



## *Application of K-Means and K-Medoids Algorithms in Clustering of Densely Populated Areas in Riau Province*

### **Penerapan Algoritma K-Means dan K-Medoids Dalam Pengelompokan Kepadatan Penduduk Provinsi Riau**

**Nabiilah<sup>1\*</sup>, Farin Junita Fauzan<sup>2</sup>, Nurazizah Nurazizah<sup>3</sup>,  
Abdul Hamid<sup>4</sup>, Sania Fitri Octavia<sup>5</sup>**

<sup>1,2,3,4,5</sup>Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,  
Universitas Islam Negeri Sultan Syarif Kasim Riau

E-Mail: <sup>1</sup>12050322475@students.uin-suska.ac.id, <sup>2</sup>12050323918@students.uin-suska.ac.id,  
<sup>3</sup>12050321684@students.uin-suska.ac.id, <sup>4</sup>12050312122@students.uin-suska.ac.id,  
<sup>5</sup>saniahfitri2017@gmail.com

*Corresponding Author: Nabiilah*

#### **Abstract**

*The problem of unequal distribution of the population is one of the problems related to the population faced by Riau. The existence of an unequal distribution can cause problems at the level of density in an area. This study utilizes information sourced from the Central Bureau of Statistics (BPS). This study sampled data from 2010-2021 and consisted of 12 districts. For this reason, a comparison of the K-Means and K-Medoids algorithms was carried out by researchers where the aim was to find the best results. Based on the research conducted, the K-Means algorithm has a more satisfactory performance compared to K-Medoids in clustering Riau population density data. The results of this study are indicated by the resulting value, namely the DBI in the K-means algorithm of 0.001 for the K=2 experiment.*

*Keyword: Clusters, DBI, K-Means, K-Medoids, Population Density*

#### **Abstrak**

Masalah tidak seimbang nya pembagian penduduk menjadi salah satu persoalan memiliki keterkaitan dengan kependudukan yang dihadapi oleh Riau. Adanya pembagian yang tidak merata dapat menyebabkan masalah pada tingkat kepadatan di sebuah area. Penelitian ini memanfaatkan informasi yang bersumber dari Badan Pusat Statistik (BPS). Penelitian ini mengambil sampel data dari tahun 2010-2021 dan terdiri dari 12 kabupaten. Untuk itu, dilakukan perbandingan algoritma K-Means dan K-Medoids oleh peneliti dimana tujuannya adalah menemukan hasil terbaik. Berdasarkan penelitian yang dilakukan, algoritma K-Means memiliki kinerja yang lebih memuaskan dibandingkan dengan K-Medoids dalam melakukan klusterisasi data kepadatan penduduk Riau. Hasil penelitian ini ditunjukkan dengan nilai yang dihasilkan yaitu DBI pada algoritma K-means sebesar 0.001 untuk percobaan K=2.

Kata Kunci: Cluster, DBI, K-Means, K-Medoids, Kepadatan Penduduk

#### **1. PENDAHULUAN**

Kepadatan penduduk merupakan akumulasi dari rata-rata penduduk dari setiap wilayah satu kilometer persegi. Setiap wilayah memiliki kepadatan penduduk yang berbeda [1]. Ada dua faktor yang dapat mempengaruhi kepadatan penduduk seperti faktor internal daerah yaitu relatif tingginya pertumbuhan penduduk serta wilayah yang strategis, maupun faktor eksternal yaitu perpindahan penduduk massal yang menyebabkan perluasan kota yang dapat menyebabkan kesenjangan ekonomi [2], sosial [3], kesehatan, dan permasalahan lingkungan [4]. Persoalan terkait kependudukan yang dialami oleh Provinsi Riau ialah pesatnya

perkembangan penduduk [5] menyebabkan tingginya kebutuhan lahan pemukiman terutama di wilayah perkotaan dan tingginya tingkat kriminalitas [6].

