



## *Comparison of K-Means and K-Medoids on Poor Data Clustering in Indonesia*

### **Perbandingan K-Means dan K-Medoids Pada Pengelompokan Data Miskin di Indonesia**

**Nanda Try Luchia<sup>1</sup>, Hani Handayani<sup>2</sup>, Fathan Surya Hamdi<sup>3</sup>,  
Dwi Erlangga<sup>4</sup>, Sania Fitri Octavia<sup>5</sup>**

<sup>1,2,3,4,5</sup>Program Studi Sistem Informasi, Fakultas Sains dan Teknologi  
<sup>1,2,3,5</sup>Puzzle Research Data Technology (Predatech), Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sultan Syarif Kasim Riau

E-mail: <sup>1</sup>nandaluchia@gmail.com, <sup>2</sup>hanihandayani2002@gmail.com, <sup>3</sup>fathan150601@gmail.com,  
<sup>4</sup>anggaundrrvol60@gmail.com, <sup>5</sup>saniahfitri2017@gmail.com

Received Aug 25th 2022; Revised Sept 05th 2022; Accepted Sept 15th 2022  
Corresponding Author: Nanda Try Luchia

#### **Abstract**

*In Indonesia, the problem of poverty has always been the center of attention and continues to increase. This can be an obstacle in both the economic, social, and political fields due to the lack of government attention in dealing with the issue. The government is making efforts to reduce poverty by providing direct cash assistance (BLT) and other policies and programs to people in need. In this study, the main data source was obtained from the Central Statistics Agency (BPS) website from 2014 to 2021. The poor population in Indonesia reached 26.50 million people or about 9.7% and is one of the largest poor people in the world. In this study, the grouping of data on the poor population in Indonesia in order to provide information to the government about the picture of poor areas with the highest number. The research was conducted by comparing two algorithms, namely K-Means and K-Medoids, which aimed to determine the best algorithm results. Based on the results of the study, it is found that the K-Means algorithm is better when compared to K-Medoids in clustering poor data in Indonesia. Proven with the best value of DBI K-means of 0.041 with experiment  $K = 8$ .*

*Keywords: Cluster, DBI, K-Means, K-Medoids, Poor*

#### **Abstrak**

Di Indonesia, masalah kemiskinan selalu menjadi pusat perhatian dan terus meningkat. Hal ini dapat menjadi hambatan baik di bidang ekonomi, sosial, maupun politik karena kurangnya perhatian pemerintah dalam menangani masalah tersebut. Pemerintah melakukan upaya untuk mengurangi angka kemiskinan dengan cara memberikan bantuan langsung tunai (BLT) dan kebijakan maupun program lainnya kepada masyarakat yang membutuhkan. Dalam penelitian ini, sumber data utama diperoleh dari website Badan Pusat Statistik (BPS) dari tahun 2014 hingga 2021. Penduduk miskin di Indonesia mencapai 26,50 juta jiwa atau sekitar 9,7% dan termasuk salah satu penduduk miskin terbesar di dunia. Pada penelitian ini dilakukan pengelompokan data penduduk miskin di Indonesia guna dapat memberikan informasi pada pemerintah mengenai gambaran wilayah miskin dengan angka tertinggi. Penelitian dilakukan dengan membandingkan dua algoritma yaitu K-Means dan K-Medoids yang bertujuan untuk mengetahui hasil algoritma terbaik. Berdasarkan hasil penelitian, didapatkan bahwa algoritma K-Means lebih baik jika dibandingkan dengan K-Medoids dalam mengklasterisasi data miskin yang ada di Indonesia. Dibuktikan dengan nilai terbaik DBI K-means sebesar 0.041 dengan percobaan  $K=8$ .

Kata Kunci: Cluster, DBI, K-Means, K-Medoids, Miskin

#### **1. PENDAHULUAN**

Kemiskinan menjadi salah satu permasalahan negara Indonesia yang terus meningkat dan belum bisa diselesaikan hingga saat ini [1]. Kemiskinan bisa dikatakan sebagai masalah multidimensi apabila terjadi hambatan akses ekonomi, masyarakat, budaya, politik dan partisipasi dalam masyarakat [2]. Permasalahan yang menyebabkan kemiskinan adalah kondisi suatu negara dan situasi global. Kemiskinan bukan hanya sekedar tidak cukupnya ekonomi, namun juga gagalannya kewajiban untuk memenuhi hak dan perlakuan

bermartabat dari setiap orang [3]. Hal yang dapat dilakukan adalah dengan menanggulangi kemiskinan tersebut.

