



Perbandingan Metode *K-Means* dan *K-Medoids* Dengan Validitas Davies-Bouldin Indeks, Dunn Indeks dan Indeks Connectivity Pada Pengelompokkan Masyarakat Penerima Bantuan Langsung Tunai

Nur Ain Kilo^{1*}, Muh Rifai Katili², Isran K Hasan³

^{1,2,3}Jurusan Matematika, Universitas Negeri Gorontalo, Bone Bolango 96554, Indonesia

Info Artikel

*Penulis Korespondensi.
Email: nurainkilo7@gmail.com

Diterima: 2 September 2024
Direvisi: 11 Desember 2024
Disetujui: 14 Januari 2025



Under the licence
CC BY-NC-SA 4.0

Diterbitkan oleh:



Copyright ©2025 by Author(s)

Abstrak

Penelitian ini membahas perbandingan antara metode K-Means dan K-Medoids dalam pengelompokan masyarakat penerima bantuan langsung tunai (BLT), dengan penilaian menggunakan tiga indeks validitas: Davies-Bouldin Indeks (DBI), Dunn Indeks, dan Indeks Connectivity. Tujuan utama dari penelitian ini adalah untuk menentukan metode clustering yang paling efektif untuk mengelompokkan data penerima BLT dengan mempertimbangkan kualitas clustering yang dihasilkan. Dalam eksperimen, metode K-Means dengan tiga cluster menghasilkan, yaitu: Cluster 1 dengan 10 anggota kepala keluarga, Cluster 2 dengan 101 anggota kepala keluarga, dan Cluster 3 dengan 118 anggota kepala keluarga. Sebaliknya, metode K-Medoids juga dengan tiga cluster, yakni: Cluster 1 dengan 67 anggota kepala keluarga, Cluster 2 dengan 59 anggota kepala keluarga, dan Cluster 3 dengan 103 anggota kepala keluarga. Berdasarkan evaluasi menggunakan Davies-Bouldin Indeks dan Indeks Connectivity, metode K-Means menunjukkan performa yang lebih baik dibandingkan K-Medoids. Nilai DBI untuk metode K-Means adalah 1.307, sedangkan nilai Indeks Connectivity adalah 40.079, yang menunjukkan bahwa hasil clustering K-Means lebih efektif dalam menghasilkan cluster yang terpisah dan berkualitas dalam konteks pengelompokan masyarakat penerima BLT.

Kata Kunci: *K-Means; K-Medoids; Davies-Bouldin Indeks; Dunn Indeks; Connectivity Indeks; BLT*

Abstract

This study discusses the comparison between the K-Means and K-Medoids methods in grouping direct cash assistance (BLT) recipients, with an assessment using three validity indices: Davies-Bouldin Index (DBI), Dunn Index, and Connectivity Index. The main objective of this study is to determine the most effective clustering method for grouping BLT recipient data by considering the quality of the resulting clustering. In the experiment, the K-Means method with three clusters produced, namely: Cluster 1 with 10 family head members, Cluster 2 with 101 family head members, and Cluster 3 with 118 family head members. In contrast, the K-Medoids method also with three clusters, namely: Cluster 1 with 67 family head members, Cluster 2 with 59 family head members, and Cluster 3 with 103 family head members. Based on the evaluation using the Davies-Bouldin Index and Connectivity Index, the K-Means method showed better performance than K-Medoids. The DBI value for the K-Means method is 1,307, while the Connectivity Index value is 40,079, which shows that the K-Means clustering results are more effective in producing separate and quality clusters in the context of grouping BLT recipient communities.

