Means Clustering.
8. Mengevaluasi Data (Evaluation Data)  
Tahap selanjutnya yaitu mengevaluasi data. Tahap ini digunakan untuk menguji kualitas dari data apakah pola atau informasi yang ditemukan bersesuaian atau bertentangan dengan fakta sebelumnya.
9. Menguji Hasil

Tahap pengujian hasil adalah suatu teknik yang digunakan untuk menentukan bahwa penerapan metode yang digunakan telah mampu memecahkan masalah. Pada penelitian ini dilakukan pengujian dengan menggunakan aplikasi Ms. Excel

10. Menarik Kesimpulan

Setelah melakukan analisa data menggunakan tahapan KDD maka dapat ditarik suatu kesimpulan apakah informasi yang ditemukan berkesesuaian dengan fakta sebelumnya atau tidak, sehingga akan diperoleh pengetahuan baru.

11. Menerapkan Pengetahuan.

Langkah selanjutnya yaitu menerapkan pengetahuan yang diperoleh, sehingga diharapkan dapat membantu pemerintah dalam proses pembuatan keputusan yang akan datang.

### 3. ANALISA DAN PEMBAHASAN

Data yang dianalisis yaitu data Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Desa yang diperoleh dari Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>.

**Tabel 1.** Desa/ Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Desa

Provinsi	Keluarga Pengguna Listrik			Sumber Penerangan Jalan Utama Desa			Jumlah
	PLN	Non PLN	Tidak Ada	Listrik Pemerintah	Listrik Non Pemerintah	Non Listrik	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Aceh	6,427	296	11	2,539	1,119	5	3,663
Sumatera Utara	5,543	1,475	36	2,820	836	6	3,662
Sumatera Barat	1,099	350	2	731	130	1	862
Riau	1,301	1,194	1	680	349	7	1,036
Jambi	1,339	613	3	595	187	2	784
Sumatera Selatan	2,886	1,123	-	1,508	570	8	2,086
Bengkulu	1,470	244	10	501	188	4	693
Lampung	2,402	779	-	1,095	605	1	1,701
Kepulauan Bangka Belitung	377	153	-	291	16	-	307
Kepulauan Riau	294	293	-	215	42	-	257
DKI Jakarta	267	2	-	264	-	-	264
Jawa Barat	5,960	257	-	4,364	695	5	5,064
Jawa Tengah	8,566	115	-	5,478	2,846	6	8,330
DI Yogyakarta	438	9	-	280	148	-	428
Jawa Timur	8,457	291	-	5,557	2,491	7	8,055
Banten	1,551	34	-	716	234	-	950
Bali	716	20	-	492	208	-	700
Nusa Tenggara Barat	1,114	122	3	688	151	1	840
Nusa Tenggara Timur	2,624	1,694	20	249	49	-	298
Kalimantan Barat	1,380	1,239	15	444	73	4	521
Kalimantan Tengah	838	1,079	1	258	160	3	421
Kalimantan Selatan	1,903	401	6	1,122	510	2	1,634
Kalimantan Timur	647	662	-	351	105	6	462
Kalimantan Utara	180	380	7	95	30	8	133
Sulawesi Utara	1,789	258	-	843	288	1	1,132
Sulawesi Tengah	1,601	897	1	903	350	4	1,257
Sulawesi Selatan	2,777	734	2	1,887	277	1	2,165
Sulawesi Tenggara	1,786	896	11	497	286	2	785
Gorontalo	690	298	-	320	214	-	534
Sulawesi Barat	403	440	1	148	36	-	184
Maluku	654	540	12	136	225	5	366
Maluku Utara	785	598	1	276	175	2	453
Papua Barat	443	914	262	118	242	4	364
Papua	824	2,093	2,114	203	187	22	412

(Sumber : Badan Pusat Statistik , url : <https://www.bps.go.id>)

Dari data awal kolom no 4 dan kolom no 7 tidak digunakan karena tidak sesuai dengan kebutuhan proses perhitungan data mining. Sehingga kolom yang digunakan adalah kolom PLN, Non PLN, Listrik Pemerintah, dan Listrik Non Pemerintah. Untuk hasil selengkapnya dapat dilihat pada table berikut :