Upaya pemerintah dalam menanggulangi kemiskinan dilakukan dengan memberikan bantuan tunai langsung (BLT) kepada masyarakat yang membutuhkan [4]. Program yang digagas pemerintah dalam menangani kemiskinan tersebut seringkali diberikan tidak tepat sasaran ke masyarakat yang membutuhkan. Badan Pusat Statistik (BPS) telah mencatat persentase dari penduduk miskin bulan September 2021 di Indonesia yaitu 26,50 juta jiwa atau sekitar 9,71%. Hal ini menunjukkan bahwa Indonesia termasuk dalam kategori salah satu negara dengan jumlah penduduk miskin yang tinggi di dunia. Tingkat kemiskinan penduduk Indonesia di setiap provinsi dapat dilihat pada gambar 1 di bawah ini.



**Gambar 1.** Persentase Tingkat Kemiskinan Penduduk di Indonesia  
(Sumber: Badan Pusat Statistik, 2021)

Pada gambar 1. menunjukkan nilai rata-rata tingkat kemiskinan di Indonesia sangat tinggi. Provinsi Papua menempati peringkat pertama tingkat kemiskinan tertinggi dengan jumlah persentase 21,70%. Begitu pula dengan Bali yang jumlah kemiskinannya berada di peringkat terakhir mencapai 4,45%. Hal ini disebabkan karena berbagai faktor seperti tingginya pertumbuhan penduduk, meningkatnya angka pengangguran, rendahnya pertumbuhan ekonomi dan taraf Pendidikan yang rendah. Mengingat kemiskinan yang dialami masyarakat di Indonesia, maka pemerintah perlu melakukan upaya pengurangan persentase penduduk miskin dengan berbagai bantuan[5]. Oleh karena itu, diperlukan rencana provinsi mana yang harus dijadikan prioritas. Pemerintah mengalami kesulitan untuk menentukan gambaran wilayah miskin dengan angka tertinggi. Masalah tersebut dapat diatasi dengan menggunakan pengelompokan data mining[6].

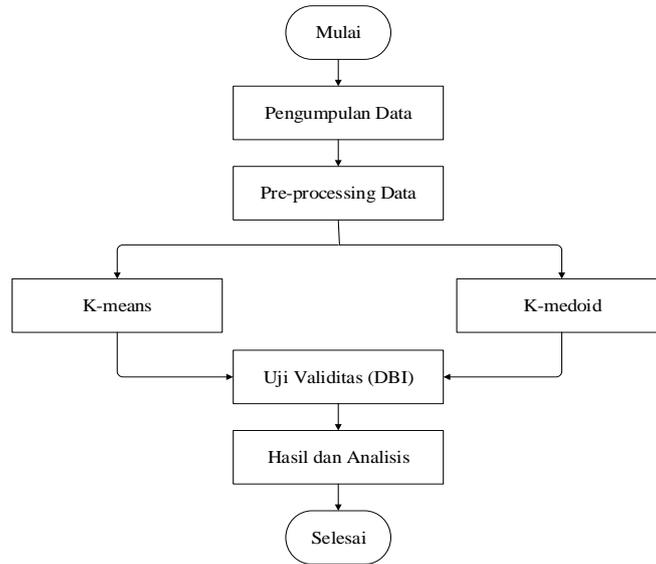
Data mining merupakan proses mencari dan menemukan fakta pada suatu database besar. Tujuannya yaitu menemukan pola eksklusif yang wajib dicapai agar bisa dipakai untuk mengetahui pengambilan keputusan yang akan datang [7]. Data Mining berdasarkan fungsi terbagi menjadi enam kelompok yaitu klasifikasi,regresi,association rule,anomaly detection,deployment dan clustering[8].Clustering merupakan proses pengelompokan beberapa titik dalam dua atau lebih kelompok hingga data dalam kelompok yang sama menjadi mirip satu sama lain berdasarkan informasi yang tersedia dengan poin data [9]. Ada beberapa algoritma yang terdapat dalam clustering seperti K-Means dan K-Medoids. Tujuannya agar partisi yang dihasilkan berkurang sensitivitasnya karena dataset memiliki nilai yang ekstrim, sehingga penggunaan medoid tidak didasarkan pada rata-rata yang diamati dari setiap *cluster* [10].