Keywords: *K-Means; K-Medoids; Davies-Bouldin Indeks; Dunn Indeks; Connectivity Indeks; BLT*

1. Pendahuluan

Kemiskinan adalah masalah yang meluas di hampir semua negara berkembang. Pemerintah Indonesia secara aktif menangani kemiskinan, yang merupakan masalah utama di daerah perkotaan

dan pedesaan. Masalah kemiskinan di Indonesia merupakan tantangan berat bagi pemerintah untuk mengatasinya [1]. Badan Perencanaan Pembangunan Nasional (Bappenas) telah melaporkan bahwa masalah kemiskinan di Indonesia memburuk selama periode 2020-2022 karena adanya hambatan serius terhadap pembangunan yang disebabkan oleh pandemi *Covid-19*. Perekonomian telah mengalami kontraksi, yang menyebabkan peningkatan pengangguran dan kemudian meningkatkan tingkat kemiskinan. Dampak ini diperkirakan akan terus berlanjut, dengan mengakibatkan sebagian penduduk menjadi masyarakat miskin ekstrim [2]. Pemerintah telah menerapkan berbagai strategi untuk mengentaskan kemiskinan, termasuk program Bantuan Langsung Tunai (BLT).

Bantuan Langsung Tunai (BLT) merupakan sebuah inisiatif pemerintah yang bertujuan untuk meningkatkan jaring pengaman sosial (JPS) melalui pemanfaatan dana desa (BLT). Inisiatif BLT bertujuan untuk mengatasi berbagai tantangan yang dihadapi oleh semua warga negara, terutama yang berkaitan dengan ekonomi. Kebijakan pemberian bantuan tunai melalui Bantuan Langsung Tunai (BLT) dituangkan dalam Instruksi Presiden (Inpres) Nomor 4 Tahun 2022, yang berfokus pada percepatan penanggulangan kemiskinan ekstrim. Alokasi Bantuan Langsung Tunai (BLT) Dana Desa merupakan langkah yang bertujuan untuk meningkatkan pendapatan rumah tangga miskin yang tinggal di desa. Menurut mentri Keuangan Republik Indonesia 2022 tentang tata kelola dana desa, Pasal 35 ayat (a) menetapkan anggaran yang dialokasikan untuk kegiatan prioritas yang akan didanai oleh Dana Desa. Alokasi keuangan untuk program pemulihan ekonomi, khususnya perlindungan sosial dan penanggulangan kemiskinan yang parah melalui inisiatif BLT Desa, harus berkisar antara minimal 10 persen hingga maksimal 25 persen dari anggaran Dana Desa [3].

Program BLT harus dilaksanakan dengan benar, transparan dan teratur agar BLT yang diberikan dapat diterima oleh masyarakat miskin yang sangat membutuhkan. Proses penerimaan Bantuan Langsung Tunai dilakukan dengan memilih daftar calon penerima dengan kriteria yang telah ditentukan. Proses penilaian diberikan kepada kepala desa atau kepala RT/Dusun setempat yang bersifat subyektif dan akan dipilih lagi melalui Musyawarah desa. Hal ini dapat menimbulkan kekhawatiran akan ketidaktepatan sasaran penerima BLT, sehingga BLT tidak dapat menjangkau masyarakat miskin yang sangat membutuhkan. Dalam pelaksanaannya BLT kurang efektif karena tidak tepatnya pelaksana dalam memilih penerima bantuan secara subyektif. Hal tersebut dapat menjadi penghambat utama dalam efektivitas dana desa. Oleh karena itu diperlukan keakuratan data untuk dapat mengelompokkan penerima BLT sehingga penentuan prioritas pada kelayakan penerima dapat teratasi secara objektif [4].

Clustering adalah istilah yang digunakan dalam Matematika untuk merujuk pada proses pengelompokan data menurut kualitasnya. *Clustering* melibatkan pengelompokan elemen data yang memiliki kesamaan fitur ke dalam satu kelompok, sementara yang memiliki sifat yang berbeda ditempatkan di kelompok yang terpisah. Pemisahan dilakukan secara otomatis menggunakan teknik *clustering* daripada dilakukan secara manual. Beberapa penelitian dengan klasterisasi dengan metode statistika digunakan untuk menyelesaikan kasus yang berbeda, seperti Bula, dkk [5] yang melakukan klusterisasi pada kasus literasi informasi generasi muda dengan menggunakan metode Improve Chi-Square. Pada penelitian lainnya, Warolemba, dkk [6] menggunakan Fuzzy C-Means untuk klasterisasi kesejahteraan rakyat. Sementara Hasan, dkk [7] menggunakan K-NN dan Random Forest untuk klasifikasi lama studi mahasiswa. Beberapa penelitian ini menunjukkan bahwa setiap metode klusterisasi memiliki keunggulan masing-masing dalam terapannya untuk analisis klasterisasi di berbagai kasus yang berbeda.