Pemerintah di masing-masing daerah selayaknya mampu memastikan setiap warganya memiliki akses terhadap layanan yang memiliki kaitan dengan kebutuhan pokok secara menyeluruh demi pembangunan sumber daya manusia yang penting dalam pembangunan masa ini. Padatnya penduduk dapat memunculkan masalah baru yang dapat mempengaruhi volume lalu lintas[7], penurunan kualitas air [8], peningkatan kriminalitas dan pengangguran [9], hingga mudahnya penyebaran penyakit [10]. Masalah kepadatan penduduk dapat diselesaikan menggunakan pengolahan data mining [1].

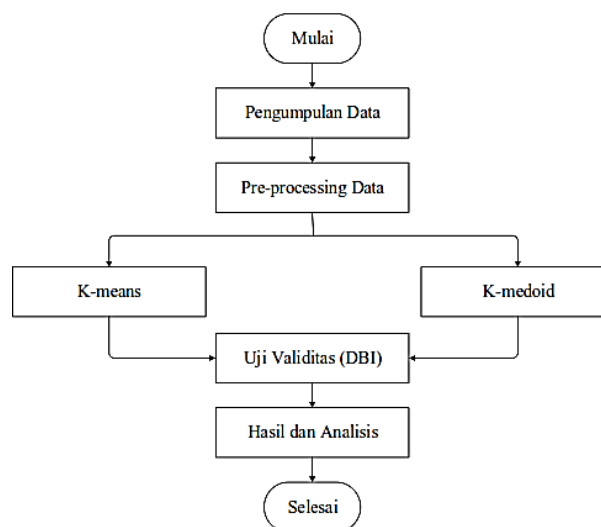
Data mining merupakan sebuah tahapan menggali pengetahuan dari pusat data [11]. Data dapat diklasifikasikan dalam berbagai bentuk pola dan aturan tertentu yang dapat membantu pemerintah untuk menganalisa data [12] dan memprediksi kebijakan yang akan dibuat kedepannya, perlu adanya pemetaan terhadap penduduk di setiap kabupaten ke dalam beberapa cluster [7]. Clustering ialah tahapan pengelompokan beberapa objek/titik dalam dua atau lebih kategori hingga data dalam kategori yang sama menjadi serupa satu dengan yang lainnya menurut informasi yang ada bersama poin data [13], [14]. K-Means dan K-Medoids merupakan algoritma dalam kelompok clustering. Tujuannya agar partisi keluaran memiliki sensitivitas yang menurun karena data set yang memiliki nilai ekstrem, sehingga setiap cluster yang diamati pada penggunaan medoid tidak berdasarkan pada rata-rata [15].

Dalam penelitian Luchia dkk [13] menyatakan bahwa algoritma K-Means memiliki performa lebih baik dari algoritma K-Medoids dalam meng-klasterisasi data penduduk miskin. Berdasarkan penelitian yang dilakukan oleh Gultom dkk [1] dengan menggunakan algoritma K-Means mendapatkan hasil 3 cluster daerah berkategori C1 untuk wilayah penduduk tinggi, C2 untuk wilayah penduduk sedang, dan C3 untuk wilayah penduduk rendah. Pada penelitian yang dilakukan oleh Sangga [16] K-Medoids lebih baik dari K-Means untuk data pengelompokan komoditas peternakan di Provinsi Jawa Tengah. Kemudian pada penelitian yang dilakukan oleh Arianto [17] pada kasus klasterisasi penduduk kurang mampu di Desa Sambirejo Timur menggunakan algoritma K-Medoids mendapatkan 2 cluster dari penelitiannya. Kemudian pada penelitian yang dilakukan oleh Herviany dkk [18] dalam klasterisasi daerah rawan tanah longsor dengan algoritma K-Means memiliki hasil lebih memuaskan dibandingkan dengan K-Medoids dengan nilai  $k = 6$  sebagai nilai paling optimal.