Dalam penelitian fitriyadi & Kurniawati [11] menyatakan bahwa tingkat akurasi K-Means lebih tinggi daripada K-Medoids. K-Means mudah dimengerti dan relatif efisien sehingga sering digunakan pada banyak studi kasus. Berdasarkan penelitian sebelumnya yang dilakukan oleh Kamila dkk [10] menyatakan bahwa dari segi waktu K-means hanya membutuhkan waktu 1 detik dalam mengolah data sedangkan K-Medoids membutuhkan waktu 1 menit 38 detik. Selain itu, nilai DBI pada K-Means lebih mendekati 0 dibanding K-Medoids.Sedangkan dalam penelitian Pujangga dkk [12] menyatakan bahwa waktu proses antara K-means dan K-medoids berbeda, dimana K-Medoids lebih cepat dibandingkan K-Means. Dalam penelitian Sangga[13], dilakukan perbandingan menggunakan nilai variansi cluster K-means sebesar 4.6 sedangkan K-medoids 1.5. hal ini menunjukan K-Medoids lebih baik dibandingkan dengan K-Means karena memiliki variansi nilai yang lebih kecil khususnya pada pengelompokan komoditas peternakan di Jawa Tengah.

Berdasarkan beberapa penelitian sebelumnya, peneliti mendapat kesimpulan yang berbeda dalam pola *cluster* dan tingkat akurasi. Pengujian performance algoritma menjadi cara untuk menemukan metode klastering terbaik. Oleh karena itu, dalam penelitian ini peneliti membandingkan klastering 2 algoritma berbeda yaitu K-Means dan K-Medoid untuk menentukan metode mana yang dapat digunakan dalam hal pengelompokkan. Penelitian ini bertujuan untuk mengidentifikasi potensi dari setiap algoritma dalam melakukan klastering hingga mendapatkan hasil yang optimal. Penelitian ini diharapkan dapat membantu pemerintah dalam memberikan informasi tentang pengelompokkan data penduduk miskin untuk penyebarluasan bantuan.

**2. METODOLOGI PENELITIAN**

Metodologi dalam penelitian ini, dibagi dalam lima tahap, yaitu : pengumpulan data, pre-processing data, proses klasterisasi data, validitas klaster dan hasil analisis. Metodologi penelitian dapat dilihat pada gambar 2 dibawah ini.



**Gambar 2.** Metodologi Penelitian

**2.1 Pengumpulan Data**

Pengumpulan sumber data utama dilakukan dengan melihat data penduduk miskin di website BPS.co.id pada tahun 2014 hingga 2021. Seluruh data diproses agar mendapatkan daftar penduduk miskin di Indonesia dari setiap provinsi. Data ini dapat membantu memberikan informasi kepada pemerintah untuk penyebaran bantuan sosial kepada masyarakat. Untuk proses klasterisasi data, dibutuhkan 9 atribut dari data bps yang telah didapat sebelumnya. Atributnya yaitu provinsi dan tahun 2014-2021. Atribut-atribut ini menjadi acuan untuk pengelompokan. Data yang telah didapat bisa dilihat ditabel 1.