Analisis *clustering* mencakup dua pendekatan yang berbeda: metodologi hirarki dan non-hirarki. Pendekatan hirarki adalah teknik yang secara khusus dikembangkan untuk tujuan memecah kumpulan data ke dalam tingkatan hirarki berdasarkan kualitas data. Metode non-hierarki adalah teknik yang digunakan untuk pengelompokan data, di mana jumlah cluster dapat diatur. Dalam ranah analisis clustering, ada dua jenis algoritma yang masih saling berhubungan erat, yaitu *K-Means* dan *K-Medoids*. Kedua algoritma tersebut merupakan teknik *clustering* non-hierarki, yang menawarkan waktu pemrosesan yang lebih cepat dibandingkan dengan pendekatan hierarki dan sangat bermanfaat ketika berhadapan dengan data yang sangat besar [8].

Algoritma *clustering K-Means* adalah teknik yang digunakan untuk mengelompokkan sekumpulan objek berdasarkan kesamaan sifat atau karakteristik ke dalam beberapa kelompok (diwakili oleh bilangan bulat positif). Sebuah klaster didefinisikan oleh sebuah massa yang berfungsi sebagai nilai rata-rata klaster. *K-means* adalah teknik analisis klaster yang banyak digunakan dalam bidang data mining. Ini adalah teknik untuk kuantisasi vektor. *K-Means* adalah teknik pengelompokan berbasis jarak yang mempartisi data ke dalam banyak kelompok. Algoritma ini secara khusus beroperasi pada properti numerik. Hanya beroperasi pada atribut yang bersifat numerik. Sedangkan *K-Medoids* adalah varian dari algoritma *K-Means* yang memanfaatkan pengukuran jarak dan lebih intensif. Meskipun masih terhubung dengan *K-Means*, *K-Medoids* adalah versi umum dari algoritma tersebut. Kedua metode ini bersifat partitional, yang berarti mereka membagi dataset menjadi beberapa kelompok, dan bertujuan untuk meminimalkan kesalahan kuadrat. Kesalahan kuadrat mengacu pada jarak antara titik-titik berlabel dalam sebuah klaster dan titik pusat yang dipilih dari klaster tersebut. Perbedaan utama antara *K-Medoids* dan *K-Means* adalah bahwa *K-Medoids* memilih titik data tertentu untuk dijadikan pusatnya. Salah satu aspek penting yang perlu diperhatikan dalam analisis cluster adalah validasi hasil clustering. Hasil klaster divalidasi untuk menentukan partisi yang optimal untuk data. Sesuaikan data untuk merepresentasikan informasi secara akurat. Kegagalan dalam melakukan validasi akan berdampak pada hasil analisis clustering. Pemeriksaan atau studi tentang sesuatu secara detail untuk memahami sifat atau strukturnya [9].

Berdasarkan uraian di atas, peneliti tertarik untuk melakukan penelitian perbandingan metode *clustering K-Means* dan *K-Medoids* untuk mengelompokkan masyarakat penerima Bantuan Langsung Tunai dengan menggunakan beberapa Indeks Validitas yaitu *Davies-Boulding*, *Indeks Dunn*, dan *Connectivity Indeks* untuk melihat hasil metode mana yang lebih optimal.

2. Metode Penelitian

2.1 Sumber Data

Data yang digunakan adalah data primer dan data sekunder. Data tersebut diambil dari Pemerintah Desa Lupoyo Kecamatan Telaga Biru. Populasi yang digunakan pada penelitian ini adalah Keluarga Penerima Bantuan Langsung Tunai, Seluruh anggota populasi adalah sampel pada penelitian ini. Jumlah Keluarga Penerima BLT di Desa Lupoyo berjumlah 229 Keluarga penerima manfaat. Adapun teknik penarikan sampel pada penelitian ini menggunakan sampling jenuh. Dimana semua populasi yang ada dijadikan sampel.