Berdasarkan penelitian terdahulu yang dibahas, peneliti dapat mengambil simpulan penelitian yang berbeda dalam penentuan pola cluster dan tingkat akurasi. Sehingga diperlukannya cara untuk menentukan metode clustering dengan cara pengujian performance algoritma. Maka didalam studi ini, peneliti membandingkan algoritma K-Means dan K-Medoids untuk menentukan metode yang dapat digunakan dalam hal mengelompokkan kepadatan penduduk Provinsi Riau. Penelitian ini untuk mengidentifikasi jumlah kepadatan penduduk dari setiap algoritma dalam klasterisasi. Penelitian ini diharapkan dapat membantu pemerintah dalam membagi wilayah kepadatan penduduk yang tinggi agar dapat menentukan kebijakan yang diambil untuk mengatasi permasalahan yang timbul seiring dengan bertambah padatnya penduduk di masing-masing wilayah yang ada di Provinsi riau.

## 2. METODE PENELITIAN

Pada penelitian ini metodologi penelitian dibagi menjadi lima alur, yaitu: pengumpulan data, pre-processing data, proses klasterisasi data, uji validitas klaster dan hasil analisis. Metodologi penelitian ini dapat dilihat pada Gambar 1:



**Gambar 1.** Metodologi Penelitian

**2.1 Pengumpulan Data**

Pengumpulan data memiliki tujuan untuk mengumpulkan berbagai data yang diperlukan dalam proses penelitian [19] dan dilakukan dengan mengambil data kepadatan penduduk pada website bps.co.id dari tahun 2010 sampai tahun 2021. Daftar kepadatan penduduk Provinsi Riau didapat melalui proses data yang telah didapat. Data ini dapat membantu pemerintah dalam menentukan kebijakan yang diambil untuk mengatasi permasalahan yang timbul seiring dengan bertambah padatnya penduduk. Dalam clusterisasi data, membutuhkan 13 atribut dari data BPS yang telah didapat. Atribut data ini ialah jumlah kepadatan penduduk setiap kabupaten yang ada di provinsi riau dari tahun 2010-2021. Data ini didapatkan berdasarkan jumlah penduduk dari perluas wilayah, jumlah penduduk dari perong dan luas wilayah perkilometer persegi. Acuan untuk pengelompokan berdasarkan atribut dari data yang didapat. Data dapat dilihat pada Tabel 1.

**Tabel 1.** Data Kepadatan Penduduk Provinsi Riau

No	Kabupaten/ Kota	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1	Kuantan Singingi	55,8	56,0	57,5	58,3	59,1	59,8	60,5	61,1	61,7	62,2	64	65
2	Indragiri Hulu	47,3	48,2	49,7	50,8	51,9	53,0	54,1	55,1	56,2	57,2	58	59
3	Indragiri Hilir	53,2	53,6	53,6	54,3	55,1	55,8	56,5	57,3	58,0	58,7	52	52
4	Pelalawan	23,9	25,8	26,6	28,1	29,6	31,1	32,7	34,4	36,1	37,9	31	31
...	...	...	...	...	...	...	...	...	...	...	...	...	...
10	Kepulauan Meranti	47,5	47,7	47,9	48,2	48,5	48,8	49,1	49,4	49,7	50,0	56	56
11	Pekanbaru	1428,2	1469,7	1515,7	1557,4	1599,7	1641,9	1683,7	1725,7	1767,2	1808,3	155	157
12	Dumai	157,1	160,1	165,1	168,8	172,5	176,2	179,8	183,3	186,8	190,2	195	199

**2.2 Pre-processing Data**

Normalisasi data merupakan proses lanjutan setelah mendapatkan data. Penggunaan hasil normalisasi ditujukan untuk kemudahan perhitungan. Hasil perubahan dari normalisasi data dapat dilihat pada Tabel 2.