**Tabel 1.** Data Penduduk Miskin di Indonesia

No	Provinsi	2014	2015	2016	2017	2018	2019	2020	2021
1	Aceh	35.03	34.19	33.16	32.81	31.65	30.33	30.42	30.86
2	Sumatera Utara	19.23	21.32	20.62	19.5	18.16	17.46	17.89	17.5
3	Sumatera Barat	14.3	14.02	14.23	13.62	13.2	12.71	12.84	12.67
4	Riau	16.11	17.24	15.65	15.19	14.6	13.98	13.86	14.12
5	Jambi	16.31	17.98	16.78	16.09	15.77	15.11	15.55	15.76
6	Sumatera Selatan	27.53	28.02	26.93	26.29	25.62	25.27	25.64	25.63
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
32	Maluku Utara	14.71	13.06	12.74	12.79	13.26	13.68	13.75	13.27
33	Papua Barat	53.39	51.55	50.31	48.22	45.67	43.68	43.07	43.66
34	Papua	57.85	56.57	56.94	55.38	55.17	54.08	53.44	54.24

**2.2 Pre-processing Data**

Setelah data didapatkan, langkah selanjutnya melakukan normalisasi data. Hasil data normalisasi didapatkan untuk memudahkan melakukan perhitungan. Perubahan data hasil normalisasi dapat dilihat pada tabel 2.

**Tabel 2.** Normalisasi Data

No	Provinsi	2014	2015	2016	2017	2018	2019	2020	2021
1	Aceh	0.5421	0.5435	0.519	0.528	0.5105	0.4967	0.491	0.48
2	Sumatera Utara	0.2251	0.2811	0.265	0.25	0.2298	0.224	0.214	0.183
3	Sumatera Barat	0.1262	0.1322	0.136	0.127	0.1265	0.1233	0.102	0.076
4	Riau	0.1625	0.1978	0.165	0.16	0.1557	0.1502	0.125	0.108
5	Jambi	0.1665	0.2129	0.188	0.179	0.18	0.1742	0.162	0.145
6	Sumatera Selatan	0.3917	0.4177	0.393	0.392	0.385	0.3895	0.385	0.364
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

No	Provinsi	2014	2015	2016	2017	2018	2019	2020	2021
32	Maluku Utara	0.1344	0.1126	0.106	0.11	0.1278	0.1439	0.122	0.089
33	Papua Barat	0.9105	0.8976	0.866	0.85	0.8023	0.7796	0.771	0.765
34	Papua	10.000	10.000	10.000	10.000	10.000	10.000	10.000	10.000

### 2.3 Clustering

Dalam data mining, klastering keilmuan merupakan proses sejumlah data yang dikelompokkan menjadi *cluster* (group) hingga tiap *cluster* berisi data yang mirip maupun berbeda dengan *cluster* lainnya [14]. Klastering sendiri sering dijadikan metodologi analisis secara otomatis untuk mengklasifikasikan data menjadi kelompok dengan ukuran asosiasi [15]. Klastering juga menjadi tahap pengolahan dataset melalui perhitungan yang dimodelkan dengan algoritma-algoritma untuk menghasilkan pengelompokan yang sesuai [16].

#### 2.3.1 K-Means

Metode algoritma k-means merupakan cara objek ataupun informasi yang didapat dikelompokkan ke dalam *cluster*. Untuk melaksanakan klasterisasi, nilai k wajib ditetapkan terlebih dulu. Umumnya user (pengguna) telah memiliki data dini tentang objek yang dipelajari [17]. Teori jarak *Euclidean* digunakan untuk bisa menghitung jarak data ke tiap titik pusat *cluster* dengan persamaan (1).

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

$D(i,j)$  merupakan jarak anatar data dan pusat *cluster*  $j$ . [18].

#### 2.3.2 K-Medoids

K-Medoid adalah proses partisi klastering untuk mengelompokkan kumpulan objek ke dalam *cluster* [15]. Algoritma K-Medoids sangat efektif untuk dataset kecil. Langkah pertama dalam K-Medoids yaitu menemukan titik representative pada dataset dengan menghitung jarak kelompok. Rumus *Euclidian Distance* untuk menghitung jarak vektor antar dokumen adalah pada persamaan 2 [19].

$$deu(xij, ckj) = \sqrt{\sum_{i=1}^n \sum_{j=1}^p (xij - ckj)^2} \quad (2)$$

### 2.4 Uji Validitas

Uji validitas yang digunakan pada penelitian ini adalah DBI yang berguna untuk mengevaluasi hasil dari pengelompokan untuk memeriksa hasil dari sebuah cluster itu sendiri [20]. Evaluasi dengan DBI memiliki evaluasi *cluster* internal skema, hal ini berarti hasil *cluster* tersebut baik atau tidaknya dinilai dari kedekatan dan kuantitas antara hasil *cluster* itu sendiri [21].