2.2 Tahapan Penelitian

Penelitian ini dilakukan berdasarkan prosedur sebagai berikut :

1. Melakukan Penginputan Data
2. Pra-prosesing
3. Melakukan analisis *Clustering*

Pada tahapan ini di lakukan analisis clustering dengan menggunakan *K-Means* dan *K-Medoids*.

- a) Algoritma *K-Means*
 - 1) Menentukan jumlah klaster yang diinginkan (K) dan acak pilih K sebagai titik pusat awal klaster (centroid).
 - 2) Hitung jarak antara setiap data dengan setiap titik pusat klaster
 - 3) Menghitung jarak *euclidian* objek berdasarkan jarak minimum
 - 4) Mengelompokan objek berdasarkan jarak minimum
 - 5) Jika masih terdapat objek yang berpindah, maka lakukan iterasi dengan mengulang langkah 1 sampai 4
- b) Algoritma *K-Medoids*
 - 1) Menentukan jumlah klaster
 - 2) Dari n titik data, carilah medoid awal sebanyak K
 - 3) Menentukan jarak antara masing-masing objek dan medoid sementara, kemudian mencatat jarak terdekat objek ke medoid dan menghitung jumlahnya.
 - 4) Melakukan iterasi medoid

- 5) Menentukan simpangan keseluruhan (S). Simpangan totalnya adalah $S = b - a$. Jika a adalah banyaknya jarak terpendek antara objek dengan medoid awal dan b adalah banyaknya jarak terpendek antara benda dengan medoid baru. Objek dan data ditukar untuk membuat kumpulan K baru sebagai medoid jika $S < 0$.
 - 6) Jika $S < 0$, ulangi prosedur 3 hingga 5 dan berhenti.
4. Melakukan validasi clustering.
 5. Melakukan Interpretasi Hasil

3. Hasil dan Pembahasan

3.1 Statistik Deskriptif Data Penelitian

Data yang digunakan pada penelitian ini merupakan data Bantuan Langsung Tunai Desa Lupoyo dengan sebanyak 229 sampel yang terdiri dari 5 Variabel yaitu Jenis Pekerjaan (X_1) Tingkat Kesehatan Keluarga (X_2), Status Penerima Bantuan (X_3), Jumlah Anggota Keluarga (X_4) dan Pendapatan Keluarga (X_5). Adapun statistik deskriptif disajikan pada Tabel 1.

Tabel 1. Statistik Deskriptif

Variabel	Min	Median	Mean	Max	Q1	Q2
X_1	1	4	3,572	8	2	4
X_2	2	3	2,956	3	3	3
X_3	1	1	1,476	2	1	2
X_4	0	4	3,594	8	3	5
X_5	0	900.000	997.817	2.500.000	650.000	1.250.000

Berdasarkan Tabel 1, dilihat bahwa pada variabel X_5 terdapat perbedaan rentang nilai yang cukup besar dibandingkan dengan variabel lainnya yang dapat mempengaruhi hasil analisis data, sehingga perlu dilakukan standarisasi data.

3.2 Standarisasi Data

Standarisasi dilakukan untuk mencegah terjadinya bias dalam analisis cluster. Hasil standarisasi disajikan dalam Tabel 2.

Tabel 2. Standarisasi Data

No	X1	X2	X3	X4	X5
1	-1,39232	0,21322	-0,95098	0,954078	-1,13926
2	2,396972	0,21322	-0,95098	0,954078	1,149254
3	-0,30967	0,21322	1,046953	0,954078	-0,22385
4	1,855644	0,21322	-0,95098	0,954078	1,149254
5	-0,30967	0,21322	1,046953	0,275557	0,4627
6	-0,30967	0,21322	1,046953	0,954078	-0,22385
7	-0,851	0,21322	1,046953	-0,40296	-0,45271
8	0,23166	0,21322	1,046953	0,275557	0,119422
:	:	:	:	:	:
299	-1,39232	0,21322	-0,95098	-1,08149	-1,25369

3.3 Pembentukan Cluster K-Means

K-means clustering yaitu dimana K merupakan konstanta jumlah kelompok yang dibentuk. Sedangkan mean adalah rataan dari grup data yang akan disajikan dalam kelompok. Pada penelitian ini jumlah $K = 3$ yaitu cluster 1, cluster 2 dan cluster 3. Setelah itu, menentukan nilai centroid awal yang dipilih secara acak pada 229 data.