**Tabel 2.** Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Desa Yang Sudah Diseleksi

Provinsi	Keluarga Pengguna Listrik		Sumber Penerangan Jalan Utama Desa	
	PLN	Non PLN	Listrik Pemerintah	Listrik Non Pemerintah
Aceh	6,427	296	2,539	1,119
Sumatera Utara	5,543	1,475	2,820	836
Sumatera Barat	1,099	350	731	130
Riau	1,301	1,194	680	349
Jambi	1,339	613	595	187
Sumatera Selatan	2,886	1,123	1,508	570
Bengkulu	1,470	244	501	188
Lampung	2,402	779	1,095	605
Kepulauan Bangka Belitung	377	153	291	16
Kepulauan Riau	294	293	215	42
DKI Jakarta	267	2	264	0
Jawa Barat	5,960	257	4,364	695

Jawa Tengah	8,566	115	5,478	2,846
DI Yogyakarta	438	9	280	148
Jawa Timur	8,457	291	5,557	2,491
Banten	1,551	34	716	234
Bali	716	20	492	208
Nusa Tenggara Barat	1,114	122	688	151
Nusa Tenggara Timur	2,624	1,694	249	49
Kalimantan Barat	1,380	1,239	444	73
Kalimantan Tengah	838	1,079	258	160
Kalimantan Selatan	1,903	401	1,122	510
Kalimantan Timur	647	662	351	105
Kalimantan Utara	180	380	95	30
Sulawesi Utara	1,789	258	843	288
Sulawesi Tengah	1,601	897	903	350
Sulawesi Selatan	2,777	734	1,887	277
Sulawesi Tenggara	1,786	896	497	286
Gorontalo	690	298	320	214
Sulawesi Barat	403	440	148	36
Maluku	654	540	136	225
Maluku Utara	785	598	276	175
Papua Barat	443	914	118	242
Papua	824	2,093	203	187

(Sumber : Badan Pusat Statistik , url : <https://www.bps.go.id>)

Kemudian data tersebut akan masuk ke tahapan clustering dengan menerapkan algoritma K-Means untuk mengcluster data

### 3.1 Proses Pengelompokan Algoritma K-Means

#### a. Menentukan Jumlah Cluster

Pada pengelompokan Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama ini dibuat menjadi 3 cluster. Penentuan cluster tersebut terdiri dari 4 variabel yaitu Keluarga Pengguna Listrik PLN, Keluarga Pengguna Listrik Non PLN, Sumber Penerangan Jalan Utama Desa Listrik Pemerintah, dan Sumber Penerangan Jalan Utama Desa Listrik Non Pemerintah. Dengan cluster sebagai berikut :

1. Cluster Tingkat Pengguna Listrik Tinggi
2. Cluster Tingkat Pengguna Listrik Sedang
3. Cluster Tingkat Pengguna Listrik Rendah.

#### b. Menentukan Titik Pusat Awal Cluster (Centroid)

Dalam penerapan algoritma K-means dihasilkan nilai titik tengah atau centroid dari data yang didapat dengan ketentuan bahwa clusterisasi yang diinginkan adalah 3, Penentuan cluster dibagi atas tiga bagian yakni cluster tingkat pengguna listrik tinggi (C1), cluster tingkat pengguna listrik sedang (C2) dan cluster tingkat pengguna listrik rendah (C3). maka nilai titik tengah atau centroid juga terdapat 3 titik. Penentuan titik cluster ini dilakukan dengan mengambil nilai terbesar (maksimum) untuk cluster tingkat pengguna listrik tinggi (C1), nilai rata-rata (average) untuk cluster tingkat pengguna listrik sedang (C2) dan nilai terkecil (minimum) untuk cluster tingkat pengguna listrik rendah (C3). Nilai titik tersebut dapat diketahui pada Tabel 3 berikut:

**Tabel 3. Centroid Data Iterasi 1**

	PLN	Non PLN	Listrik Pemerintah	Listrik Non Pemerintah
C1	8,566	2,093	5,557	2,846
C2	2,045	603	1,078	412
C3	180	2	95	0

