

Klasterisasi Data Produksi Daging Sapi Menggunakan Algoritma K-Means Orange Data Mining

Achmes Dade Ramadani¹, Farras Hilmy Ibrahim², Manarul Hidayat³, Ahmad Habibullah⁴,
Sumanto⁵, Andi Diah Kuswanto⁶

Informatika, Teknik Informatika, Universitas Bina Sarana Informatika

¹dade.ramadani@gmail.com, ²farrashilmy90@gmail.com,

³manarulh093@gmail.com, ⁴ahmadhabib3336@gmail.com, ⁵sumanto.sto@bsi.ac.id, ⁶andi.ahk@bsi.ac.id

Abstract

Beef production is an important component of the livestock sector that supports national food security. Given the fluctuations in production from year to year and the differences in characteristics between regions, an appropriate analysis method is needed to process the data effectively. This research aims to cluster beef production data in Indonesia for the period 2021 to 2024 using the K-Means Orange Data Mining algorithm. The analysis process follows the CRISP-DM stages, starting from business understanding to deployment. The data used was obtained from the Central Bureau of Statistics and processed to produce three main clusters: regions with high, low, and medium beef production. The results showed that Orange Data Mining's K-Means algorithm was able to effectively group beef production data into different clusters. Orange Data Mining also assists the data analysis process with an innovative visual interface and easy-to-interpret results. The findings are expected to serve as a reference in the formulation of strategic livestock policies and data-driven production distribution planning. The clustering results provide the government with an overview of the level of beef production in each region, enabling more appropriate policies or strategic measures to be taken based on the clustering results.

Keywords: Beef Production Clustering, K-Means Algorithm, Orange Data Mining.

Abstrak

Produksi daging sapi merupakan salah satu komponen penting dalam sektor peternakan yang mendukung ketahanan pangan nasional. Mengingat fluktuasi produksi dari tahun ke tahun dan perbedaan karakteristik antar wilayah, diperlukan metode analisis yang tepat untuk mengolah data secara efektif. Penelitian ini bertujuan untuk mengelompokkan data produksi daging sapi di Indonesia selama periode 2021 hingga 2024 menggunakan algoritma K-Means Orange Data Mining. Proses analisis mengikuti tahapan CRISP-DM, mulai dari pemahaman bisnis hingga deployment. Data yang digunakan diperoleh dari Badan Pusat Statistik dan diproses untuk menghasilkan tiga kluster utama: wilayah dengan produksi daging sapi tinggi, rendah, dan sedang. Hasil penelitian menunjukkan bahwa algoritma K-Means Orange Data Mining mampu mengelompokkan data produksi daging sapi secara efektif ke dalam beberapa kluster yang berbeda. Orange Data Mining turut membantu proses analisis data dengan tampilan antarmuka visual yang inovatif dan hasil yang mudah diinterpretasikan. Temuan ini diharapkan menjadi acuan dalam perumusan kebijakan strategis peternakan dan perencanaan distribusi produksi berbasis data. Hasil klasterisasi ini memberikan gambaran kepada pemerintah mengenai tingkat produksi daging sapi di setiap wilayah, sehingga memungkinkan pengambilan kebijakan atau langkah-langkah strategis yang lebih tepat dan sesuai dengan kondisi masing-masing wilayah berdasarkan hasil klasterisasi.

Kata kunci: Klasterisasi Produksi daging Sapi, Algoritma K-Means, Orange Data Mining.



1. Pendahuluan

Sektor peternakan merupakan bagian penting dalam pembangunan pertanian dan ketahanan pangan nasional. Salah satu komoditas utama dalam subsektor peternakan adalah sapi, baik sapi potong maupun sapi perah. Produksi sapi memiliki peran strategis dalam memenuhi kebutuhan daging dan susu nasional, serta menjadi sumber pendapatan bagi peternak rakyat di berbagai daerah di Indonesia [1].

