

# PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK ANALISIS POLA POPULASI TERNAK DI KOTA BANDUNG BERBASIS DATA MINING

Depi Lasandi<sup>1</sup>, Rini Astuti<sup>2</sup>, Khaerul Anam<sup>3</sup>, Aris Pratama Putra<sup>4</sup>, Bani Nurhakim<sup>5</sup>

<sup>1</sup> Program Studi Sistem Informasi STMIK IKMI Cirebon ; Jl.Perjuangan No.10 B, Majasem, Cirebon , Jawa Barat.

<sup>2</sup> Program Studi Sistem Informasi STMIK LIKMI Bandung ; Jl. Ir.H. Juanda No. 96, Lebakgede, Coblong, Bandung Jawa Barat.

<sup>3</sup>Program Studi Teknik Informatika STMIK IKMI Cirebon ; Jl.Perjuangan No.10 B, Majasem, Cirebon , Jawa Barat.

<sup>4</sup> Program Manajemen Informatika STMIK IKMI Cirebon ; Jl.Perjuangan No.10 B, Majasem, Cirebon , Jawa Barat.

<sup>5</sup>Program Studi Rekayasa Perangkat Lunak STMIK IKMI Cirebon ; Jl.Perjuangan No.10 B, Majasem, Cirebon , Jawa Barat.

## Keywords:

Algoritma K-Means;  
Algoritma Klasterisasi;  
Penambangan Data; Analisis  
Pola; Kompleksitas  
Komputasi

## Correspondent Email:

lasandidepi@gmail.com

Abstrak - Perkembangan teknologi informatika memungkinkan pengolahan data besar untuk mendukung pengambilan keputusan berbasis bukti. Populasi ternak di Kota Bandung penting bagi pembangunan pertanian dan peternakan, namun data yang tersedia masih bersifat deskriptif. Penelitian ini bertujuan mengidentifikasi pola perubahan populasi ternak menggunakan algoritma K-Means Clustering dalam kerangka Knowledge Discovery in Databases (KDD). Data dikumpulkan dari BPS dan DKPP Kota Bandung periode 2021–2024, kemudian melalui preprocessing, normalisasi, dan transformasi ke format numerik. Analisis dilakukan dengan K-Means menggunakan Euclidean distance,  $K = 5$ , max runs = 10, dan max optimization steps = 100. Hasil klasterisasi membentuk lima cluster dengan distribusi tidak merata; satu cluster dominan, sementara cluster lain menunjukkan karakteristik wilayah spesifik. Temuan ini memberikan gambaran pola pertumbuhan populasi ternak di tingkat kecamatan dan menjadi dasar bagi pemerintah daerah dalam merumuskan kebijakan pengelolaan ternak yang lebih efektif dan berbasis data.

*Abstract- The development of information technology enables the processing of large-scale data to support evidence-based decision making. Livestock populations in Bandung City are important for agricultural and livestock development; however, the available data remain largely descriptive. This study aims to identify patterns of livestock population changes using the K-Means Clustering algorithm within the Knowledge Discovery in Databases (KDD) framework. Data were collected from BPS and DKPP Bandung City for the period 2021–2024, followed by preprocessing, normalization, and transformation into a numerical format. Analysis was performed using K-Means with Euclidean distance,  $K = 5$ , max runs = 10, and max optimization steps = 100. The clustering results formed five clusters with uneven distribution; one cluster dominated, while the others exhibited specific regional characteristics. These findings provide insights into livestock population growth at the sub-district level and can serve as a basis for local government in formulating more effective, data-driven livestock management policies.*