#### c. Clustering Data

Dengan menggunakan centroid tersebut maka dapat dicluster data yang telah didapat menjadi 3 cluster. Proses cluster dengan mengambil jarak terdekat dari setiap data yang diolah. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 1 dan Clustering dapat digambarkan pada tabel dan gambar berikut:

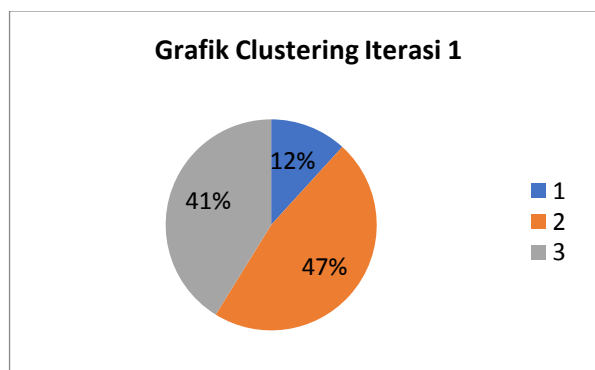
**Tabel 4. Perhitungan Jarak Pusat Cluster Iterasi 1**

Provinsi	PLN	Non PLN	Listrik Pemerintah	Listrik Non Pemerintah	C1	C2	C3	Jarak Terpendek
Aceh	6,427	296	2,539	1,119	4460.42	4682.79	6807.11	4460.42
Sumatera Utara	5,543	1,475	2,820	836	4588.22	4026.09	6249.48	4026.09
Sumatera Barat	1,099	350	731	130	9458.39	1076.69	1177.74	1076.69
Riau	1,301	1,194	680	349	9143.77	1032.41	1772.44	1032.41
Jambi	1,339	613	595	187	9279.64	884.89	1414.77	884.89
Sumatera Selatan	2,886	1,123	1,508	570	7401.21	1089.65	3301.60	1089.65
Bengkulu	1,470	244	501	188	9295.17	918.18	1386.67	918.18
Lampung	2,402	779	1,095	605	8040.71	442.57	2628.12	442.57
Kepulauan Bangka Belitung	377	153	291	16	10322.93	1939.50	316.67	316.67
Kepulauan Riau	294	293	215	42	10395.45	2011.12	337.40	337.40
DKI Jakarta	267	2	264	0	10457.59	2086.99	190.08	190.08

Jawa Barat	5,960	257	4,364	695	4026.44	5130.48	7223.63	4026.44
Jawa Tengah	8,566	115	5,478	2,846	1979.58	8248.65	10364.08	1979.58
DI Yogyakarta	438	9	280	148	10272.94	1908.49	350.35	350.35
Jawa Timur	8,457	291	5,557	2,491	1839.87	8098.73	10228.92	1839.87
Banten	1,551	34	716	234	9149.18	854.78	1523.50	854.78
Bali	716	20	492	208	9926.38	1578.45	698.92	698.92
Nusa Tenggara Barat	1,114	122	688	151	9507.23	1148.32	1123.03	1123.03
Nusa Tenggara Timur	2,624	1,694	249	49	8453.68	1531.65	2976.93	1531.65
Kalimantan Barat	1,380	1,239	444	73	9284.41	1168.20	1759.91	1168.20
Kalimantan Tengah	838	1,079	258	160	9800.21	1555.78	1282.60	1282.60
Kalimantan Selatan	1,903	401	1,122	510	8507.91	268.88	2107.79	268.88
Kalimantan Timur	647	662	351	105	9968.64	1606.72	854.55	854.55
Kalimantan Utara	180	380	95	30	10536.73	2154.34	379.19	379.19
Sulawesi Utara	1,789	258	843	288	8835.16	505.24	1815.73	505.24
Sulawesi Tengah	1,601	897	903	350	8822.21	564.27	1896.21	564.27
Sulawesi Selatan	2,777	734	1,887	277	7445.00	1106.91	3250.88	1106.91
Sulawesi Tenggara	1,786	896	497	286	8919.55	712.08	1903.12	712.08
Gorontalo	690	298	320	214	9980.33	1594.81	666.44	666.44
Sulawesi Barat	403	440	148	36	10320.87	1931.32	495.66	495.66
Maluku	654	540	136	225	10063.22	1691.76	752.61	752.61
Maluku Utara	785	598	276	175	9889.50	1512.56	885.79	885.79
Papua Barat	443	914	118	242	10185.12	1901.24	979.80	979.80
Papua	824	2,093	203	187	9781.32	2128.11	2198.56	2128.11