Pemerintah menghadapi tantangan dalam menghitung dan mengelompokkan total produksi daging sapi di seluruh Indonesia akibat keterbatasan ketersediaan data dan akurasi data produksi di berbagai daerah. Secara global, permintaan terhadap daging sapi terus meningkat seiring pertumbuhan populasi dan perubahan pola konsumsi masyarakat. *Food and Agriculture Organization* (FAO) memproyeksikan bahwa konsumsi daging sapi akan meningkat secara signifikan, terutama di negara-negara berkembang termasuk Indonesia, yang menghadapi tantangan untuk mencukupi permintaan domestik secara berkelanjutan [2]. Kondisi ini menegaskan pentingnya penguatan produksi dalam negeri, baik dari sisi kuantitas, distribusi, maupun efisiensi.

Berdasarkan data dari Direktorat Jenderal Peternakan dan Kesehatan Hewan, produksi sapi menunjukkan fluktuasi dari tahun ke tahun yang dipengaruhi oleh berbagai faktor seperti cuaca, ketersediaan pakan, kebijakan pemerintah, serta manajemen peternakan [3]. Oleh karena itu, analisis data produksi daging sapi secara komprehensif menjadi penting untuk memetakan wilayah-wilayah yang memiliki karakteristik produksi yang serupa. Hal ini dapat membantu dalam perencanaan distribusi sumber daya, pengambilan kebijakan, serta peningkatan efisiensi produksi di daerah-daerah tertentu.

Dalam upaya memahami pola produksi daging sapi di berbagai daerah di Indonesia, pengolahan data produksi memerlukan teknik analisis yang mampu mengekstraksi informasi penting dari kumpulan data yang kompleks. Salah satu pendekatan yang relevan dalam konteks ini adalah data mining, yaitu suatu tingkatan proses untuk mencari nilai tambah dari kumpulan data yang didalamnya terdapat pengetahuan yang selama ini tidak diketahui secara manual. Informasi penting ini didapat dari suatu proses yang amat rumit seperti menggunakan artificial intelligence, teknik statistik, ilmu matematika, machine learning, dan lain sebagainya. Teknik-teknik rumit tersebut nantinya akan

mengidentifikasi dan mengekstraksi informasi yang bermanfaat dari suatu database besar [4].

Salah satu teknik dalam data mining yang banyak digunakan dalam pengelompokan data adalah klusterisasi (clustering). Klusterisasi merupakan suatu teknik untuk mengelompokkan data ke dalam suatu kluster tertentu yang memungkinkan data dalam kluster tersebut memiliki kesamaan dan memiliki perbedaan yang jelas dengan data pada kluster lainnya [5]. Masing-masing klusterisasi tersebut adalah untuk mencari pola pengelompokan data. Sehingga pada akhirnya hasil pengelompokan data tersebut dapat digunakan untuk mendukung pengambilan keputusan bagi manajemen lembaga itu sendiri [6].

Salah satu algoritma klusterisasi yang banyak digunakan dalam penelitian data mining adalah K-Means Clustering. Algoritma K-Means merupakan studi formal untuk mengelompokkan atau clustering benda-benda sesuai dengan karakteristik yang diukur berdasarkan kemiripan satu sama lain. Clustering adalah pengelompokan menggunakan teknik unsupervised learning dimana tidak diperlukan fase learning serta tidak menggunakan pelabelan pada setiap kelompok. Metode clustering mempartisi data ke dalam kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama [7]. K-Means merupakan metode klusterisasi dengan basis jarak dengan melakukan pembagian data ke beberapa cluster serta algoritma ini bekerja di atribut numerik [8].