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

## 1. PENDAHULUAN

Perkembangan teknologi informatika telah membawa transformasi digital yang signifikan dalam berbagai sektor, termasuk pemerintahan daerah[1]. Pemanfaatan teknologi seperti data mining, machine learning, dan big data analytics memungkinkan pengolahan data dalam jumlah besar secara lebih sistematis, cepat, dan akurat untuk mendukung pengambilan keputusan berbasis bukti (data-driven decision making)[2]. Dalam konteks sektor pertanian dan peternakan, ketersediaan data populasi ternak seharusnya dapat dimanfaatkan secara optimal untuk perencanaan pembangunan yang lebih terarah dan berkelanjutan[3]. Pengelolaan data yang baik tidak hanya meningkatkan efisiensi administrasi, tetapi juga membantu pemerintah dalam merumuskan kebijakan strategis berbasis analisis kuantitatif yang objektif.[4]

Di Kota Bandung, data populasi ternak seperti sapi, kambing, ayam, dan domba telah dikumpulkan secara rutin oleh instansi terkait, seperti Badan Pusat Statistik (BPS) dan Dinas Ketahanan Pangan dan Pertanian. Namun, data tersebut umumnya masih disajikan dalam bentuk tabel deskriptif tanpa analisis pola yang mendalam, sehingga belum mampu menggambarkan dinamika pertumbuhan atau penurunan populasi ternak antar kecamatan secara komprehensif[5]. Kondisi ini berpotensi menyebabkan pengambilan kebijakan dilakukan berdasarkan asumsi atau pendekatan konvensional, bukan hasil analisis berbasis data. Padahal, penerapan teknik data mining dapat mengidentifikasi pola tersembunyi serta mengelompokkan wilayah berdasarkan karakteristik populasi ternak, sehingga strategi pengembangan dapat disesuaikan dengan kondisi riil di lapangan[6]. Meskipun beberapa penelitian sebelumnya telah menerapkan metode klasterisasi pada data pertanian dan peternakan, sebagian besar masih berfokus pada skala provinsi atau nasional dan belum banyak mengkaji konteks peternakan perkotaan secara spesifik[7].

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk menganalisis pola perubahan populasi ternak di Kota Bandung menggunakan pendekatan kuantitatif dengan metode data mining, khususnya algoritma K-Means Clustering[8]. Metode ini dipilih karena mampu mengelompokkan data berdasarkan

kesamaan karakteristik sehingga terbentuk klaster wilayah dengan tingkat pertumbuhan populasi ternak yang serupa. Proses penelitian meliputi tahapan pengumpulan data dari sumber resmi, pembersihan dan transformasi data, penerapan algoritma K-Means, serta interpretasi hasil klasterisasi untuk menghasilkan pemetaan pola perkembangan populasi ternak[9]. Hasil penelitian diharapkan dapat memberikan kontribusi akademik dalam pengembangan penerapan data mining pada bidang Sistem Informasi, sekaligus memberikan manfaat praktis bagi pemerintah daerah dalam merancang kebijakan pengelolaan sektor peternakan yang lebih efektif, berbasis data, dan berorientasi pada pembangunan berkelanjutan.[10]

