

IMPLEMENTASI K-MEANS DAN K-MEDOIDS DALAM PENGELOMPOKAN WILAYAH POTENSIAL PRODUKSI DAGING AYAM

IMPLEMENTATION OF K-MEANS AND K-MEDOIDS IN GROUPING POTENTIAL AREAS OF CHICKEN MEAT PRODUCTION

Miwan Kurniawan Hidayat^{1,2)} dan Rina Fitriana^{2,3)*}

¹⁾Prodi Teknik Industri Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, Senen, Jakarta

²⁾Magister Teknik Industri Universitas Trisakti, Jl. Kyai Tapa No.1, Grogol, Jakarta

³⁾Jurusan Teknik Industri Universitas Trisakti, Jl. Kyai Tapa No.1, Grogol, Jakarta

*Email: rinaf@trisakti.ac.id

Makalah: Diterima 28 Juni 2022; Diperbaiki 19 September 2022; Disetujui 10 Oktober 2022

ABSTRACT

Livestock is the main sector in the effort to fulfill food needs for people in Indonesia and has the potential to maintain the availability of animal food. Guidance and socialization to provide information and knowledge in the field of food production from animals, especially in areas with low levels of chicken meat production need to be done. The research objectives were the use of the K-Means and K-Medoids algorithms for grouping chicken meat production areas in the province of West Java and the use of the Davies Bouldin Index (DBI) value in choosing the best algorithm. The application of K-Means and K-Medoids was carried out through the data mining process phase, namely data collection, data preprocessing, data mining implementation, evaluation of the number of clusters, determination of the best algorithm, and clustering results. The K-Means algorithm with 5 clusters can optimally classify potential areas for chicken meat production in West Java province with a DBI value of 0.273. The results of clustering can be used in business processes related to information on the amount of chicken meat production in the West Java region as a reference in the pattern of guidance to increase animal food production, develop chicken farming potential, and develop animal feed distribution potential.

Keywords: clustering, chicken meat k-means, data mining, k-medoids

ABSTRAK

Peternakan merupakan sektor utama dalam upaya pemenuhan keperluan pangan untuk masyarakat di Indonesia dan memiliki potensi untuk menjaga ketersediaan pangan hewani. Pembinaan dan sosialisasi dalam rangka memberikan informasi dan pengetahuan dalam bidang produksi pangan dari hewan terutama pada wilayah-wilayah dengan tingkat produksi daging ayam yang masih rendah perlu dilakukan. Tujuan penelitian yaitu penggunaan algoritma K-Means dan K-Medoids untuk pengelompokan wilayah produksi daging ayam di wilayah provinsi Jawa Barat serta penggunaan nilai *Davies Bouldin Index* (DBI) dalam memilih algoritma terbaik. Penerapan *K-Means* dan *K-Medoids* dilakukan melalui fase proses *data mining*, yaitu pengumpulan data, *data preprocessing*, implementasi *data mining*, evaluasi jumlah *cluster*, penentuan algoritma terbaik dan hasil *clustering*. Algoritma K-Means dengan 5 *cluster* mampu mengelompokkan wilayah potensial produksi daging ayam di provinsi Jawa Barat secara optimal dengan nilai DBI sebesar 0,273. Hasil *clustering* dapat dimanfaatkan dalam proses bisnis terkait informasi jumlah produksi daging ayam di wilayah Jawa Barat sebagai acuan dalam pola pembinaan untuk meningkatkan produksi pangan hewani, pengembangan potensi budidaya ayam, dan pengembangan potensi distribusi pakan ternak

Kata kunci: data mining, daging ayam klastering, k-means, k-medoids

PENDAHULUAN

Peternakan merupakan sektor utama dalam upaya pemenuhan keperluan pangan untuk masyarakat di Indonesia dan memiliki potensi untuk menjaga ketersediaan pangan hewani. Daging ayam adalah sumber makanan yang mengandung protein hewani dan memiliki tingkat harga yang relatif lebih rendah daripada sumber protein hewani yang lain. Hal ini merupakan salah satu penyebab meningkatnya kebutuhan konsumsi daging ayam. Industri ayam pedaging (*broiler*) adalah bagian dari usaha sektor

peternakan yang memiliki peran utama pada pasokan pangan yang bersumber dari hewan (Nurhayati *et al.*, 2016).