Tabel 3. Centroid awal

Cluster	X1	X2	X3	X4	X5
C1	-1,392324	-4,66951	1,046953	-2,438528	-2,283517
C2	-1,392324	0,21322	-0,95098	0,9540776	-1,13926
C3	0,7729881	0,21322	1,046953	2,3111197	1,3781052

Selanjutnya dilakukan perhitungan jarak antara setiap objek dan centroid masing-masing cluster. Perhitungan jarak yaitu jarak Euclidean. Berikut ini adalah perhitungan jarak objek 1 dengan nilai C1, C2 dan C3.

Jarak *Euclidean* data 1 dan C1

$$d(1,C1) = \sqrt{(-1,39232 + 1,392324)^2 + (0,21322 + 4,66951)^2 + (-0,95098 - 1,046953)^2} \\ \sqrt{+(0,954078 + 2,438528)^2 + (-1,13926 + 2,283517)^2} = 6,3758$$

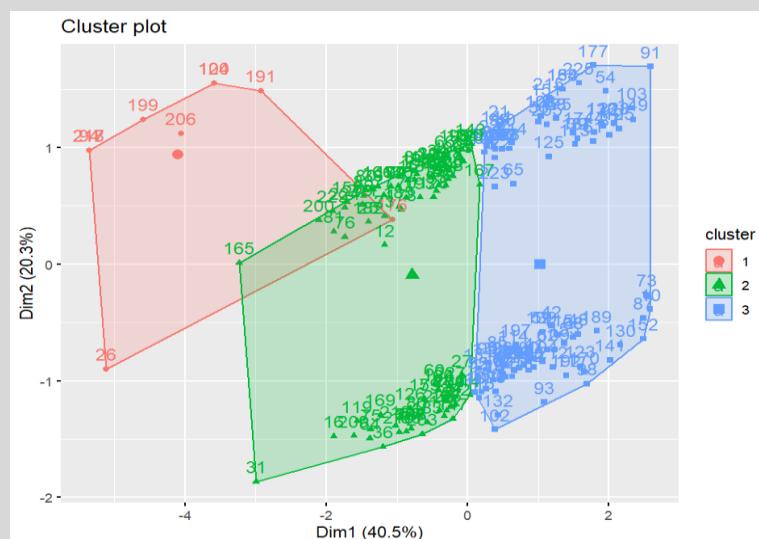
Jarak *Euclidean* data 1 dan C2

$$d(1,C2) = \sqrt{(-1,39232 + 1,392324)^2 + (0,21322 - 0,21322)^2 + (-0,95098 + 0,95098)^2} \\ \sqrt{+(0,954078 - 0,9540776)^2 + (-1,13926 + 1,13926)^2} = 0$$

Jarak *Euclidean* data 1 dan C3

$$d(1,C3) = \sqrt{(-1,39232 - 0,7729881)^2 + (0,21322 - 0,21322)^2 + (-0,95098 - 1,046953)^2} \\ \sqrt{+(0,954078 - 2,3111197)^2 + (-1,13926 - 1,3781052)^2} = 4,1059$$

Perhitungan jarak Euclidean dilakukan sampai pada objek ke 229 dengan setiap centroid. Dari hasil perhitungan diatas diperoleh jarak antara objek 1 dengan masing-masing centroid. Nilai tersebut menunjukkan tingkat kemiripan objek 1 dengan nilai centroid yang ada. Terlihat bahwa nilai terkecil ditunjukkan oleh jarak objek 1 dengan C2 yang artinya objek 1 lebih dekat jaraknya dekat C2 atau cluster kedua. Hal ini berarti, diantara ketiga centroid yang ada objek 1 memiliki kemiripan karakteristik yang lebih tinggi dengan C2, sehingga objek 1 merupakan anggota cluster kedua. Selanjutnya menentukan centroid baru dengan menghitung rata-rata data dalam cluster. Kemudian langkah menghitung jarak euclidean dan cluster dilakukan kembali (melakukan iterasi). Iterasi dihentikan ketika tidak terjadi perubahan cluster atau terjadi konvergensi pada dua iterasi terakhir. Berdasarkan bantuan aplikasi RStudio menghasilkan cluster 1 terdiri dari 10 anggota kepala keluarga, cluster 2 terdiri dari 101 anggota kepala keluarga dan cluster 3 terdiri dari 118 anggota kepala keluarga.