## 3. HASIL DAN ANALISIS

Hasil penelitian yang dilakukan berupa hasil klastering melalui proses dengan menggunakan algoritma K-Means dan K-Medoids yang kemudian akan di uji menggunakan DBI sehingga didapatkan hasil algoritma terbaik.

### 3.1 Proses Klasterisasi data

Dalam penelitian ini, dilakukan perbandingan dua algoritma menggunakan K-Means dan K-Medoids dengan data hasil normalisasi sebelumnya. Beberapa percobaan klastering menggunakan algoritma K-Means dan K-Medoids. Hal ini bertujuan untuk mendapatkan nilai *cluster* yang tepat.

#### 3.1.1 K-Means

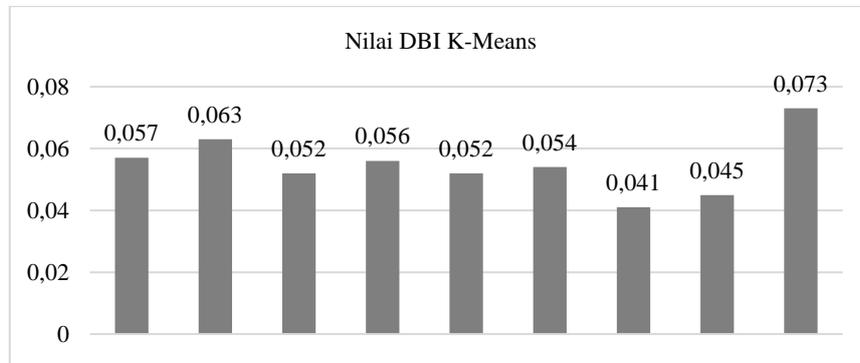
Proses percobaan menggunakan 2 *cluster* hingga 10 *cluster* untuk mendapatkan *cluster* yang tepat. Hasil percobaan klastering algoritma K-Means dapat pada tabel 3.

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

Percobaan	Klaster									
	1	2	3	4	5	6	7	8	9	10
K=2	18	2	12	-	-	-	-	-	-	-
K=3	16	2	6	6	-	-	-	-	-	-
K=4	10	1	16	10	2	-	-	-	-	-
K=5	5	9	10	8	8	2	-	-	-	-
K=6	8	1	7	4	2	4	1	-	-	-
K=7	10	3	1	6	1	6	8	-	-	-
K=8	4	6	1	1	7	4	-	-	-	-
K=9	1	6	8	1	4	8	1	2	1	2

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

Setelah klusterisasi, selanjutnya dilakukan validitas *cluster* untuk mendapatkan klaster terbaik dengan teknik Davies Bouldin Index (DBI). Nilai DBI hasil proses clustering data dengan K-Means dapat dilihat pada gambar 3.



Gambar 3. Nilai DBI K-Means

Berdasarkan nilai DBI, *cluster* terbaik K-Means ada pada percobaan k=8 dengan nilai DBI yaitu 0,041 yang membagi data menjadi 8 *cluster* dengan 4 anggota dari *cluster* 1, 6 anggota dari *cluster* 2, 1 anggota dari *cluster* 3 dan 4, 7 anggota dari *cluster* 5, dan 4 anggota dari *cluster* 6.