Pemerintah Indonesia selalu mengupayakan peningkatan hasil produksi dari sektor peternakan dalam rangka memenuhi ketersediaan pangan hewani untuk masyarakat, salah satunya adalah produksi daging ayam. Berdasar pada data Badan Pusat Statistik (BPS) yang bersumber dari Direktorat Jenderal Peternakan dan Kesehatan Hewan, Kementerian Pertanian, jumlah daging ayam ras pedaging yang diproduksi di Indonesia pada tahun

2021 mencapai 3.426.042 ton (Badan Pusat Statistik, 2022).

Pembinaan dan sosialisasi produksi pangan hewani dilakukan untuk meningkatkan hasil produksi, memperluas area pemasaran, mencukupi kebutuhan lapangan kerja, meningkatkan penghasilan pelaku usaha daging, serta mengembangkan daya dukung wilayah secara terpadu dengan subsektor lain. Pengelompokan wilayah berdasarkan jenis ayam dan jumlah produksi perlu dilakukan sebagai dasar pengambilan keputusan untuk menetapkan tindakan yang tepat dalam rangka pembinaan pelaku usaha produksi daging ayam. Tujuan penelitian yang dilakukan yaitu penggunaan algoritma *K-Means* dan *K-Medoids* untuk pengelompokan wilayah produksi daging ayam di wilayah provinsi Jawa Barat serta memilih algoritma terbaik berdasarkan nilai *Davies Bouldin Index* (DBI).

Data mining adalah proses menemukan informasi yang bermanfaat pada repositori data yang besar secara otomatis (Tan *et al.*, 2019). Penggunaan data mining dalam menjelajahi banyak data bertujuan untuk memperoleh pola-pola baru yang bermanfaat serta menyediakan kemampuan untuk memprediksi hasil pengamatan di masa depan, seperti jumlah yang akan dibelanjakan pelanggan di toko. *Clustering* dikenal sebagai teknik segmentasi yang membantu membagi kumpulan data besar. Teknik ini mengelompokkan data masa lalu. *Clustering* juga merupakan bagian dari teknik kecerdasan buatan (Maheshwari, 2015). Tujuan *clustering* adalah untuk menentukan pengelompokan intrinsik dalam satu set data yang tidak berlabel tergantung pada beberapa ukuran kesamaan, misalnya jarak Euclidean (Shukur dan Alrashid, 2014). Metode *clustering* yang digunakan dalam penelitian adalah *K-Means* dan *K-Medoids* karena kedua metode tersebut termasuk *partitioning-based clustering* yang membagi dataset besar menjadi angka K (ditentukan oleh pengguna) grup dengan menggunakan jarak sebagai ukuran kesamaan, dimana setiap kelompok mewakili sebuah *cluster* (Djouzi dan Beghdad-Bey, 2019).

K-Means clustering adalah algoritma pengelompokan yang sering digunakan. Diawali dengan menentukan suatu titik K sebagai pusat awal (*centroid*). Selanjutnya semua titik diarahkan ke *centroid* terdekat berdasarkan tingkat kedekatan. Setelah terbentuk *cluster*, dilakukan pembaruan *centroid* pada setiap *cluster*. Selanjutnya algoritma mengulangi dua langkah ini secara iteratif sampai tidak ada perubahan *centroid* (Aggarwal dan Reddy, 2014). *K-Medoids* adalah algoritma *clustering* yang bertujuan untuk meminimalkan kriteria kesalahan absolut daripada *Sum of Squared Errors* (SSE). Mirip dengan algoritma pengelompokan *K-means*, algoritma *K-Medoids* berjalan secara iteratif sampai setiap objek representatif sebenarnya adalah *medoid* dari *cluster* (Aggarwal dan Reddy, 2014).

Penggunaan algoritma *clustering* sudah pernah dilakukan pada penelitian-penelitian