**Tabel 5. Hasil Pengelompokan Iterasi 1**

Provinsi	C1	C2	C3
Aceh	1		
Sumatera Utara		1	
Sumatera Barat		1	
Riau		1	
Jambi		1	
Sumatera Selatan		1	
Bengkulu		1	
Lampung		1	
Kepulauan Bangka Belitung			1
Kepulauan Riau			1
DKI Jakarta			1
Jawa Barat	1		
Jawa Tengah	1		
DI Yogyakarta			1
Jawa Timur	1		
Banten		1	
Bali			1
Nusa Tenggara Barat			1
Nusa Tenggara Timur		1	
Kalimantan Barat		1	
Kalimantan Tengah			1
Kalimantan Selatan		1	
Kalimantan Timur			1
Kalimantan Utara			1
Sulawesi Utara	1		
Sulawesi Tengah	1		
Sulawesi Selatan	1		
Sulawesi Tenggara		1	
Gorontalo			1
Sulawesi Barat			1
Maluku			1
Maluku Utara			1
Papua Barat			1
Papua		1	



**Gambar 1. Grafik Clustering Iterasi 1**

Proses K-Means akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Pada iterasi 1 diperoleh cluster data Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama, yang dapat dilihat pada gambar 1. Proses iterasi tersebut berhenti pada iterasi ke 3, pada iterasi 3 akan dilakukan proses mencari nilai titik tengah atau centroid yang dapat diketahui pada table berikut

**Tabel 6.** Centroid Data Iterasi 3

	PLN	Non PLN	Listrik Pemerintah	Listrik Non Pemerintah
C1	6990.6	486.8	4151.6	1597.4
C2	1908.385	777.3846	849.2308	305.0769
C3	610.5625	497.0625	304.125	129.3125

Setelah mendapatkan nilai titik tengah atau centroid, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 3 dan Clustering data dapat digambarkan Pada tabel dan gambar berikut:

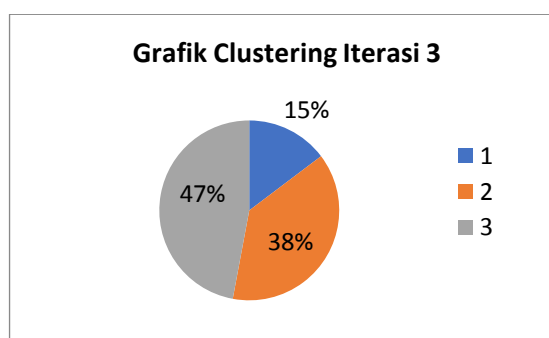
**Tabel 7.** Perhitungan Jarak Pusat Cluster Iterasi 1