**Tabel 2.** Normalisasi Data

No	Kabupaten/ Kota	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1	Kuantan Singingi	0.022	0.020	0.020	0.019	0.018	0.017	0.016	0.015	0.014	0.013	0.021	0.022
2	Indragiri Hulu	7	9	7	8	8	8	8	8	8	7	6	0
3	Indragiri Hilir	0.016	0.015	0.015	0.014	0.014	0.013	0.013	0.012	0.011	0.010	0.017	0.018
4	Pelalawan	7	4	5	8	2	6	0	3	6	9	7	1
...	...	0.020	0.019	0.018	0.017	0.016	0.015	0.014	0.013	0.012	0.011	0.013	0.013
10	Kepulauan Meranti	8	2	1	1	2	3	4	5	6	7	8	6
11	Pekanbaru	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00	00.00
12	Dumai	0.016	0.015	0.014	0.013	0.012	0.011	0.009	0.008	0.007	0.006	0.016	0.016
		8	2	3	2	1	0	9	9	9	8	4	2
		01.00	01.00	01.00	01.00	01.00	01.00	01.00	01.00	01.00	01.00	01.00	01.00
		0.094	0.093	0.093	0.092	0.091	0.090	0.089	0.088	0.087	0.086	0.107	0.108
		9	0	0	0	1	0	1	1	1	0	7	9

**2.2.1 Clustering**

Clustering merupakan teknik yang berguna dalam ilmu pengolahan data. Metode digunakan untuk menemukan struktur cluster didalam sebuah dataset yang memiliki karakter dengan kemiripan yang tinggi dalam satu cluster dan memiliki perbedaan yang besar dengan cluster yang berbeda [20]. Struktur dari clustering dapat direpresentasikan sebagai S dari sebuah subset S1, S2,...,Sk, seperti pada persamaan [21] (1).

$$S_1 \cap S_2 \cap S_3 \dots \cap S_k = \phi \tag{1}$$

**2.2.2 K-Means**

K-Means ialah algoritma yang bekerja dalam pengelompokan informasi ataupun objek yang telah didapatkan ke dalam suatu cluster. Untuk memproses klasterisasi, penetapan nilai k harus dilakukan terlebih dulu. Pada umumnya peneliti telah memiliki informasi dari objek yang dipelajari. Dalam melakukan perhitungan jarak data antar tiap titik pusat cluster dengan persamaan digunakan Teori jarak Euclidean (2).

$$D(L_j) = \sqrt{(X_{1i}-X_{1j})^2 + (X_{2i}-X_{2j})^2 + \dots + (X_{ki}-X_{kj})^2} \tag{2}$$

$D(i,j)$  adalah jarak antar data dengan pusat cluster  $j$ .

### 2.2.3 K-Medoids

K-Medoid ialah metode yang digunakan pembagian kluster objek. Algoritma K-Medoids memiliki tingkat efektifitas yang sangat baik untuk pengolahan dengan dataset yang sedikit. Langkah awal dalam K-Medoids yaitu penentuan titik representatif pada dataset dengan melakukan perhitungan jarak antar kelompok. Untuk melakukan proses perhitungan jarak vektor antar data digunakan Rumus Euclidian Distance (3).

$$deu(x_{ij}, ck_j) = \sqrt{\sum_{i=1}^n \sum_{j=1}^p (x_{ij} - ck_j)^2} \quad (3)$$

### 2.3 Davies-Bouldin Index (DBI)

Metode ini diperkenalkan oleh David L.Davies dan Donald W.Bouldin dan memiliki fungsi dalam gevaluasi cluster. Evaluasi dengan metode DBI memiliki skema evaluasi cluster internal, dengan indikator hasil cluster dinilai berdasarkan kuantitas dan jarak antar hasil cluster. Davies-Bouldin Index mengukur rata-rata kemiripan diantara setiap cluster [20], [22]. DBI berupaya untuk memaksimalkan ini antara jarak cluster sembari meminimalkan jarak antara pusat cluster dan objek data lainnya.

## 3. HASIL DAN ANALISIS

Berikut hasil dari penelitian ini dengan klustering menggunakan algoritma K-Means dan K-Medoids yang di uji dengan DBI (Davies Bouldin Index) sehingga akan didapatkan suatu hasil yang terbaik.

### 3.1 Proses Klasterisasi Data

Pada penelitian ini akan dilakukan proses perbandingan dengan dua algoritma. Algoritma yang akan diuji yaitu algoritma K-Means dan algoritma K-Medoids. Hal tersebut memiliki tujuan agar menghasilkan nilai cluster yang tepat.