sebelumnya, yaitu penggunaan sistem *clustering* untuk memudahkan proses identifikasi bahan menu makanan halal (Sucipto *et al.*, 2021). Penelitian mengolah data pandemi Covid-19 melalui proses teknik *data mining* menggunakan *K-Means clustering* dengan bantuan tiga perangkat lunak yaitu KNIME, Weka dan Microsoft Excel untuk mengelompokkan wilayah berdasarkan jumlah terjangkau dan meninggal (Indraputra dan Fitriana, 2020); Penelitian Manajemen Pengelolaan Pelanggan Dan Mutu Untuk Agroindustri Susu dengan menerapkan data mining dan OLAP Cube (Fitriana *et al.*, 2012); Penelitian mengenai analisa transaksi dari sebuah toko kopi dengan model association rule untuk mendefinisikan pola transaksi pembelian dalam 6 bulan terakhir. (Febriani *et al.*, 2021); Pengelompokan data wilayah yang terinfeksi Covid-19 dengan pengelompokan terbaik dilakukan menggunakan algoritma *K-Medoids* sebanyak 3 *cluster* menggunakan perangkat lunak Rapid Miner (Sindi *et al.*, 2020); Penerapan algoritma *K-Means* untuk proses pengelompokan provinsi berdasar pada data kependudukan di Indonesia (Ahmar *et al.*, 2018); Penerapan algoritma *K-Means* pada *business intelligence system* dalam mengelompokkan cacat produk pada saat proses produksi roti (Fitriana *et al.*, 2017); Penerapan sistem intelijensia bisnis yang dipadukan dengan metode *K-Means* pada bagian pemasaran di pabrik roti (Fitriana *et al.*, 2018); Segmentasi pelanggan dilakukan menggunakan algoritma *clustering* untuk menganalisis perilaku pelanggan dan mengelompokkannya (Aryuni *et al.*, 2018; Maryani *et al.*, 2018; Shihab *et al.*, 2019; Syukur *et al.*, 2017); Pemanfaatan *clustering* untuk mengelompokkan berbagai jenis pekerjaan yang diminati oleh pencari kerja menggunakan *K-Means clustering* (Shamrat *et al.*, 2020); Perbandingan validasi algoritma *K-Medoids* dengan *K-Means* dengan menggunakan *Silhouette Coefficient Index* dalam mengelompokkan wilayah cacat pada anak (Marlina *et al.*, 2018); Perbandingan antara algoritma *K-Medoids* dan algoritma *K-Means* untuk pengelompokan menu masakan bahan-bahan dari ikan dengan pemilihan algoritma terbaik berdasarkan nilai *Davies Bouldin Index* (Suarna *et al.*, 2021); Analisis komparatif *K-Means* dengan *K-Medoids* dari kedua algoritma dalam kelompok data yang berbeda untuk menjelaskan kekuatan dan kelemahan keduanya (Arbin *et al.*, 2015); Evaluasi kinerja algoritma *K-Means* dasar dilakukan dengan menggunakan berbagai *distance metrics* (Thakare dan Bagal, 2015); Peningkatan hasil *clustering* berdasarkan *Davies Bouldin Index* dalam menentukan *centroid* awal pada algoritma *K-Means* (Sitompul *et al.*, 2019).

METODE PENELITIAN

Penerapan *K-Means* dan *K-Medoids* dilakukan melalui tahapan proses *data mining*, yaitu:

1. Pengumpulan Data
Perhatian utama untuk akurasi informasi adalah sumber data (Jassim dan Abdulwahid, 2021). Data yang digunakan pada penelitian adalah data produksi daging ayam yang diperoleh dari Dinas Ketahanan Pangan dan Peternakan Provinsi Jawa Barat yang dipublikasikan pada situs web <https://opendata.jabarprov.go.id>.
2. *Data Preprocessing*
Teknik *data preprocessing* adalah langkah penting untuk *data mining*, hasil analisis akan baik selama kualitas data baik (Joshi dan Patel, 2020). Tahapan *data preprocessing* dilakukan untuk menghindari masalah yang timbul karena format data yang tidak konsisten sehingga dapat mengganggu ketika pemrosesan data berlangsung dan menyebabkan berkurangnya tingkat akurasi hasil dari proses data. Teknik *data preprocessing* yang digunakan pada penelitian yaitu *data cleaning* untuk membersihkan *missing value*, *data transformation* untuk mengubah bentuk data yang sesuai pada proses data mining, *data integration* untuk menggabungkan beberapa dataset dalam suatu dataset.
3. Implementasi *Data Mining*
Proses *data mining* dilakukan pada tahap ini dengan menerapkan algoritma *K-Means* dan *K-Medoids* menggunakan data hasil *preprocessing*. Memilih algoritma yang sesuai dalam proses *data mining* sangat penting karena menentukan hasil yang akan mempengaruhi keputusan bisnis (Chikohora, 2014). Algoritma yang dipilih ditentukan berdasarkan pada nilai *Davies Bouldin Index* (DBI) yang dihasilkan.
4. Evaluasi Jumlah *Cluster*
Performa algoritma diukur dengan menghitung DBI untuk mengevaluasi *cluster* secara umum berdasarkan jumlah dan kedekatan antar elemen klaster. DBI adalah salah satu teknik pengukuran validitas *clustering* (Aggarwal dan Reddy, 2014). DBI mengukur kekompakan *cluster* yang dihasilkan dan seberapa baik *cluster* ini dipisahkan (Ghany *et al.*, 2022).
5. Penentuan Algoritma Terbaik
Penentuan performa algoritma terbaik berdasarkan pada nilai DBI, performa algoritma dikatakan terbaik jika memiliki DBI lebih kecil. Semakin kecil nilai DBI memberikan hasil