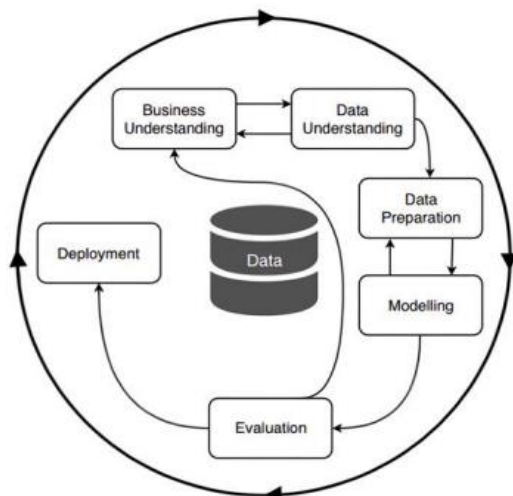
Penelitian ini menghadirkan kebaruan dengan menerapkan algoritma K-Means yang diimplementasikan melalui perangkat lunak Orange Data Mining untuk mengelompokkan data produksi daging sapi di berbagai wilayah Indonesia selama kurun waktu 2021 hingga 2024. Penggunaan Orange Data Mining memberikan keunggulan tersendiri karena kemampuannya dalam melakukan pengolahan data secara visual dan interaktif, sehingga memudahkan proses interpretasi hasil klusterisasi oleh para pengambil kebijakan [9]. Dengan pendekatan ini, setiap wilayah akan dikelompokkan ke dalam kluster-kluster yang merepresentasikan kesamaan dalam karakteristik produksinya. Hasil analisis klusterisasi ini diharapkan dapat menjadi acuan penting bagi para pemangku kebijakan dan pelaku industri peternakan dalam merumuskan strategi pembangunan dan pengembangan sektor peternakan yang lebih terarah, efisien, dan sesuai dengan potensi daerah masing-masing.

Secara keseluruhan, pemanfaatan algoritma K-Means dalam Orange Data Mining pada penelitian ini diharapkan mampu mengubah data produksi daging sapi yang semula bersifat mentah dan tersebar menjadi informasi yang bernilai strategis. Hal ini sejalan dengan arah transformasi digital dalam sektor pertanian dan peternakan, yang menuntut pemanfaatan teknologi data untuk meningkatkan efisiensi, produktivitas, dan ketahanan pangan nasional secara menyeluruh.

2. Metode Penelitian

Metode atau cara kerja ilmiah dalam penelitian merupakan sarana atau alat untuk mencapai tujuan penelitian. Oleh karenanya, sebagai alat, maka dapat dipergunakan secara fleksibel, tidak rigid atau kaku, sehingga cara penelitian dapat dilakukan secara variatif, tergantung pada objek format disiplin ilmu yang diteliti [10].

Penelitian ini mengadopsi metode *Cross Industry Standard Process for Data Mining* (CRISP-DM) yang terdiri dari tahap business understanding, tahap data understanding, tahap data preparation, tahap modeling, tahap evaluation, dan tahap deployment. [11].



Gambar 1. Metode CRISP-DM [14]

2.1 Business Understanding

Pada tahap ini dilakukan pemahaman atau analisa latar belakang pada klusterisasi data produksi daging sapi guna mengidentifikasi permasalahan yang ada. Kemudian digunakan sebagai acuan untuk menentukan tujuan dan perencanaan strategi dalam penerapan proses pemodelan data mining pada Orange Data Mining.

2.2 Data Understanding

Pada fase pemahaman data ini, data didapat dari Badan Pusat Statistik (BPS) yaitu data produksi daging sapi selama 4 tahun, yaitu pada tahun 2021

sampai dengan tahun 2024 dengan setiap wilayah akan dikelompokkan ke dalam klaster-klaster yang merepresentasikan kesamaan dalam karakteristik produksinya. Data yang diterima dalam bentuk excel dan data yang didapat perlu dibersihkan dan akan dilakukan pemilihan data dengan atributnya.