## 2. TINJAUAN PUSTAKA

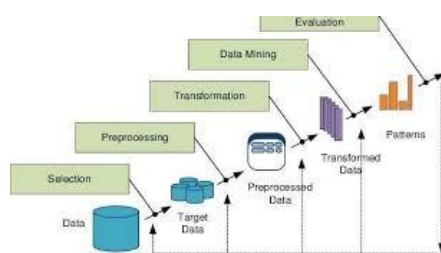
### 2.1 Clustering

Clustering adalah salah satu metode dalam data mining yang bertujuan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan tingkat kesamaan karakteristik tanpa memerlukan label atau kelas tertentu (unsupervised learning)[11]. Prinsip dasar teknik ini adalah meningkatkan kemiripan antar data dalam satu kelompok (intra-cluster similarity) serta memperbesar perbedaan antar kelompok yang berbeda (inter-cluster dissimilarity), sehingga pola yang sebelumnya tersembunyi dapat diidentifikasi secara terstruktur. Secara umum, clustering terdiri atas beberapa pendekatan, seperti partitioning methods (misalnya K-Means), hierarchical methods, dan density-based methods seperti DBSCAN, yang penggunaannya disesuaikan dengan karakteristik serta distribusi data. Dalam pengolahan data numerik berskala besar, algoritma K-Means banyak diterapkan karena prosedurnya relatif sederhana, komputasinya efisien, dan mampu menghasilkan pengelompokan melalui tahapan penentuan jumlah cluster (k), perhitungan jarak ke pusat cluster (centroid), serta proses iterasi hingga tercapai kondisi stabil. Dalam konteks sektor pertanian dan peternakan, teknik clustering dapat dimanfaatkan untuk menganalisis pola persebaran populasi ternak serta karakteristik pertumbuhan wilayah, sehingga mendukung perumusan kebijakan yang berbasis data. Meski demikian, untuk memperoleh hasil klasterisasi

yang valid dan dapat dipertanggungjawabkan secara ilmiah, diperlukan tahapan normalisasi data serta evaluasi kualitas cluster menggunakan metrik pengujian yang sesuai.[12]

## 2.2 Algoritma *K-Means*

Algoritma K-Means adalah salah satu teknik clustering berbasis partisi yang digunakan untuk mengelompokkan data ke dalam sejumlah cluster ( $k$ ) berdasarkan kedekatan terhadap pusat cluster atau centroid. Tahap awal metode ini dimulai dengan menentukan jumlah cluster yang akan dibentuk, kemudian menginisialisasi centroid awal, baik secara acak maupun dengan pendekatan tertentu[13]. Selanjutnya, setiap data dihitung jaraknya terhadap seluruh centroid, biasanya menggunakan metode Euclidean Distance, lalu ditempatkan pada cluster dengan jarak paling dekat. Setelah proses pengelompokan awal selesai, nilai centroid diperbarui dengan menghitung rata-rata data dalam masing-masing cluster. Langkah ini dilakukan secara berulang hingga tidak terjadi perubahan keanggotaan data atau hingga sistem mencapai kondisi stabil (konvergen)[14]. K-Means banyak diterapkan karena prosesnya relatif sederhana, memiliki efisiensi komputasi yang baik untuk data numerik dalam jumlah besar, serta menghasilkan struktur cluster yang mudah dianalisis dan diinterpretasikan dalam berbagai kebutuhan penelitian.



Gambar 1.1 Knowledge Discovery in Databases

## 3. METODE PENELITIAN

Knowledge Discovery in Databases (KDD) sebagai kerangka kerja dalam proses pengolahan data untuk menghasilkan informasi yang bermakna[15]. Tahapan diawali dengan selection, yaitu proses pemilihan dan pengumpulan data yang relevan dari sumber resmi seperti BPS dan instansi terkait.

Selanjutnya dilakukan preprocessing yang mencakup pembersihan data dari kesalahan, data ganda, maupun nilai yang tidak lengkap, agar kualitas data terjamin.[16] Tahap berikutnya adalah transformation, yaitu penyesuaian atau normalisasi data ke dalam format yang sesuai untuk proses analisis. Setelah itu dilakukan tahap inti, yaitu data mining, yang merupakan proses ekstraksi pola atau pengetahuan tersembunyi dari data menggunakan teknik tertentu, dalam penelitian ini melalui penerapan algoritma *K-Means Clustering* untuk mengelompokkan data berdasarkan kesamaan karakteristik. Tahap akhir adalah interpretation and evaluation, di mana hasil pengelompokan dianalisis dan dievaluasi untuk memastikan validitasnya serta diinterpretasikan menjadi informasi yang dapat mendukung pengambilan keputusan berbasis data.

### 3.1. Selection

Tahapan ini mencakup proses pengumpulan serta verifikasi data yang diperoleh dari berbagai sumber resmi, seperti Badan Pusat Statistik (BPS) dan Dinas Ketahanan Pangan dan Pertanian (DKPP) Kota Bandung[17]. Data yang dihimpun meliputi jumlah populasi ternak berdasarkan jenisnya, antara lain sapi, kambing, ayam, domba, dan itik, dalam rentang waktu empat tahun. Selanjutnya, data tersebut melalui tahap data cleaning yang bertujuan untuk menghilangkan duplikasi, menangani nilai yang hilang (missing values), serta memastikan konsistensi dan keseragaman format agar siap digunakan dalam proses analisis..