Provinsi	PLN	Non PLN	Listrik Pemerintah	Listrik Non Pemerintah	C1	C2	C3	Jarak Terpendek
Aceh	6,427	296	2,539	1,119	1784.2071	4916.036	6312.331	1784.207
Sumatera Utara	5,543	1,475	2,820	836	2329.1573	4226.453	5666.951	2329.157
Sumatera Barat	1,099	350	731	130	6970.1816	939.357	665.1476	665.1476
Riau	1,301	1,194	680	349	6817.7848	757.0025	1073.299	757.0025
Jambi	1,339	613	595	187	6826.0633	655.5887	794.9835	655.5887
Sumatera Selatan	2,886	1,123	1,508	570	5029.5788	1256.719	2685.691	1256.719
Bengkulu	1,470	244	501	188	6771.2085	782.0818	919.1721	782.0818
Lampung	2,402	779	1,095	605	5609.6606	627.7062	2034.822	627.7062
Kepulauan Bangka Belitung	377	153	291	16	7826.6331	1769.232	431.2104	431.2104
Kepulauan Riau	294	293	215	42	7924.5251	1819.979	396.762	396.762
DKI Jakarta	267	2	264	0	7943.9882	1931.562	617.6195	617.6195
Jawa Barat	5,960	257	4,364	695	1405.1277	5402.965	6743.641	1405.128
Jawa Tengah	8,566	115	5,478	2,846	2436.8954	8523.168	9878.478	2436.895
DI Yogyakarta	438	9	280	148	7762.4034	1761.006	518.5692	518.5692
Jawa Timur	8,457	291	5,557	2,491	2227.6303	8370.322	9735.466	2227.63
Banten	1,551	34	716	234	6592.1531	838.5385	1131.128	838.5385
Bali	716	20	492	208	7410.237	1460.291	529.334	529.334
Nusa Tenggara Barat	1,114	122	688	151	6982.5572	1053.713	736.1736	736.1736
Nusa Tenggara Timur	2,624	1,694	249	49	6176.7594	1333.47	2344.372	1333.47
Kalimantan Barat	1,380	1,239	444	73	6936.4782	842.8231	1079.463	842.8231
Kalimantan Tengah	838	1,079	258	160	7445.2285	1267.792	627.2546	627.2546
Kalimantan Selatan	1,903	401	1,122	510	6020.9593	508.0265	1579.071	508.0265
Kalimantan Timur	647	662	351	105	7546.1099	1375.743	176.9758	176.9758
Kalimantan Utara	180	380	95	30	8081.3628	1946.73	502.6768	502.6768
Sulawesi Utara	1,789	258	843	288	6306.3741	533.2387	1327.19	533.2387
Sulawesi Tengah	1,601	897	903	350	6428.4882	337.1978	1244.295	337.1978
Sulawesi Selatan	2,777	734	1,887	277	4968.6419	1354.3	2697.572	1354.3
Sulawesi Tenggara	1,786	896	497	286	6506.2414	391.7627	1265.921	391.7627
Gorontalo	690	298	320	214	7505.2089	1415.15	230.9982	230.9982
Sulawesi Barat	403	440	148	36	7865.4591	1715.849	281.8165	281.8165
Maluku	654	540	136	225	7626.5231	1464.562	202.8609	202.8609
Maluku Utara	785	598	276	175	7454.2158	1280.502	208.555	208.555
Papua Barat	443	914	118	242	7820.5245	1644.595	499.2545	499.2545
Papua	824	2,093	203	187	7628.0734	1827.099	1614.35	1614.35

**Tabel 8.** Hasil Pengelompokan Iterasi 3

Provinsi	C1	C2	C3
Aceh	1		
Sumatera Utara	1		
Sumatera Barat			1
Riau		1	
Jambi		1	
Sumatera Selatan		1	
Bengkulu		1	
Lampung		1	
Kepulauan Bangka Belitung			1
Kepulauan Riau			1
DKI Jakarta			1
Jawa Barat	1		
Jawa Tengah	1		
DI Yogyakarta			1
Jawa Timur	1		
Banten		1	

Bali	1
Nusa Tenggara Barat	1
Nusa Tenggara Timur	1
Kalimantan Barat	1
Kalimantan Tengah	1
Kalimantan Selatan	1
Kalimantan Timur	1
Kalimantan Utara	1
Sulawesi Utara	1
Sulawesi Tengah	1
Sulawesi Selatan	1
Sulawesi Tenggara	1
Gorontalo	1
Sulawesi Barat	1
Maluku	1
Maluku Utara	1
Papua Barat	1
Papua	1



Gambar 2. Grafik Clustering Iterasi 3

### 3.2 Analisa Data

Pada iterasi 4, pengelompokan data yang dilakukan terhadap 3 cluster dengan iterasi 3 didapatkan hasil yang sama. Dari 34 data Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama dapat diketahui, 4 provinsi cluster tingkat tinggi yakni Sumatera Utara, Jawa Barat, Jawa Tengah, Jawa Timur, 14 provinsi cluster tingkat sedang dan 16 provinsi lainnya termasuk cluster tingkat rendah.