2.3 Data Preparation

Pada penelitian ini, sumber data yang digunakan berasal dari Badan Pusat Statistik (BPS), yang diakses melalui website resminya (<https://a/www.bps.go.id/id/statistics-table/2/NDgwIzI=/produksi-daging-sapi-menurut-provinsi--ton-.html>). Data yang digunakan merupakan data produksi daging sapi di berbagai wilayah Indonesia selama periode 2021 hingga 2024. Data tersebut kemudian dikelompokkan (*selection*) ke dalam beberapa kelompok yang telah ditentukan, serta diolah untuk keperluan pemodelan data [12].

2.4 Modeling

Tahap modeling diterapkan algoritma yang sesuai untuk mencapai tujuan penelitian. Salah satu teknik yang sering digunakan dalam analisis data adalah K-Means Clustering, yang bertujuan untuk mengelompokkan data produksi daging sapi di berbagai wilayah Indonesia periode 2021 hingga 2024. Pemilihan jumlah cluster menjadi aspek penting yang mempengaruhi hasil akhir.

2.5 Evaluation

Evaluation bertujuan untuk menilai kualitas model yang telah dibangun. Beberapa metode yang digunakan dalam evaluasi clustering adalah Silhouette Score.

2.6 Deployment

Deployment mencakup penyajian hasil penelitian dalam bentuk laporan atau visualisasi yang dapat dilihat dan diperbaharui secara berkala dengan data terbaru untuk pemantauan produksi daging sapi kedepannya.

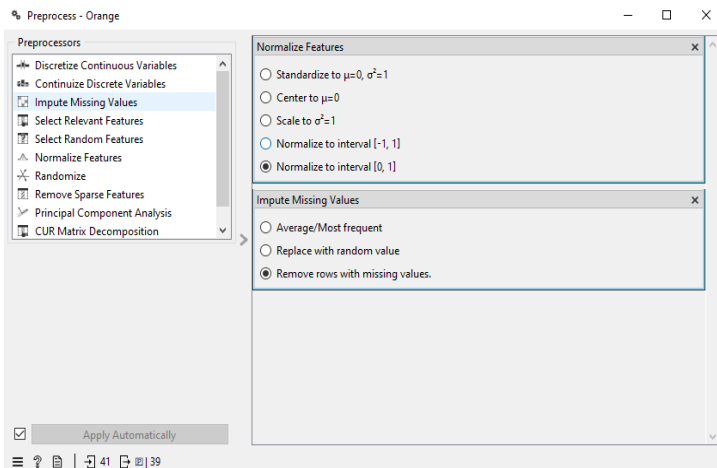
3. Hasil dan Pembahasan

Penelitian ini memanfaatkan perangkat lunak Orange Data Mining sebagai alat bantu untuk mengolah dataset sehingga dapat menghasilkan proses pengelompokan data produksi daging sapi di Indonesia dengan algoritma K-Means Clustering. Penelitian berfokus pada data produksi daging sapi di Indonesia pada tahun 2021-2024.