#### 3.1.1. Data Normalization

Proses transformasi data pada penelitian ini dilakukan dengan menerapkan normalisasi untuk menyeragamkan skala antar atribut, sehingga tidak ada variabel yang mendominasi hasil klasterisasi akibat perbedaan rentang nilai. Langkah ini krusial dalam algoritma K-Means karena perhitungan jarak Euclidean sangat peka terhadap variasi skala data[18]. Oleh sebab itu, digunakan metode Min-Max Normalization untuk mengonversi nilai asli ke dalam rentang tertentu. Dengan proses ini, setiap variabel berada pada skala yang

seimbang sehingga hasil pengelompokan menjadi lebih adil, konsisten, dan representatif.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

di mana  $X'$  adalah nilai hasil normalisasi,  $X_{min}$  nilai minimum, dan  $X_{max}$  nilai maksimum dari atribut.

Gambar 1.2 Data *Normalization*

### 3.1.2. Data Mining

Tahap data mining merupakan proses utama dalam analisis yang berfokus pada penggalian pola atau pengetahuan tersembunyi dari data yang telah melewati tahap seleksi, pembersihan, dan transformasi[19]. Dalam penelitian ini, teknik yang digunakan adalah algoritma K-Means Clustering untuk mengelompokkan data berdasarkan tingkat kesamaan karakteristiknya. Melalui perhitungan jarak antara data dan pusat cluster (centroid), proses dilakukan secara berulang hingga terbentuk kelompok dengan pola yang serupa. Hasil pengelompokan tersebut kemudian digunakan sebagai dasar analisis lebih lanjut dalam mendukung pengambilan keputusan yang berbasis data.

### 3.1.3 Interpretasi

Tahap terakhir adalah menginterpretasikan hasil klasterisasi untuk mengidentifikasi pola perubahan populasi ternak di Kota Bandung selama periode 2021–2024. Setiap cluster dianalisis untuk melihat karakteristik serta kecenderungan pertumbuhan atau penurunan populasi di masing-masing wilayah. Hasilnya kemudian divisualisasikan dalam bentuk diagram batang, scatter plot, atau grafik garis guna mempermudah pemahaman terhadap distribusi dan tren populasi setiap jenis ternak.

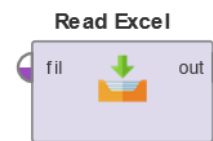
## 4. HASIL DAN PEMBAHASAN

Tahapan dalam ini mengikuti metodologi yang diterapkan, yaitu metodologi KDD (*knowledge Discovery in Databases*)

### 4.1 Data Selection

Pada tahap ini dilakukan proses import data ke dalam perangkat lunak RapidMiner agar dataset dapat digunakan dalam proses analisis. Dataset yang dimasukkan bernama

“Data\_dataset\_ternak\_bandung.xlsx”, yang berisi data populasi ternak berdasarkan jenis dan wilayah. Tujuan dari proses ini adalah untuk memuat data ke dalam repository RapidMiner sehingga sistem dapat mengidentifikasi struktur atribut, tipe variabel, serta jumlah data yang tersedia. Setelah proses impor selesai, data siap untuk melalui tahapan berikutnya seperti preprocessing, transformasi, dan penerapan algoritma K-Means Clustering.