Gambar 1. Hasil K-Means Clustering

3.4 Pembentukan Cluster K-Medoids

Teknik pengelompokan partisi yang disebut K-Medoids clustering berupaya mengidentifikasi sekelompok klaster di antara data yang paling menggambarkan objek dalam kumpulan data. Dalam K-Medoids penentuan titik pusat (centroid) tidak menggunakan rata-rata. Pada penelitian ini jumlah K=3. Dengan centroid data disajikan pada Tabel 4.

Tabel 4. Centroid awal

Cluster	X1	X2	X3	X4	X5
C1	-0,851	0,21322	-0,95098	-0,40296	-0,68156
C2	0,23166	0,21322	-0,95098	0,275557	0,691551
C3	0,23166	0,21322	1,046953	0,275557	0,119422

Selanjutnya dilakukan perhitungan jarak antara setiap objek dan centroid masing-masing cluster. Perhitungan jarak yaitu jarak Euclidean.

Jarak *Euclidean* data 1 dan C1

$$d(1,C1) = \sqrt{(-1,39232 + 0,851)^2 + (0,21322 - 0,21322)^2 + (-0,95098 + 0,95098)^2} \\ \sqrt{+(0,954078 + 0,40296)^2 + (-1,13926 + 0,68156)^2} = 1,531043$$

Jarak *Euclidean* data 1 dan C2

$$d(1,C2) = \sqrt{(-1,39232 - 0,23166)^2 + (0,21322 - 0,21322)^2 + (-0,95098 + 0,95098)^2} \\ \sqrt{+(0,954078 - 0,275557)^2 + (-1,13926 - 0,691551)^2} = 2,539603$$

Jarak *Euclidean* data 1 dan C3

$$d(1,C3) = \sqrt{(-1,39232 + 0,23166)^2 + (0,21322 - 0,21322)^2 + (-0,95098 - 1,046953)^2} \\ \sqrt{+(0,954078 - 0,275557)^2 + (-1,13926 - 0,119422)^2} = 2,945122$$

Dari hasil perhitungan diatas diperoleh jarak antara objek 1 dengan masing-masing centroid. Nilai tersebut menunjukkan tingkat kemiripan objek 1 dengan nilai centroid yang ada. Terlihat bahwa nilai terkecil ditunjukkan oleh jarak objek 1 dengan C1 yang artinya objek 1 lebih dekat jaraknya dekat dengan cluster pertama. Hal ini berarti, diantara ketiga centroid yang ada objek 1 memiliki kemiripan karakteristik yang lebih tinggi dengan C1, sehingga, objek 1 merupakan anggota cluster pertama. Proses perhitungan yang sama dilakukan pada objek 2 sampai dengan objek 229 dengan masing-masing nilai centroid. Kemudian kita hitung total jarak kedekatan. Proses ini merupakan iterasi pertama dari pengelompokan clustering. Dilanjutkan dengan menginisiasikan data sebagai centroid kemudian menghitung kembali jarak euclidean. Proses iterasi berhenti jika simpangan total (S)>0.

Adapun hasil pengelompokan K-Medoids pada cluster 1 terdiri dari 67 kepala keluarga, cluster 2 terdiri dari 59 kepala keluarga dan cluster 3 terdiri dari 103 kepala keluarga.



sehat akan tetapi pada cluster 3 semua kepala keluarga telah menerima bantuan lainnya dengan rata-rata jumlah tanggungan sebanyak 4 orang dan penghasilan sebesar 1.293.644.