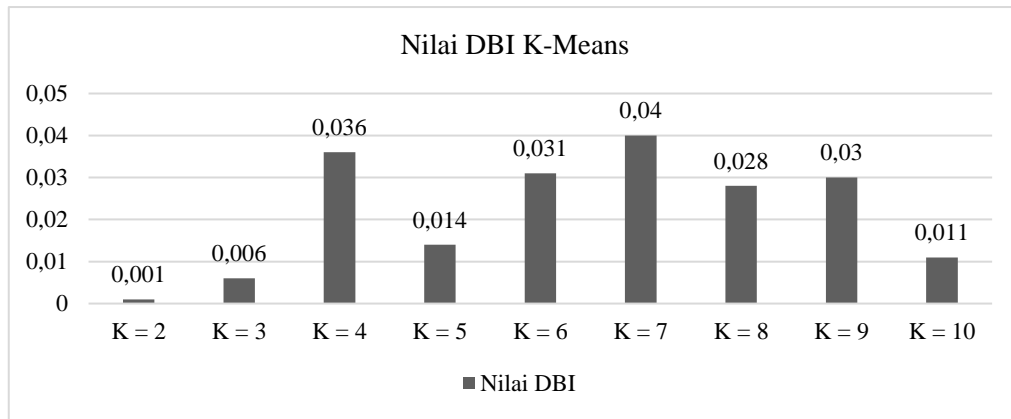
#### 3.1.1 K-Means

Pada penelitian ini yang akan dilakukan dengan menggunakan 2 sampai 10 cluster untuk mendapatkan hasil yang tepat. Pada hasil percobaan dari klustering dengan menggunakan algoritma *K-means* terdapat pada tabel 3.

**Tabel 3.** Hasil Klasterisasi K-Means

Percobaan	Kluster									
	1	2	3	4	5	6	7	8	9	10
K = 2	11	1	-	-	-	-	-	-	-	-
K = 3	10	1	1	-	-	-	-	-	-	-
K = 4	8	1	1	2	-	-	-	-	-	-
K = 5	1	1	5	4	1	-	-	-	-	-
K = 6	1	4	1	1	3	2	-	-	-	-
K = 7	2	1	1	1	3	2	2	-	-	-
K = 8	1	2	1	1	1	3	2	1	-	-
K = 9	2	1	1	1	2	1	1	2	1	-
K = 10	1	1	1	1	1	2	1	2	1	1

Pada tabel 3, didapatkan hasil klasterisasi K-Means dengan K=2 hingga k=10 menggunakan bantuan tools RapidMiner. Setelah di klasterisasi, selanjutnya dilakukan uji validas cluster dalam menemukan nilai yang terbaik dengan menggunakan metode DBI (*Davies Bouldin Index*). Nilai dari metode tersebut, didapatkan dari hasil clustering dengan *K-Means* yang dapat dilihat pada Gambar 2.



Gambar 2. Nilai DBI K-Means

Berdasarkan uji validitas DBI diatas, *cluster* terbaik dari algoritma *K-Means* ada di percobaan k = 2 dengan hasil nilai DBI sebesar 0,001 yang membagi keseluruhan data menjadi 2 cluster yaitu 11 anggota di cluster 1 dan 1 anggota di cluster 2.