pengelompokan yang lebih baik (Aggarwal dan Reddy, 2014; Sivaguru dan Punniyamoorthy, 2021). Penelitian ini membandingkan kinerja algoritma *clustering K-Means* dan *K-Medoids* sebanyak 6 *cluster*. Nilai k dimasukkan secara bertahap mulai dari 2, 3, 4, 5 dan 6 untuk mendapatkan nilai DBI dari setiap jumlah *cluster* dan hasilnya dibandingkan.

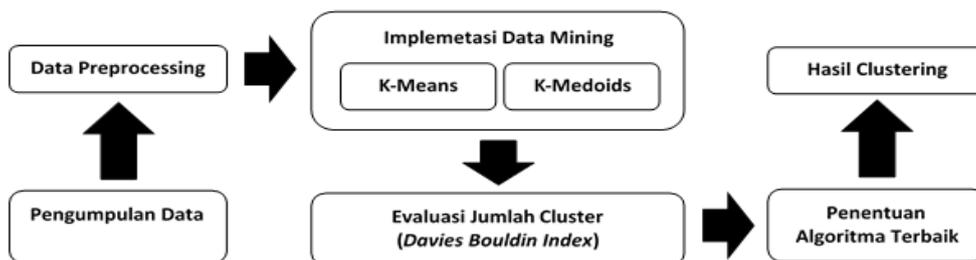
6. Hasil *Clustering*
Pada tahap akhir diperoleh hasil dari proses *data mining* dari penerapan algoritma terbaik yaitu pengelompokan wilayah potensial produksi daging ayam. Alur proses *data mining* dapat dilihat pada Gambar 1.

Hasil pengumpulan data diperoleh data produksi daging ayam dalam format dokumen Microsoft Excel. Pada tahap *data preprocessing* dilakukan pemeriksaan data untuk memastikan data sudah bersih serta dilakukan juga transformasi dan integrasi data menggunakan Microsoft Excel sehingga dihasilkan data yang siap digunakan pada tahap proses *data mining*. Implementasi *data mining* menggunakan perangkat lunak *RapidMiner Studio*. Diawali dengan *import data* dari format Microsoft Excel ke *RapidMiner Studio*, dilanjutkan perancangan proses *data mining* dengan menggunakan beberapa operator untuk pengolahan data. Rancangan proses yang telah disusun dengan baik dapat dijalankan dan menghasilkan output. Hasil proses *data mining* yang diperoleh dapat dianalisis dan disimpulkan.

HASIL DAN PEMBAHASAN

Data untuk keperluan penelitian diperoleh dari Dinas Ketahanan Pangan dan Peternakan Provinsi Jawa Barat yang dipublikasikan melalui situs web <https://opendata.jabarprov.go.id> seperti pada Gambar 2 dan Gambar 3. Informasi tentang data penelitian bisa dilihat di Tabel 1.

Dataset yang diperoleh tentang jumlah produksi daging untuk masing-masing jenis ayam terdiri dari delapan atribut yang menjelaskan jumlah produksi daging ayam per tahun di tiap wilayah kabupaten/kota.



Gambar 4. Proses *Data Mining*

id	kode_provinsi	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	produksi_daging_ayam_buras	satuan
1	32	JAWA BARAT	3201	KABUPATEN BOGOR	1200528	KILOGRAM
2	32	JAWA BARAT	3202	KABUPATEN SUKABUMI	1160064	KILOGRAM
3	32	JAWA BARAT	3203	KABUPATEN CIANJUR	4135276	KILOGRAM
4	32	JAWA BARAT	3204	KABUPATEN BANDUNG	1855093	KILOGRAM
5	32	JAWA BARAT	3205	KABUPATEN GARUT	1713317	KILOGRAM
6	32	JAWA BARAT	3206	KABUPATEN TASIKMALAYA	2002979	KILOGRAM
7	32	JAWA BARAT	3207	KABUPATEN CIAMIS	1497506	KILOGRAM