## 4. KESIMPULAN

Berdasarkan penelitian ini didapat kesimpulan sebagai berikut:

1. Telah berhasil dilakukan pengelompokan data Desa<sup>1</sup> / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama menggunakan algoritma K-means clustering menjadi 3 cluster, yaitu untuk cluster 1 berjumlah 5 data yaitu provinsi Aceh, Sumatera Utara, Jawa Barat, Jawa Tengah dan Jawa Timur, cluster 2 berjumlah 13 Provinsi, dan cluster 3 berjumlah 16 Provinsi
2. Penentuan titik pusat awal (centroid) sangat berpengaruh terhadap jumlah iterasi yang akan dihitung.
3. Berdasarkan penelitian ini dapat dijadikan acuan yang dapat dilakukan pemerintah dalam hal membangun sumber penerangan jalan utama.

## UCAPAN TERIMA KASIH

Terima kasih kepada program studi sistem informasi, STIKOM Tunas Bangsa Pematangsiantar atas dukungan dalam terlaksanakannya penelitian ini serta bapak pembimbing Agus Perdana Windarto.

## REFERENCES

- [1] A. P. Windarto, M. R. Lubis, and L. R. Komprehensif, "IMPLEMENTASI JST PADA PREDIKSI TOTAL LABA RUGI KOMPREHENSIF IMPLEMENTATION OF NEURAL NETWORK IN PREDICTING TOTAL COMPREHENSIVE INCOME OF CONVENTIONAL COMMERCIAL BANKS USING," vol. 5, no. 4, pp. 411–418, 2018.
- [2] A. P. Windarto, M. R. Lubis, and Solikhun, "IMPLEMENTASI JST PADA PREDIKSI TOTAL LABA RUGI KOMPREHENSIF BANK UMUM KONVENSIONAL DENGAN BACKPROPAGATION," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 4, pp. 411–418, 2018.
- [3] Solikhun, A. P. Windarto, Handrizal, and M. Fauzan, "Jaringan Saraf Tiruan Dalam Memprediksi Suku Negara Ritel Berdasarkan Kelompok Profesi Dengan Backpropogation Dalam Mendorong Laju Pertumbuhan Ekonomi," *Kumpul. J. Ilmu Komput.*, vol. 4, no. 2, pp. 184–197, 2017.
- [4] A. P. Windarto, L. S. Dewi, and D. Hartama, "Implementation of Artificial Intelligence in Predicting the Value of Indonesian Oil and Gas Exports With BP Algorithm," *Int. J. Recent Trends Eng. Res.*, vol. 3, no. 10, pp. 1–12, 2017.
- [5] Sumijan, A. P. Windarto, A. Muhammad, and Budiharjo, "Implementation of Neural Networks in Predicting the Understanding Level of Students Subject," *Int. J. Softw. Eng. Its Appl.*, vol. 10, no. 10, pp. 189–204, 2016.
- [6] A. P. Windarto, M. R. Lubis, and Solikhun, "MODEL ARSITEKTUR NEURAL NETWORK DENGAN BACKPROGATION PADA PREDIKSI TOTAL LABA RUGI KOMPREHENSIF BANK UMUM KONVENSIONAL," *Kumpul. J. Ilmu Komput.*, vol. 5, no. 2, pp. 147–158, 2018.
- [7] I. S. Damanik and A. P. Windarto, "PENERAPAN METODE TOPSIS DALAM MENENTUKAN PEMILIHAN KARTU PRABAYAR HANDPHONE GOBAL SYSTEM FOR MOBILE ( HP GSM )," vol. 2, pp. 231–237, 2018.
- [8] F. Syahputra, I. Lubis, and A. P. Windarto, "SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN GURU BERPRESTASI KOTA MEDAN MENERAPKAN METODE PREFERENCES SELECTION INDEX ( STUDI KASUS : DINAS PENDIDIKAN KOTA MEDAN )," vol. 2, pp. 147–