Gambar 4.1 Operator Read Excel pada RapidMiner

### 4.2 Preprocessing

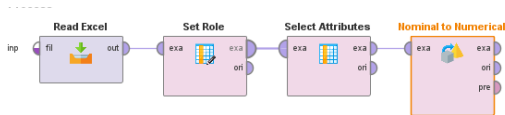
Tahap preprocessing meliputi proses pembersihan data, khususnya untuk menangani kemungkinan adanya nilai yang hilang (missing values). Sebelum dilakukan penanganan lebih lanjut, terlebih dahulu dilakukan evaluasi terhadap atribut dalam dataset untuk memastikan ada atau tidaknya data yang kosong. Berdasarkan hasil statistik dataset yang ditampilkan pada gambar, dapat disimpulkan bahwa seluruh atribut tidak mengandung nilai yang hilang, sehingga data dapat langsung digunakan pada tahap analisis berikutnya.

Tabel 4.2 *Result dari statistik Set Role*

<i>NAME</i>	<i>TYPE</i>	<i>MISSING</i>
BPS_NAMA_KECAMATAN,	<i>NOMINAL</i>	0
KATEGORI_TERNAK,	<i>NOMINAL</i>	0
JUMLAH	<i>INTEGER</i>	0
TAHUN	<i>INTEGER</i>	0

### 4.3 Data Transformation

Pada tahap ini dilakukan transformasi data agar seluruh atribut kompatibel dengan algoritma K-Means yang hanya memproses data numerik. Oleh karena itu, atribut berbentuk kategori atau nominal dikonversi menjadi numerik menggunakan operator Nominal to Numeric pada RapidMiner. Proses ini bertujuan menyederhanakan representasi data sehingga dapat digunakan sebagai input dalam proses klasterisasi.



Gambar 4.3 Model Proses *Transformation* di *RapidMiner*.

#### 4.4 Data Mining

Pada tahap data mining, diterapkan algoritma K-Means Clustering untuk membentuk model pengelompokan data. Proses pengolahan data dilakukan untuk mengelompokkan data berdasarkan pola tertentu,

#### 4.5 Evaluation

Melalui proses yang dilakukan, diperoleh lima cluster utama yang menunjukkan bahwa populasi ternak di Kota Bandung dapat dikelompokkan ke dalam lima kategori berdasarkan kesamaan karakteristik nilai numeriknya. Proses pengelompokan menggunakan Euclidean Distance sebagai ukuran jarak antara data dan centroid, sehingga setiap wilayah tergabung dalam cluster dengan tingkat kemiripan paling tinggi. Model cluster yang terbentuk selanjutnya digunakan sebagai dasar untuk analisis dan interpretasi pola data.

Tabel 4.4 Model cluster hasil pengelompokan

CLUSTER MODEL	
Cluster 0	542 items
Cluster 1	42 items
Cluster 2	3 items
Cluster 3	12 items

Cluster 4	1 item
Total Jumlah Data 600 items	

Hasil *cluster model* menunjukkan bahwa dari 600 data terbentuk lima cluster dengan distribusi yang tidak seimbang. Cluster 0 mendominasi dengan 542 item, yang mengindikasikan bahwa sebagian besar wilayah memiliki karakteristik populasi ternak yang relatif homogen[20]. Sementara itu, cluster lainnya memiliki jumlah anggota yang jauh lebih sedikit, dengan beberapa cluster hanya terdiri dari sedikit data, yang menunjukkan adanya pola berbeda atau karakteristik yang lebih spesifik[21]. Ketimpangan ini mengindikasikan konsentrasi populasi pada kategori tertentu serta kemungkinan adanya wilayah dengan nilai ekstrem (*outlier*) dibandingkan mayoritas data.

Berdasarkan evaluasi performa clustering menggunakan operator Performance (*Clustering*) pada RapidMiner, diperoleh nilai Davies-Bouldin Index (DBI) sebesar -0,280. Rata-rata jarak data terhadap centroid dalam masing-masing cluster (*Average Within Centroid Distance*) tercatat sebesar -12021,963. Secara rinci, nilai rata-rata jarak pada Cluster 0 sebesar -4197,862, Cluster 1 sebesar -71333,554, Cluster 2 sebesar -32671,556, Cluster 3 sebesar -153659,410, dan Cluster 4 sebesar -0,000. Evaluasi ini dilakukan dengan algoritma K-Means menggunakan ukuran jarak Euclidean Distance, dengan parameter K = 5, jumlah max runs sebanyak 10 kali, serta max optimization steps sebesar 100.