Gambar 2. Sampel dataset produksi daging ayam buras (Dinas Ketahanan Pangan dan Peternakan, 2022)

id	kode_provinsi	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	produksi_daging_ayam_buras	satuan
1	32	JAWA BARAT	3201	KABUPATEN BOGOR	1200528	KILOGRAM
28	32	JAWA BARAT	3201	KABUPATEN BOGOR	1480892	KILOGRAM
55	32	JAWA BARAT	3201	KABUPATEN BOGOR	1794946	KILOGRAM
82	32	JAWA BARAT	3201	KABUPATEN BOGOR	1886576	KILOGRAM
109	32	JAWA BARAT	3201	KABUPATEN BOGOR	2026862	KILOGRAM

Gambar 3. Sampel dataset produksi daging ayam buras di kabupaten Bogor (Dinas Ketahanan Pangan dan Peternakan, 2022)

Tabel 1. Informasi dataset produksi daging ayam

No.	Dataset	Dibuat	Diperbarui	Dimensi Awal	Dimensi Akhir	Frekuensi
1	Jumlah Produksi Ayam Buras	03 March 2021 23:54:20	24 May 2022 14:54:07	2016	2020	Tahunan
2	Jumlah Produksi Ayam Pedaging	03 March 2021 23:54:22	24 May 2022 14:39:07	2016	2020	Tahunan
3	Jumlah Produksi Ayam Petelur	03 March 2021 23:54:24	24 May 2022 14:52:07	2016	2020	Tahunan

Data preprocessing dilakukan terhadap dataset, kumpulan data diperiksa untuk memastikan data sudah bersih serta melakukan transformasi dan integrasi agar data yang akan digunakan dapat diolah dengan baik pada tahap proses data mining. Langkah selanjutnya setelah data preprocessing yaitu melakukan import data dari format Microsoft Excel ke RapidMiner Studio. Hasil import data pada RapidMiner dapat dilihat pada Gambar 4 dan data jumlah produksi divisualisasikan pada Gambar 5.

Pembuatan rancangan proses data mining dilakukan sebelum menjalankan proses data lebih lanjut seperti pada Gambar 6. Beberapa operator untuk pengolahan data digunakan pada RapidMiner Studio, yaitu Select Attributes, Multiply, K-Means, K-

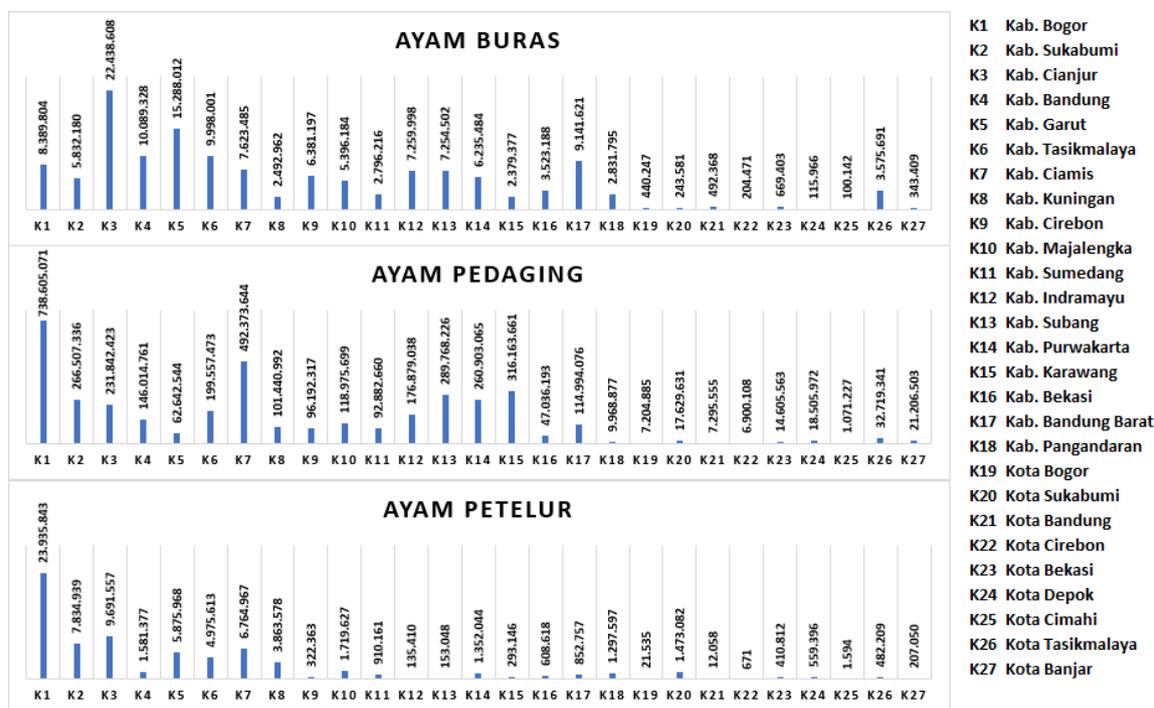
Medoids dan Cluster Distance Performance. Operator Select Attributes digunakan untuk memilih atribut yang akan digunakan pada pengolahan clustering data yaitu buras, pedaging dan petelur. Multiply berfungsi untuk membuat salinan obyek pada RapidMiner. Operator k-Means dan K-Medoids adalah operator yang akan menjalankan proses clustering sesuai algoritma masing-masing. Cluster Distance Performance diperlukan untuk evaluasi centroid pada metode clustering.