38 Provinsi	2021	2022	2023	2024
1 ?	Produksi Daging...	?	?	?
2 ?	2021	2022	2023.0	2024.0
3 ACEH	11674.1	12014.96	15644.5	11006.35
4 SUMATERA UT...	13745.06	13859.09	15902.8	18245.02
5 SUMATERA BA...	21375.13	21514.66	31194.4	14901.15
6 RIAU	8948.7	9127.69	10974.8	13457.12
7 JAMBI	4638.46	5760.1	4920.9	3571.72
8 SUMATERA SEL...	13832.62	15458.87	12171.5	11810.74
9 BENGKULU	4348.01	3329.27	3355.5	1762.84
10 LAMPUNG	21130.03	21176.2	22895.2	18624.97
11 KEP. BANGKA B...	3233.04	3395.04	3485.7	2490.59
12 KEP. RIAU	1453.07	1390.39	1820.9	2202.23
13 DKI JAKARTA	16381.81	17617.61	17664.8	14925.17
14 JAWA BARAT	78134.54	84960.62	76287.2	85241.7
15 JAWA TENGAH	65150.77	61393.95	65995.4	83275.69
16 DI YOGYAKARTA	7830.73	7899.9	7998.2	6700.69
17 JAWA TIMUR	108284.07	110991.18	102711.7	96907.31
18 BANTEN	17933.48	17242.69	20754.6	19259.7
19 BALI	4650.67	4694.03	4076.0	4882.25
20 NUSA TENGGA...	11014.08	11159.01	10083.2	11356.76
21 NUSA TENGGA...	8176.8	8680.44	9395.6	6234.53
22 KALIMANTAN ...	5119.52	5095.89	4726.1	3890.08
23 KALIMANTAN ...	3965.53	4113.87	4026.7	1448.19
24 KALIMANTAN ...	7328.53	7351.35	7892.4	5272.95
25 KALIMANTAN ...	7319.28	7465.56	7875.6	6466.78
26 KALIMANTAN ...	808.02	895.57	891.6	632.91
27 SULAWESI UTA...	3103.44	3169.39	2032.2	1840.64
28 SULAWESI TEN...	5926.7	5967.2	5472.3	3848.72
29 SULAWESI SEL...	15366.14	16277.51	16181.3	13722.51
30 SULAWESI TEN...	4522.93	4525.62	5213.8	5985.22
31 GORONTALO	2899.38	2918.75	3137.5	1900.72
32 SULAWESI BAR...	1636.87	1725.34	1358.5	1174.18
33 MALUKU	2306.54	2308.12	2555.7	1279.24
34 MALUKU UTARA	1300.92	1320.51	950.1	1530.33

Gambar 2. Data Table

Gambar 2 adalah gambar data set yang telah dimasukkan kedalam perangkat lunak Orange Data Mining. Namun, data tersebut belum diolah dan masih terdapat missing value. Dari data tersebut kita bisa mengetahui total produksi daging sapi di tiap provinsi yang ada di indonesia mulai dari tahun 2021-2024.



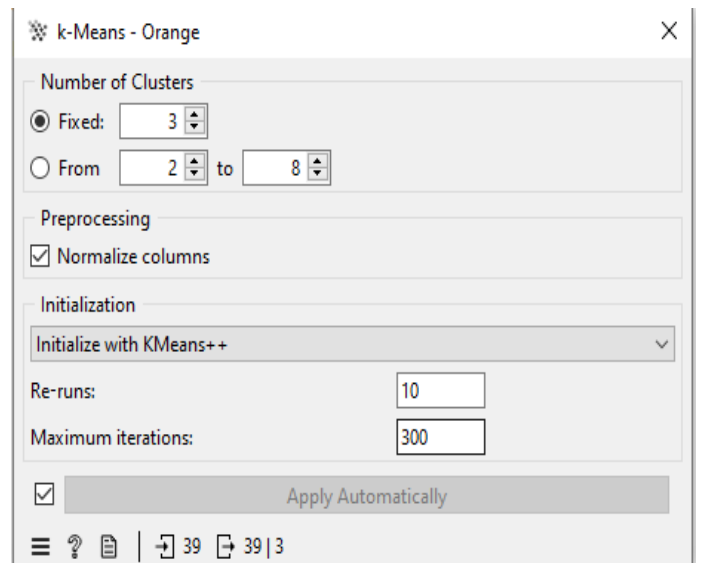
Gambar 3. Tahap Preprocessing

Gambar 3 memperlihatkan tahap preprocessing yang bertujuan untuk membersihkan data agar data siap dianalisis. Pada tahap preprocessing ini dimasukkan fitur normalisasi dan fitur input missing value yang ditujukan untuk melihat data normalisasi dan menghapus missing value pada dataset.