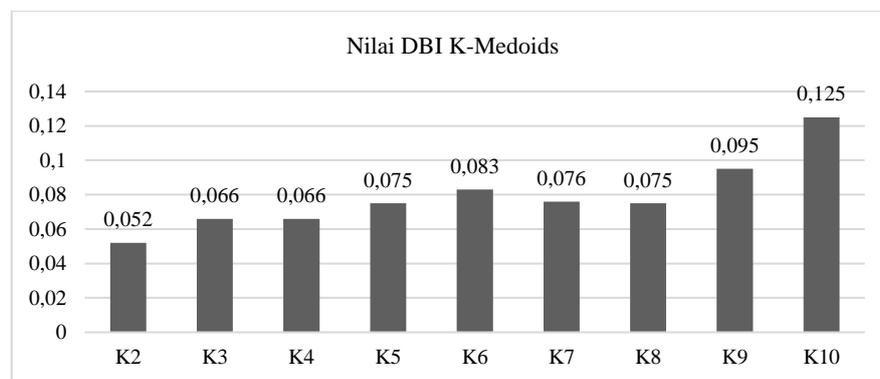
### 3.1.2 K-Medoids

Pada penelitian ini juga menggunakan algoritma K-Medoid dengan percobaan *cluster* yang sama dengan algoritma sebelumnya. Hasil *cluster* dapat dilihat pada tabel 4.

Tabel 4. Hasil Klusterisasi K-Medoids

Percobaan	Klaster									
	1	2	3	4	5	6	7	8	9	10
K=2	4	30	-	-	-	-	-	-	-	-
K=3	2	10	22	-	-	-	-	-	-	-
K=4	2	16	10	6	-	-	-	-	-	-
K=5	2	6	10	10	6	-	-	-	-	-
K=6	2	7	5	6	8	6	-	-	-	-
K=7	2	4	5	5	6	5	7	-	-	-
K=8	2	2	5	5	4	5	5	6	-	-
K=9	2	2	3	4	5	6	3	2	7	-
K=10	5	4	2	5	3	3	3	5	2	2

Setelah proses *cluster* data dilakukan, selanjutnya dilakukan hal yang sama yaitu validitas *cluster* untuk menemukan klaster terbaik dengan teknik Davies Bouldin Index (DBI). Nilai DBI hasil proses clustering data dengan algoritma K-Medoids dapat dilihat pada gambar 4.

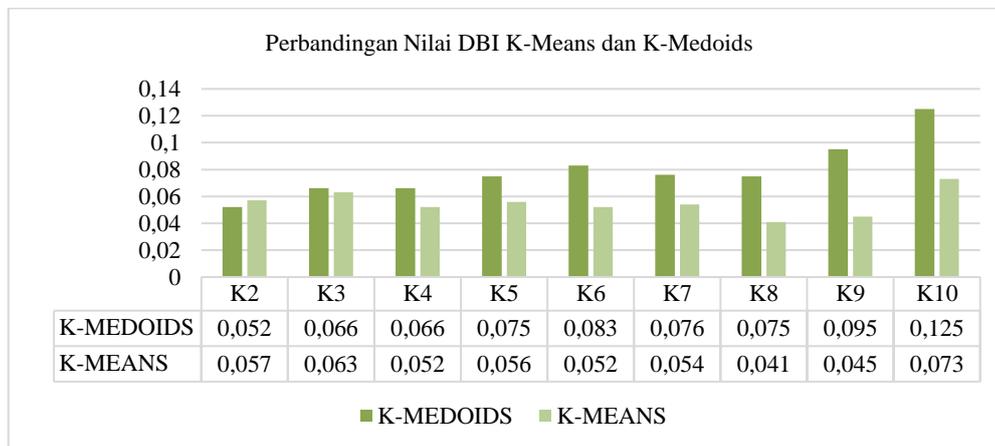


Gambar 4. Nilai DBI K-Medoids

Berdasarkan hasil dari nilai DBI, *cluster* terbaik pada K-Medoids ada pada percobaan  $k=2$  dengan nilai DBI yaitu 0,52 yang membagi data menjadi 2 *cluster* dengan jumlah anggota pada *cluster* 1 yaitu 4 anggota dan *cluster* 2 sebanyak 30 anggota.

### 3.2 Perbandingan DBI

Setelah melakukan proses klusterisasi, data diolah dengan rapidminer untuk memperlihatkan nilai Davies Bouldin Index (DBI) yang menjadi rujukan untuk pengelompokan *cluster*  $k=2$  hingga  $k=10$ . Nilai dari pengolahan data dapat dilihat pada gambar 5.