Performance Vector	Hasil
Avg. Within Centroid Distance	-12021,963
Avg. Within Centroid Distance (Cluster 0)	-4197,862
Avg. Within Centroid Distance (Cluster 1)	-71333,554
Avg. Within Centroid Distance (Cluster 2)	-32671,556
Avg. Within Centroid Distance (Cluster 3)	-153659,410
Avg. Within Centroid Distance (Cluster 4)	-0,000
<b>Davies-Bouldin Index (DBI)</b>	<b>-0,280</b>

Gambar 4.4 Hasil *Running Operator Performance*

## 5. KESIMPULAN

Berdasarkan rumusan masalah dan hasil penelitian yang telah dilakukan, maka dapat disimpulkan bahwa:

- a. Penerapan algoritma K-Means dalam kerangka KDD efektif digunakan untuk mengelompokkan data populasi ternak di Kota Bandung berdasarkan kesamaan karakteristik numeriknya.
- b. Proses klasterisasi menghasilkan lima kelompok, dengan satu cluster utama yang mencerminkan bahwa sebagian besar wilayah memiliki pola populasi ternak yang serupa.
- c. Pendekatan ini dapat dijadikan sebagai alat analisis untuk memahami pola distribusi populasi ternak serta mendukung perumusan kebijakan berbasis data di sektor peternakan

## UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak terkait yang telah memberi dukungan terhadap penelitian ini.

## DAFTAR PUSTAKA

- [1] A. Rezi and M. Allam, "Techniques in array processing by means of transformations, " in *Control and Dynamic Systems*, Vol. 69, Multidimensional Systems, C. T. Leondes, Ed. San Diego: Academic Press, 1995, pp. 133-180.
- [2] W.K. Chen, *Linear Networks and Systems*. Belmont, CA: Wadsworth, 1993, pp. 123-135
- [3] T. J. van Weert and R. K. Munro, Eds., *Informatics and the Digital Society: Social, ethical and cognitive issues: IFIP TC3/WG3.1&3.2 Open Conf.e on Social, Ethical and Cognitive Issues of Informatics and ICT*, July 22-26, 2002, Dortmund, Germany. Boston: Kluwer Academic, 2003.
- [4] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*, 2nd ed. Reading, MA: Addison Wesley, 2003. [Online] Available: Safari e-book.
- [5] M. W. Dixon, "Application of neural networks to solve the routing problem in communication networks," Ph.D. dissertation, Murdoch Univ., Murdoch, WA, Australia, 1999.
- [6] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier", *JITET*, vol. 10, no. 1, Jan. 2022.
- [1] S. Studi, K. Di, P. Jawa, and E. T. Setyasih, "TRANSFORMASI DIGITAL PEMERINTAH DAERAH DI ERA," vol. 5, no. 3, pp. 59–66.
- [2] P. Gonnade and S. Ridhorkar, "INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Business Decision making through Big Data Analytics using Machine Learning Technique," vol. 12, no. 4, pp. 3057–3063, 2025.
- [3] S. Pembibitan *et al.*, "Analisis Potensi Pakan untuk Pengembangan Ternak Ruminansia di Provinsi Jawa Timur Analysis of Feed Potency for Development of Ruminant Livestock in East Java Province," vol. 15, no. 3, pp. 251–258, 2020.
- [4] P. Ayu, W. Purnama, N. Pohan, and A. Restiady, "Predicting New Student Admissions with the SVM Regression Model in Data Mining," vol. 8, no. 2, pp. 843–850, 2025.
- [5] D. Dari, P. Teknologi, and I. Dan, "Dampak dari perkembangan teknologi informasi dan komunikasi," vol. 3, pp. 206–210, 2022.
- [6] I. I. N. Ernawati, "Jurnal Teknologi Informasi dan Pendidikan," vol. 12, no. 2, 2019.
- [7] R. R. Marbun, C. Maisyarah, P. Studi, and S. Informasi, "No Title," vol. 8, pp. 2084–2091, 2025.
- [8] K. B. Agbemadon *et al.*, "Churn detection using machine learning in the retail industry To cite this version: HAL Id: hal-04257366 Churn detection using machine learning in the retail industry," 2023.
- [9] W. B. Laksono, Y. Syahidin, and Y. Yunengsih, "Implementasi Data Mining Klasterisasi Data Pasien Rawat Inap dengan Algoritma K-Means Clustering," vol. 7, no. 2, pp. 621–627, 2024, doi: 10.32493/jtsi.v7i2.39354.
- [10] B. Data, A. Dan, P. Di, and P. Dan, "Jurnal Ekonomi dan Bisnis Dharma Andalas," vol. 25, no. 1, pp. 185–194, 2023.
- [11] D. Rizaldy, A. Fauzan, E. Rilvani, and U. P. Bangsa, "Analisa Dampak Kinerja Menteri ESDM terhadap Kesejahteraan Masyarakat : Studi Data Mining dengan Metode Clustering," vol. 03, no. 04, pp. 2413–2426, 2025.
- [12] M. I. St Nur Rahmah, Muliani.S, Andi