38 Provinsi	2021	2022	2023	2024
1 ACEH	11674.1	12014.96	15644.5	11006.35
2 SUMATERA UT...	13745.06	13859.09	15902.8	18245.02
3 SUMATERA BA...	21375.13	21514.66	31194.4	14901.15
4 RIAU	8948.7	9127.69	10974.8	13457.12
5 JAMBI	4638.46	5760.1	4920.9	3571.72
6 SUMATERA SEL...	13832.62	15458.87	12171.5	11810.74
7 BENGKULU	4348.01	3329.27	3355.5	1762.84
8 LAMPUNG	21130.03	21176.2	22895.2	18624.97
9 KEP. BANGKA B...	3233.04	3395.04	3485.7	2490.59
10 KEP. RIAU	1453.07	1390.39	1820.9	2202.23
11 DKI JAKARTA	16381.81	17617.61	17664.8	14925.17
12 JAWA BARAT	78134.54	84960.62	76287.2	85241.7
13 JAWA TENGAH	65150.77	61393.95	65995.4	83275.69
14 DI YOGYAKARTA	7830.73	7899.9	7998.2	6700.69
15 JAWA TIMUR	108284.07	110991.18	102711.7	96907.31
16 BANTEN	17933.48	17242.69	20754.6	19259.7
17 BALI	4650.67	4694.03	4076.0	4882.25
18 NUSA TENGGA...	11014.08	11159.01	10083.2	11356.76
19 NUSA TENGGA...	8176.8	8680.44	9395.6	6234.53
20 KALIMANTAN ...	5119.52	5095.89	4726.1	3890.08
21 KALIMANTAN ...	3965.53	4113.87	4026.7	1448.19
22 KALIMANTAN ...	7328.53	7351.35	7892.4	5272.95
23 KALIMANTAN ...	7319.28	7465.56	7875.6	6466.78
24 KALIMANTAN ...	808.02	895.57	891.6	632.91
25 SULAWESI UTA...	3103.44	3169.39	2032.2	1840.64
26 SULAWESI TEN...	5926.7	5967.2	5472.3	3848.72
27 SULAWESI SEL...	15366.14	16277.51	16181.3	13722.51
28 SULAWESI TEN...	4522.93	4525.62	5213.8	5985.22
29 GORONTALO	2899.38	2918.75	3137.5	1900.72
30 SULAWESI BAR...	1636.87	1725.34	1358.5	1174.18
31 MALUKU	2306.54	2308.12	2555.7	1279.24
32 MALUKU UTARA	1300.92	1320.51	950.1	1530.33
33 PAPUA BARAT ...	-	-	682.6	325.41
34 PAPUA BARAT ...	-	-	682.6	325.41

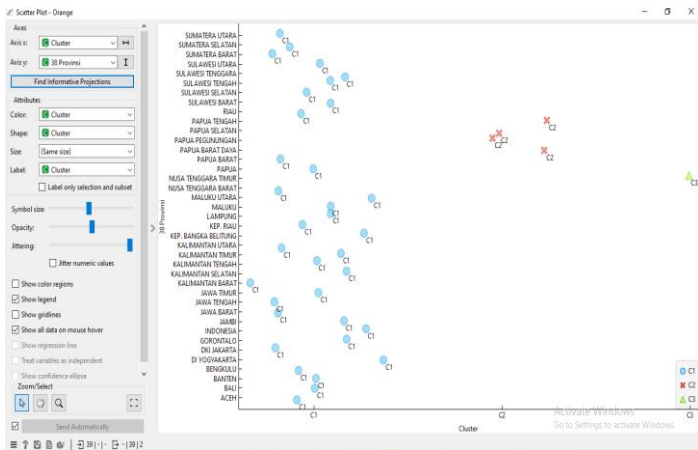
Gambar 4. Data Hasil Normalisasi

Pada tahap normalisasi ini data berhasil dibuat dan sudah menunjukkan hasil dari preprocessing sebelumnya. Gambar diatas menunjukkan bahwa missing value pada dataset produksi daging sapi telah hilang, dan data sudah siap untuk diolah pada tahap berikutnya.