**Gambar 5.** Perbandingan Nilai DBI antara K-Means dan K-Medoids

Pada gambar 5, memperlihatkan perbandingan nilai Davies Bouldin Index (DBI) dari dua algoritma, menunjukkan hasil  $k=8$  merupakan *cluster* terbaik pada algoritma K-Means karena memiliki nilai yang paling mendekati 0. Hal ini membuktikan bahwa K-Means lebih unggul dibanding K-Medoids dalam pengelompokan data penduduk miskin di Indonesia.

## 4. KESIMPULAN

Berdasarkan hasil analisis serta pengolahan data menggunakan perbandingan dua algoritma, yaitu K-Means dan K-Medoids, didapatkan *cluster* terbaik dengan nilai  $k=8$  pada algoritma K-Means. Sedangkan *cluster* terbaik pada K-Medoids dengan nilai  $k=2$ . Hasil klusterisasi yang dilakukan sudah divalidasi dengan Davies Bouldin Index (DBI). Maka pada penelitian ini, algoritma K-Means lebih unggul dibanding K-Medoids pada pengelompokan data miskin berdasarkan provinsi dengan nilai DBI terbaik yaitu 0.041 dengan nilai  $k=8$ . Percobaan K-Means dengan nilai  $k=8$  membagi data menjadi 8 *cluster* dengan 4 anggota dari *cluster* 1, 6 anggota dari *cluster* 2, 1 anggota dari *cluster* 3 dan 4, 7 anggota dari *cluster* 5, dan 4 anggota dari *cluster* 6.

## REFERENSI

- [1] D. N. Alfiansyah, V. R. S. Nastiti, and N. Hayatin, "Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar," *J. Repos.*, vol. 4, no. 1, pp. 49–58, 2022, doi: 10.22219/repositor.v4i1.1416.
- [2] I. Nasution, ... A. W.-B. of I., and undefined 2020, "Penerapan Algoritma K-Means Dalam Pengelompokan Data Penduduk Miskin Menurut Provinsi," *ejurnal.seminar-id.com*, vol. 2, no. 2, pp. 76–83, 2020, Accessed: Jun. 08, 2022. [Online]. Available: <http://ejurnal.seminar-id.com/index.php/bits/article/view/492>
- [3] G. Dwilestari, Mulyawan, Martanto, and I. Ali, "Analisis Clustering menggunakan K-Medoid pada Data Penduduk Miskin Indonesia," *JURSIMA J. Sist. Inf. dan Manaj.*, vol. 9, no. 3, pp. 282–290, 2021.
- [4] Febriyana, *Analisis Klaster K-Means dan K-Median Pada Data Indikator Kemiskinan*. 2011.
- [5] D. V. Ferezagia, "Analisis Tingkat Kemiskinan di Indonesia," *J. Sos. Hum. Terap.*, vol. 1, no. 1, pp. 1–6, 2018, doi: 10.7454/jsht.v1i1.6.
- [6] N. Normah, S. Nurajizah, and A. Salbinda, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *J. Tek. Komput. AMIK BSI*, vol. 7, no. 2, pp. 158–163, 2021.
- [7] "View of Penerapan K-Means dan Fuzzy C-Means untuk Pengelompokan Data Kasus Covid-19 di Kabupaten Indragiri Hilir." <http://ejurnal.seminar-id.com/index.php/bits/article/view/1005/698> (accessed Jun. 10, 2022).
- [8] "Data Mining dan Penerapan Algoritma - Ni Luh Wiwik Sri Rahayu Ginantra, Fatimah Nur Arifah, Anggi Hadi Wijaya, Ri Sabti Septarini, Nazaruddin Ahmad, Dewa Putu Yudhi Ardiana, Faried Effendy,