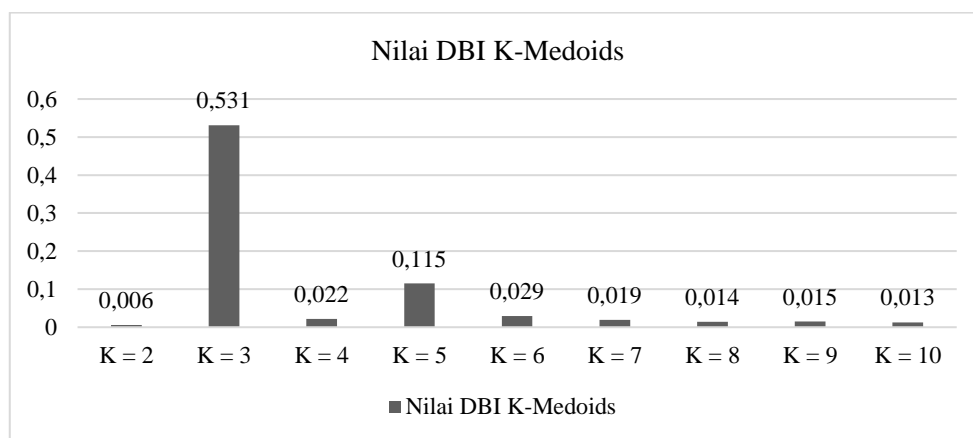
### 3.1.2 K-Medoids

Penelitian ini menggunakan algoritma *K-Medoids* dengan melakukan proses percobaan klasterisasi yang sama dengan yang sebelumnya. Hasil dari proses pengelompokan terdapat pada Tabel 4.

Tabel 4. Hasil Klasterisasi K-Medoids

Percobaan	Klaster									
	1	2	3	4	5	6	7	8	9	10
K = 2	11	1	-	-	-	-	-	-	-	-
K = 3	6	4	2	-	-	-	-	-	-	-
K = 4	6	4	1	1	-	-	-	-	-	-
K = 5	4	2	1	4	1	-	-	-	-	-
K = 6	4	1	1	4	1	1	-	-	-	-
K = 7	1	1	3	1	4	1	1	-	-	-
K = 8	1	1	1	1	2	4	1	1	-	-
K = 9	1	1	1	1	2	1	3	1	1	-
K = 10	1	1	3	1	1	1	1	1	1	1

Pada tabel 4, didapatkan hasil klasterisasi K-Medoids dengan K=2 hingga K=10 menggunakan tools RapidMiner. Setelah dilakukan tahapan klasterisasi, selanjutnya akan dilakukan nilai validitas *cluster* untuk menentukan *cluster* terbaik menggunakan metode DBI (*Davies Boulding Index*). Nilai DBI didapat dari proses clustering dengan *K-Medoids* dapat dilihat pada gambar 3.

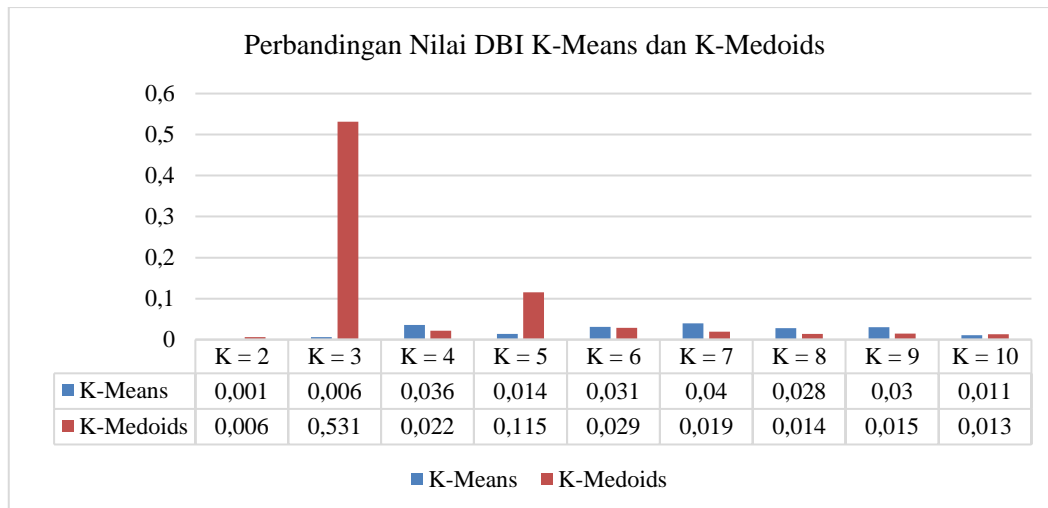


Gambar 3. Nilai DBI K-Medoids

Berdasarkan hasil uji validitas dengan DBI diatas, cluster yang terbaik dalam algoritma *K-Medoids* ada pada percobaan  $k = 2$  dengan nilainya sebesar 0,531 yang membagi data menjadi 2 *cluster* yaitu dengan jumlah 11 anggota pada cluster 1 dan 1 anggota pada cluster 2