Gambar 5. Penerapan Algoritma K-Means

Gambar 5 merupakan tahapan penerapan algoritma K-Means Clustering. Pada Tahap ini ada beberapa menu yang diterapkan seperti fixed yang berarti akan menerapkan clustering dari c1-c3, normalize columns, dan initialize with K-means.



Gambar 6. Tahap Visualisasi Cluster

Pada tahap ini per modelan yang dipakai adalah scatter plot tujuannya agar terlihat peta sebaran cluster di seluruh provinsi Indonesia mulai dari C1-C3.

Cluster 1: Direpresentasikan dengan lingkaran berwarna biru.

Cluster 2: Direpresentasikan dengan tanda silang berwarna merah.

Cluster 3: Direpresentasikan dengan segitiga berwarna hijau.

Data Hasil Clustering - Orange						
	Cluster	Silhouette	2021	2022	2023	2024
1	C1	0.5	11674.1	12014.96	15644.5	11006.35
2	C1	0.5	13745.06	13859.09	15902.8	18245.02
3	C1	0.5	21375.13	21514.66	31194.4	14901.15
4	C1	0.5	8948.7	9127.69	10974.8	13457.12
5	C1	0.5	4638.46	5760.1	4820.9	3571.72
6	C1	0.5	13832.62	15458.87	12171.5	11810.74
7	C1	0.5	4348.01	3329.27	3355.5	1762.84
8	C1	0.5	21130.03	21176.2	22895.2	18624.97
9	C1	0.5	3233.04	3395.04	3485.7	2490.59
10	C1	0.5	1453.07	1390.39	1820.9	2202.23
11	C1	0.5	16381.81	17617.61	17664.8	14925.17
12	C1	0.5	78134.54	84960.62	76287.2	85241.7
13	C1	0.5	65150.77	61393.95	65995.4	83275.69
14	C1	0.5	7830.73	7899.9	7998.2	6700.69
15	C1	0.5	108284.07	110991.18	102711.7	96907.31
16	C1	0.5	17933.48	17242.69	20754.6	19259.7
17	C1	0.5	4650.67	4694.03	4076.0	4882.25
18	C1	0.5	11014.08	11159.01	10083.2	11356.76
19	C3	0.5	8176.8	8680.44	9395.6	6234.53
20	C1	0.5	5119.52	5095.89	4726.1	3890.08
21	C1	0.5	3965.53	4113.87	4026.7	1448.19
22	C1	0.5	7328.53	7351.35	7892.4	5272.95
23	C1	0.5	7319.28	7465.56	7875.6	6466.78
24	C1	0.5	808.02	895.57	891.6	632.91
25	C1	0.5	3103.44	3169.39	2032.2	1840.64
26	C1	0.5	5926.7	5967.2	5472.3	3848.72
27	C1	0.5	15366.14	16277.51	16181.3	13722.51
28	C1	0.5	4522.93	4525.62	5213.8	5985.22
29	C1	0.5	2899.38	2918.75	3137.5	1900.72
30	C1	0.5	1636.87	1725.34	1358.5	1174.18
31	C1	0.5	2306.54	2308.12	2555.7	1279.24
32	C1	0.5	1300.92	1320.51	950.1	1530.33
33	C1	0.5	1832.35	1506.02	379.6	742.54
34	C2	0.570569	-	-	682.6	325.41
35	C1	0.5	2430.88	2616.77	1150.9	683.22
36	C2	0.570569	-	-	721.6	439.03
37	C2	0.570569	-	-	710.8	667.16
38	C2	0.570569	-	-	214.4	145.81
39	C1	0.5	487802.21	488923.14	503506.8	478852.17

Gambar 7. Data Hasil Clustering

Cluster 1 adalah cluster kelompok provinsi tertinggi yang menghasilkan produksi daging sapi di Indonesia dari tahun ke tahun dengan ciri-ciri angka produksinya mulai dari puluhan ribu sampai dengan ratusan ribu.