- Akbar Iskandar, Hazriani Hazriani, Ika Yusnita Sari, Zelvi Gustiana, Cahyo Prianto, Dudih Gustian, Edi Surya Negara - Google Buku.” [https://books.google.co.id/books?hl=id&lr=&id=v0gtEAAAQBAJ&oi=fnd&pg=PA5&dq=data+mining+terdiri+dari&ots=RiY4qCddgF&sig=ZzT8gop8Jcn04usrpi060G-Hn0g&redir\\_esc=y#v=onepage&q=data+mining+terdiri+dari&f=false](https://books.google.co.id/books?hl=id&lr=&id=v0gtEAAAQBAJ&oi=fnd&pg=PA5&dq=data+mining+terdiri+dari&ots=RiY4qCddgF&sig=ZzT8gop8Jcn04usrpi060G-Hn0g&redir_esc=y#v=onepage&q=data+mining+terdiri+dari&f=false) (accessed Jun. 28, 2022).
- [9] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [10] I. Kamila, U. Khairunnisa, and 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, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [11] A. U. Fitriyadi, “Algoritma K-Means dan K-Medoids Analisis Algoritma K-Means dan K-Medoids Untuk Clustering Data Kinerja Karyawan Pada Perusahaan Perumahan Nasional,” *Kilat*, vol. 10, no. 1, pp. 157–168, 2021, doi: 10.33322/kilat.v10i1.1174.
- [12] R. Pujangga and A. S. Kurniawansyah, “Analisis Perbandingan Algoritma K-Means dan Algoritma K-Medoid Dalam Klasterisasi Penduduk Lanjut Usia Comparative Analysis of the K-Means Algorithm and the K-Medoid Algorithm in Clustering the Elderly Population,” vol. 2, no. 2, pp. 123–128, 2021.
- [13] V. A. P. Sangga, “Perbandingan algoritma K-Means dan algoritma K-Medoids dalam pengelompokan komoditas peternakan di provinsi Jawa Tengah tahun 2015,” 2018.
- [14] 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.
- [15] I. Parlina, A. P. Windarto, A. Wanto, and M. R. Lubis, “Memanfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Aseessment Center,” *Memanfaatkan Algoritm. K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Aseessment Cent. Untuk Clust. Progr. Sdp*, vol. 3, no. 1, pp. 87–93, 2018.
- [16] Yuni Radana Sembiring, Saifullah, and Riki Winanjaya, “Implementasi Data Mining Dalam Mengelompokkan Jumlah Penduduk Miskin Berdasarkan Provinsi Menggunakan Algoritma,” *KESATRIA J. Penerapan Sist. Inf. (Komputer Manajemen) Vol. 2, No. 2*, vol. 2, no. 2, pp. 125–132, 2021.
- [17] “Santosa B.(2007). Data Mining (Teori dan Aplikasi)... - Google Cendekia.” [https://scholar.google.com/scholar?hl=id&as\\_sdt=0%2C5&q=Santosa+B.%282007%29.+Data+Minig+%28Teori+dan+Aplikasi%29.+Graha+Ilmu%2C+Yogyakarta.&btnG=](https://scholar.google.com/scholar?hl=id&as_sdt=0%2C5&q=Santosa+B.%282007%29.+Data+Minig+%28Teori+dan+Aplikasi%29.+Graha+Ilmu%2C+Yogyakarta.&btnG=) (accessed May 27, 2022).
- [18] R. Sari, D. H.-S. N. S. dan Teknologi, and undefined 2018, “Data Mining: Algoritma K-Means Pada Pengelompokkan Wisata Asing ke Indonesia Menurut Provinsi,” *seminar-id.com*, 2018, Accessed: Jun. 12, 2022. [Online]. Available: <http://seminar-id.com/prosiding/index.php/sensasi/article/view/47>
- [19] H. Zayuka, S. Michrandi Nasution, and Y. Purwanto, “Perancangan Dan Analisis Clustering Data Menggunakan Metode K-Medoids Untuk Berita Berbahasa Inggris Design And Analysis Of Data Clustering Using K-Medoids Method For English News”.
- [20] S. Butsianto and N. Saepudin, “Penerapan Data Mining Terhadap Minat Siswa Dalam Mata Pelajaran Matematika Dengan Metode K-Means,” *J. Nas. Komputasi dan Teknol. Inf.*, vol. 3, no. 1, pp. 51–59, 2020, doi: 10.32672/jnkti.v3i1.2008.
- [21] S. D. Agustina, M. Mustakim, C. Bella, and M. A. Ramadhan, “Support vector regression algorithm modeling to predict the availability of foodstuff in Indonesia to face the demographic bonus,” in *Journal of Physics: Conference Series*, 2018, vol. 1028, no. 1, p. 12240.