- Nilwana, "Jurnal Ilmiah Pemerintahan," vol. 12, no. Idm, pp. 27–38, 2024.
- [13] D. A. Setiady and H. Leong, "Implementation of K-Means Algorithm Elbow Method and Silhouette Coefficient for Rainfall Classification," *Proxies J. Inform.*, vol. 4, no. 1, pp. 18–25, 2024, doi: 10.24167/proxies.v4i1.12433.
- [14] D. A. Tarigan, "Optimization of the K-Means Clustering Algorithm Using Davies Bouldin Index in Iris Data Classification," *Media Online*, vol. 4, no. 1, pp. 545–552, 2023, doi: 10.30865/klik.v4i1.964.
- [15] F. Juliawati, R. Buatun, R. Saragih, and S. Kaputama, "Pengelompokan Data Mining Penerimaan Bantuan Pangan Non Tunai (BPNT) Menggunakan Metode Clustering (Studi Kasus: Kantor Desa Payabakung Hampan Perak)," *J. Comput. Sci. Inf. Technol. E-ISSN*, vol. 3, no. 2, p. 69, 2023.
- [16] A. Singh, S. Mittal, P. Malhotra, and Y. Srivastava, *Clustering Evaluation by Davies-Bouldin Index(DBI) in Cereal data using K-Means*. 2020. doi: 10.1109/ICCMC48092.2020.ICCMC-00057.
- [17] O. Gavrilenko, O. Khomenko, O. S. Zhurakovska, A. Piskun, and Y. Khavikova, "Establishing the grouping principle of public services based on the analysis of similarity coefficients," *Eastern-European J. Enterp. Technol.*, 2023, doi: 10.15587/1729-4061.2023.280218.
- [18] D. Shifman, I. Cohen, K. Huang, X. Xian, and G. Singer, "An adaptive machine learning algorithm for the resource-constrained classification problem," *Eng. Appl. Artif. Intell.*, 2023, doi: 10.1016/j.engappai.2022.105741.
- [19] W. Hu, F. Liu, and J. Peng, "An Efficient Data Classification Decision Based on Multimodel Deep Learning," *Comput. Intell. Neurosci.*, 2022, doi: 10.1155/2022/7636705.
- [20] C.-K. Bogor, "Implementasi Algoritma K-Means Clustering dengan Jarak Euclidean dalam Mengelompokkan Daerah Penyebaran," vol. 2, no. 1, pp. 47–56, 2023.
- [21] A. Anas and A. J. Zebua, "Clustering Nilai Mahasiswa Menggunakan Algoritma K-Means," vol. 19, no. 1, pp. 7–14, 2025.