Oleh karena itu dapat disimpulkan bahwa untuk cluster 1 dapat dikategorikan sebagai kepala keluarga calon penerima yang di prioritaskan untuk penerimaan BLT, untuk cluster 2 kepala keluarga calon penerima yang dapat dipertimbangkan untuk penerimaan BLT, sedangkan untuk cluster 3 kepala keluarga calon penerima yang tidak layak untuk penerimaan BLT.

4. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa pengelompokan masyarakat menggunakan metode *K-Means* dengan 3 cluster yakni terdiri dari 10 anggota kepala keluarga pada kluster pertama, 101 anggota kepala keluarga pada cluster kedua dan 118 anggota kepala keluarga pada cluster ketiga. Sedangkan pada pengelompokan masyarakat menggunakan metode *K-Medoids* dengan 3 cluster yakni terdiri dari 67 anggota kepala keluarga pada kluster pertama, 59 anggota kepala keluarga kluster kedua dan 103 anggota kepala keluarga pada kluster ketiga. Untuk penetuan metode terbaik dengan menggunakan tiga uji validitas yaitu *Davies-Bouldin Indeks*, *Dunn Indeks*, *Connectivity Indeks* dengan melihat masing-masing nilai indeks diperoleh bahwa metode *K-Means* memberikan hasil yang lebih baik dari pada *K-Medoids* dengan nilai *Davies-Bouldin Indeks* sebesar 1,307 dan nilai *Connectivity Indeks* 40,079.

Referensi

- [1] S. Sari and J. N. Utamajaya, “Sistem Pendukung Keputusan Penerima Bantuan Langsung Tunai Dana Desa Menggunakan Metode Algoritma K-Means Clustering,” *J. JUPITER*, vol. 14, no. 1, pp. 150–160, 2022.
- [2] D. Ortega, “Pelaksanaan Bantuan Langsung Tunai Dari Dana Desa Pada Masyarakat Terdampak Covid-19,” 2021, *Sekolah Tinggi Pembangunan Masyarakat Desa STPMD APMD*.
- [3] Menter Keuangan Republik Indonesia, “Peraturan Menteri Keuangan Republik Indonesia Nomor 201/PMK.07/2022 Tentang Pengelolaan Dana Desa,” p. 1295, 2022.
- [4] I. Sofi, “Efektivitas Bantuan Langsung Tunai Dana Desa Dalam Pemulihan Ekonomi Di Desa,” *Indones. Treas. Rev. J. Perbendaharaan, Keuang. Negara dan Kebijak. Publik*, vol. 6, no. 3, pp. 247–262, 2021, doi: 10.33105/itrev.v6i3.280.
- [5] Z. Bula, R Resmawan, L. O. Nashar, and S. K. Nasib, “Implementasi Imrpoved Chi-Square Automatic Interaction Detection pada Klasifikasi Faktor-Faktor Yang Mempengaruhi Literasi Informasi Generasi Muda,” *J. Stat. App.*, vol. 6, no. 2, pp. 214–222, 2022. doi: <https://doi.org/10.21009/JSA.06207>.
- [6] M. W. Warolemba, R. Resmawan, and D. R. Isa, “Analisis Cluster Fuzzy C-Means dan Diskriminan untuk Pengelompokan Data Kesejahteraan Rakyat,” *SAINSMAT: J. Ilm. Ilmu. Peng. Alam.*, vol. 12, no. 2, pp. 141–152, 2023. doi: <https://doi.org/10.35580/sainsmat122446492023>.
- [7] I. K. Hasan, R. Resmawan, and J. Ibrahim, “Perbandingan K-Nearest Neighbor dan Random Forest dengan Seleksi Fitur Information Gain untuk Klasifikasi Lama Studi Mahasiswa,” *Indonesian. J. App. Stat.*, vol. 5, no. 1, pp. 58–66, 2022. doi: <https://doi.org/10.13057/ijas.v5i1.58056>.
- [8] D. Widaydhana, R. B. Hastuti, I. Kharisudin, and F. Fauzi, “Perbandingan Analisis Klaster K-Means dan Average Linkage untuk Pengklasteran Kemiskinan di Provinsi Jawa Tengah,” *Prism. Pros. Semin. Nas. Mat.*, vol. 4, pp. 584–594, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [9] R. Yuliani, “Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur,” *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 39–50, 2021.