### 3.2 Perbandingan DBI

Setelah melalui proses klusterisasi pada algoritma yang dipilih, data yang diolah dengan *rapidminer* untuk mendapatkan hasil dari nilai DBI (*Davies Bouldin Index*) pengelompokkan cluster  $k = 2$  hingga  $k = 10$  yang menjadi rujukan untuk mendapatkan penggunaan algoritma yang paling efektif. Pada gambar 4 dapat dilihat nilai dari pengolahan data.



**Gambar 4.** Perbandingan Nilai DBI antara K-Means dan K-Medoids

## 4. KESIMPULAN

Perbandingan menggunakan dua algoritma berdasarkan analisis dan pengolahan data, yakni K-Means dan K-Medoids. Maka dibuktikan bahwa algoritma K-Means dengan nilai  $k = 2$  adalah cluster terbaik. Davies Bouldin Index (DBI) telah memvalidasi hasil kalsterisasi. Pada penelitian ini, lebih unggul k-means daripada k-medoids dalam mengelompokkan data miskin berdasarkan provinsi dimana nilai DBI yang terbaik yakni 0.001 untuk nilai  $k = 2$ , percobaan k-means terbagi menjadi 2 cluster yaitu 11 anggota di cluster 1 dan 1 anggota di cluster 2 untuk nilai  $k = 2$

## REFERENCES

- [1] D. Gultom, I. Gunawan, I. Purnamasari, S. R. Andani, dan Z. A. Siregar, "Penerapan Algoritma K-Means Dalam Pengelompokan Kepadatan Penduduk Menurut Kecamatan di Kabupaten Simalungun," *TIN Terap. Inform. Nusant.*, vol. 2, no. 10, hal. 622–628, 2022, doi: 10.47065/tin.v2i10.1375.
- [2] G. J. Vietz, I. D. Rutherford, C. J. Walsh, Y. E. Chee, dan B. E. Hatt, "The unaccounted costs of conventional urban development: protecting stream systems in an age of urban sprawl," *7th Aust. Stream Manag. Conf. - Full Pap.*, hal. 1–7, 2014, [Daring]. Tersedia pada: <http://7asm-2014.p.asnevents.com.au/schedule/abstract/10793%5Cnpapers2://publication/uuid/59DBA754-180C-4174-8B64-493905B9D1DF>
- [3] C. J. Heckman, "Public Parks and Shady Areas in Times of Climate Change, Urban Sprawl, and Obesity," *Am. J. Public Health*, vol. 107, no. 12, hal. 1856–1858, 2017, doi: 10.2105/AJPH.2017.304140.
- [4] L. Liu, Z. Peng, H. Wu, H. Jiao, Y. Yu, dan J. Zhao, "Fast identification of urban sprawl based on K-means clustering with population density and local spatial entropy," *Sustain.*, vol. 10, no. 8, 2018, doi: 10.3390/su10082683.
- [5] D. Bakce, A. Syahza, dan B. Asmit, "Pembangunan Ekonomi Wilayah Perbatasan Antar-Negara di Provinsi Riau," *Unri Conf. Ser. Agric. Food Secur.*, vol. 1, no. 1, hal. 182–189, 2019, doi: 10.31258/unricsagr.1a24.
- [6] R. M. Sabiq dan N. Nurwati, "Pengaruh Kepadatan Penduduk Terhadap Tindakan Kriminal," *J. Kolaborasi Resolusi Konflik*, vol. 3, no. 2, hal. 161, 2021, doi: 10.24198/jkrk.v3i2.35149.
- [7] P. Marpaung dan R. F. Siahaan, "Penerapan Algoritma K-Means Clustering Untuk Pemetaan Kepadatan Penduduk Berdasarkan Jumlah Penduduk Kota Medan," *J. Sains Komput. Inform.*, vol. 5, no. 1, hal. 503–521, 2021.
- [8] I. Puspita, L. Ibrahim, dan D. Hartono, "Penurunan Kualitas Air Sungai Karang Anyar Kota Tarakan (