- 155, 2018.
- [9] S. R. Ningsih and A. P. Windarto, "Penerapan Metode Promethee II Pada Dosen Penerima Hibah P2M Internal," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 3, no. 1, pp. 20–25, 2018.
- [10] T. Imandasari, A. Wanto, and A. P. Windarto, "Analisis Pengambilan Keputusan Dalam Menentukan Mahasiswa PKL Menggunakan Metode PROMETHEE," *J. Ris. Komput.*, vol. 5, no. 3, pp. 234–239, 2018.
- [11] T. Novika, A. Widiastari, V. Miralda, and A. P. Windarto, "SPK: ANALISA REKOMENDASI BANK KONVENSIONAL DENGAN PROMETHEE SEBAGAI SOLUSI CERDAS UNTUK MENABUNG," *JUSTIM*, vol. 3, no. 1, pp. 38–45, 2018.
- [12] F. Adelia, D. Wahyuli, T. Imanda, and A. P. Windarto, "Analisis Promethee II Pada Faktor Penyebab Mahasiswa Sulit Menemukan Judul Artikel Ilmiah," *Jurnal Ilmiah KOMPUTASI*, vol. 17, no. 2, pp. 131–135, 2018.
- [13] T. Imandasari and A. P. Windarto, "Sistem Pendukung Keputusan dalam Merekomendasikan Unit Terbaik di PDAM Tirta Lihou Menggunakan Metode Promethee," *J. Teknol. dan Sist. Komput.*, vol. 5, no. 4, p. 159, 2017.
- [14] A. P. Windarto *et al.*, "MODEL ARSITEKTUR NEURAL NETWORK DENGAN BACKPROPAGATION PADA PREDIKSI TOTAL LABA," vol. 5, no. 2, pp. 147–158, 2018.
- [15] M. G. Sadewo *et al.*, "PENERAPAN ALGORITMA CLUSTERING DALAM MENGELOMPOKKAN BANYAKNYA DESA / KELURAHAN MENURUT UPAYA ANTISIPASI / MITIGASI BENCANA ALAM MENURUT PROVINSI DENGAN K-MEANS," vol. 2, pp. 311–319, 2018.
- [16] M. G. Sadewo, A. P. Windarto, and D. Hartama, "PENERAPAN DATAMINING PADA POPULASI DAGING AYAM RAS PEDAGING DI INDONESIA BERDASARKAN PROVINSI MENGGUNAKAN K-MEANS CLUSTERING," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 60–67, 2017.
- [17] H. Siahaan, H. Mawengkang, S. Efendi, A. Wanto, and A. P. Windarto, "Application of Classification Method C4 . 5 on Selection of Exemplary Teachers," in *IOP Conference Series*, 2018, pp. 1–6.
- [18] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method," *Int. J. Artif. Intell. Res.*, vol. 1, no. 2, pp. 26–33, 2017.
- [19] B. Supriyadi, A. P. Windarto, T. Soemartono, and Mungad, "Classification of natural disaster prone areas in Indonesia using K-means," *Int. J. Grid Distrib. Comput.*, vol. 11, no. 8, pp. 87–98, 2018.
- [20] S. Sudirman, A. P. Windarto, and A. Wanto, "Data Mining Tools | RapidMiner : K-Means Method on Clustering of Rice Crops by Province as Efforts to Stabilize Food Crops In Indonesia," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 420, no. 12089, pp. 1–8, 2018.
- [21] J. O. Ong, "Implementasi Algoritma K-means clustering untuk menentukan strategi marketing president university," *J. Ilm. Tek. Ind.*, vol. vol.12, no. no. juni, pp. 10–20, 2013.
- [22] D. B. B. N. Fauziah Nur, Prof. M. Zarlis, "PENERAPAN ALGORITMA K-MEANS PADA SISWA BARU SEKOLAHMENENGAH KEJURUAN UNTUK CLUSTERING JURUSAN," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, no. 9, pp. 100–105, 2015.
- [23] A. K. Wardhani, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Kajen Pekalongan," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- [24] R. W. Sari, A. Wanto, and A. P. Windarto, "IMPLEMENTASI RAPIDMINER DENGAN METODE K-MEANS ( STUDY KASUS : IMUNISASI CAMPAK PADA BALITA BERDASARKAN PROVINSI )," vol. 2, pp. 224–230, 2018.
- [25] A. I. G. Heni Sulastri, "Jurnal Teknologi dan Sistem Informasi PENERAPAN DATA MINING DALAM PENGELOMPOKAN PENDERITA THALASSAEMIA," *Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 299–305, 2017.