Cluster 2 adalah cluster kelompok provinsi yang paling rendah dengan ciri-ciri angka mulai dari 0 sampai ribuan, biasanya kondisi ini terjadi karena faktor geografis yang sulit dengan contoh wilayah provinsi Papua.

Cluster 3 adalah cluster kelompok provinsi dengan produksi daging sapi sedang dengan ciri-ciri angka produksinya hanya sampai belasan ribu.

4. Kesimpulan

Penelitian ini bertujuan untuk mengelompokkan provinsi di Indonesia berdasarkan data produksi daging sapi selama tahun 2021-2024 menggunakan algoritma K-Means Clustering dalam aplikasi Orange Data Mining. Hasil klasterisasi menghasilkan tiga kelompok provinsi berdasarkan kemiripan pola produksi daging sapi tahunan, yaitu cluster dengan produksi daging sapi tinggi dan cenderung meningkat, cluster dengan produksi daging sapi rendah, dan cluster dengan produksi daging sapi sedang. Temuan ini memberikan gambaran mengenai distribusi produksi daging sapi antar provinsi yang dapat dimanfaatkan dalam perencanaan kebijakan peternakan dan ketahanan pangan. Penerapan algoritma K-Means dalam aplikasi Orange Data Mining terbukti efektif dalam mengolah dan menyajikan data strategis yang berguna bagi pemerintah, dan pihak yang berkepentingan dalam sektor peternakan. Penelitian ini juga membuka peluang untuk pengembangan analisis lanjutan pada sektor peternakan lainnya. Penelitian lanjutan disarankan untuk melibatkan variabel tambahan guna menghasilkan analisis yang lebih komprehensif.

Daftar Rujukan

- [1] Kementerian Pertanian Republik Indonesia. (2022). Transformasi Digital Pertanian. Jakarta: Kementerian Pertanian.
- [2] FAO. (2018). World Livestock: Transforming the Livestock Sector through the Sustainable Development Goals. Food and Agriculture Organization of the United Nations.
- [3] BPS (Badan Pusat Statistik). (2023). *Statistik Peternakan dan Kesehatan Hewan 2023*. Jakarta: BPS.
- [4] Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner. *JBASE-Journal of Business and Audit Information Systems*, 4(1).
- [5] Sucipto, A. (2019). Klasterisasi Calon Mahasiswa Baru Menggunakan Algoritma K-Means. *Science Tech: Jurnal Ilmu Pengetahuan dan Teknologi*, 5(2), 50-56.
- [6] Sucipto, A. (2019). Klasterisasi Calon Mahasiswa Baru Menggunakan Algoritma K-Means. *Science Tech: Jurnal Ilmu Pengetahuan dan Teknologi*, 5(2), 50-56.

- [7] Jain. Data clustering: 50 years beyond K-means. 2009
- [8] Handayani, F. (2022). Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokkan Mahasiswa Berdasarkan Gaya Belajar. Jurnal Teknologi dan Informasi, 12(1), 46-63.
- [9] Sucipto, A. (2019). Klasterisasi Calon Mahasiswa Baru Menggunakan Algoritma K-Means. Science Tech: Jurnal Ilmu Pengetahuan dan Teknologi, 5(2), 50-56.
- [10] Qamar, N., & Rezah, F. S. (2020). Metode Penelitian Hukum: Doktrinal dan Non-Doktrinal. CV. Social Politic Genius (SIGn).
-
- [11] Singgalen, Y. A. (2023). Penerapan Metode CRISP-DM dalam Klasifikasi Data Ulasan Pengunjung Destinasi Danau Toba Menggunakan Algoritma Naïve Bayes Classifier (NBC) dan Decision Tree (DT). J. Media Inform. Budidarma, 7(3), 1551-1562.
- [12] Putri, A. E., Yumarlin, M. Z., & Bororing, J. E. (2025). IMPLEMENTASI K-MEANS CLUSTERING DAN MODEL CRISP-DM UNTUK PENGELOMPOKAN DAERAH RAWAN TI.