- Influence of The Behavior of Citizens Residing in Riverbanks to The Decrease of Water Quality in The River of Karang Anyar Tarakan City ) 1Program Magister Ilmu Lingkungan , Program Pascasarjana ,” *Mns. Dan Lingkung.*, vol. 23, no. 2, hal. 249–258, 2016.
- [9] R. Handayani, “Analisis Dampak Kependudukan terhadap Tingkat Kriminalitas di Provinsi Banten,” *J. Adm. Publik*, vol. 8, no. 2, hal. 149–169, 2017.
- [10] A. P. Kusuma dan D. M. Sukendra, “Analisis Spasial Kejadian Demam Berdarah Dengue Berdasarkan Kepadatan Penduduk,” *Unnes J. Public Heal.*, vol. 5, no. 1, hal. 48, 2016, doi: 10.15294/ujph.v5i1.9703.
- [11] M. J. Hamid Mughal, “Data mining: Web data mining techniques, tools and algorithms: An overview,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 6, hal. 208–215, 2018, doi: 10.14569/IJACSA.2018.090630.
- [12] C. Pahl dan D. Donnellan, “Data Mining Technology for the Evaluation of Web-based Teaching and Learning Systems,” *Conf. E-Learning Business, Gov. High. Educ.*, hal. 1–6, 2002.
- [13] N. T. Luchia, H. Handayani, F. S. Hamdi, D. Erlangga, dan S. Fitri Octavia, “Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 2, hal. 35–41, 2022.
- [14] G. Gustientiedina, M. H. Adiya, dan Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, hal. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [15] I. Kamila, U. Khairunnisa, dan M. Mustakim, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, hal. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [16] V. A. P. Sangga, “Perbandingan Algoritma K-Means dan Algoritma K-Medoids dalam Pengelompokan Komoditas Peternakan di Provinsi Jawa Tengah Tahun 2015,” 2018.
- [17] J. Arianto, “Penerapan Data Mining Untuk Pengelompokan Penduduk Kurang Mampu Desa Sambirejo Timur Dengan Algoritma K-Medoids (Studi Kasus Kantor Kepala Desa Sambirejo Timur),” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, hal. 569–573, 2019, doi: 10.30865/komik.v3i1.1660.
- [18] M. Herviany, S. P. Delima, T. Nurhidayah, dan Kasini, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor di Provinsi Jawa Barat,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, hal. 34–40, 2021, [Daring]. Tersedia pada: <https://journal.irpi.or.id/index.php/malcom/article/view/60>
- [19] D. T. Savra, Nabiilah, I. Ramadhani, A. Syahroni, dan P. H. Noprita, “Implementation of Multi Factor Evaluation Process (MFEP) for Choosing New Students in Baabussalam Foundation Implementasi Algoritma Multi Factor Evaluation Process (MFEP) untuk Penentuan Calon Peserta Didik Baru pada Yayasan Baabussalam,” *MALCOMIndonesian J. Mach. Learn. Comput. Sci.*, vol. 2, no. 2, hal. 53–60, 2022, [Daring]. Tersedia pada: <https://journal.irpi.or.id/index.php/malcom/article/view/428>
- [20] K. P. Sinaga dan M. S. Yang, “Unsupervised K-means clustering algorithm,” *IEEE Access*, vol. 8, hal. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [21] A. Saxena et al., “A review of clustering techniques and developments,” *Neurocomputing*, vol. 267, hal. 664–681, Des 2017, doi: 10.1016/J.NEUCOM.2017.06.053.
- [22] D. L. Davies dan D. W. Bouldin, “A Cluster Separation Measure,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, hal. 224–227, 1979, doi: 10.1109/TPAMI.1979.4766909.