Nilai DBI dari algoritma clustering K-Means dan K-Medoids sebanyak 6 cluster dengan menggunakan RapidMiner dibandingkan untuk mengukur kinerja algoritma terbaik. Perbandingan hasil DBI dapat dilihat pada Tabel 2.

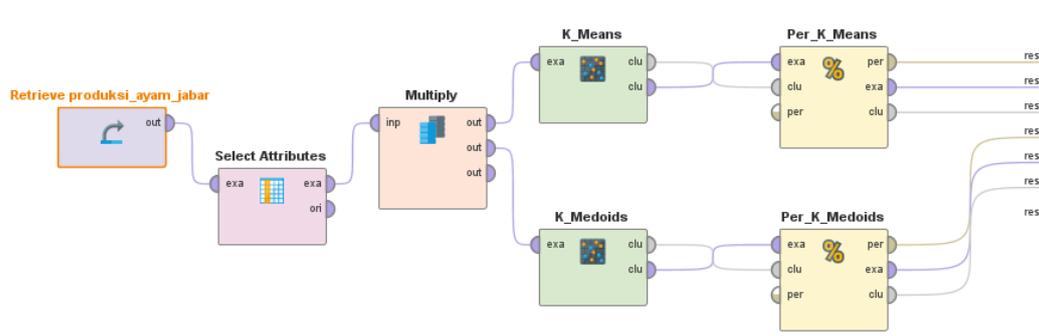
Row No.	nama_kab_kota	buras	pedaging	petelur	Row No.	nama_kab_kota	buras	pedaging	petelur
1	Kab. Bogor	8389804	738605071	23935843	15	Kab. Karawang	2379377	316163661	293146
2	Kab. Sukabumi	5832180	266507336	7834939	16	Kab. Bekasi	3523188	47036193	608618
3	Kab. Cianjur	22438608	231842423	9691557	17	Kab. Bandung Barat	9141621	114994076	852757
4	Kab. Bandung	10089328	146014761	1581377	18	Kab. Pangandaran	2831795	9968877	1297597
5	Kab. Garut	15288012	62642544	5875968	19	Kota Bogor	440247	7204885	21535
6	Kab. Tasikmalaya	9998001	199557473	4975613	20	Kota Sukabumi	243581	17629631	1473082
7	Kab. Ciamis	7623485	492373644	6764967	21	Kota Bandung	492368	7295555	12058
8	Kab. Kuningan	2492962	101440992	3863578	22	Kota Cirebon	204471	6900108	671
9	Kab. Cirebon	6381197	96192317	322363	23	Kota Bekasi	669403	14605563	410812
10	Kab. Majalengka	5396184	118975699	1719627	24	Kota Depok	115966	18505972	559396
11	Kab. Sumedang	2796216	92882660	910161	25	Kota Cimahi	100142	1071227	1594
12	Kab. Indramayu	7259998	176879038	135410	26	Kota Tasikmalaya	3575691	32719341	482209
13	Kab. Subang	7254502	289768226	153048	27	Kota Banjar	343409	21206503	207050
14	Kab. Purwakarta	6235484	260903065	1352044					

ExampleSet (27 examples, 0 special attributes, 4 regular attributes)

Gambar 4. Import data pada RapidMiner



Gambar 5. Visualisasi data jumlah produksi



Gambar 6. Rancangan RapidMiner Studio

Tabel 2. Davies Bouldin Index (DBI)

<i>K</i>	<i>K-Means</i>	<i>K-Medoids</i>
2	0,632	0,540
3	0,454	0,338
4	0,411	1,519
5	0,273	0,598
6	0,311	0,918

Tabel 3. Jumlah elemen cluster

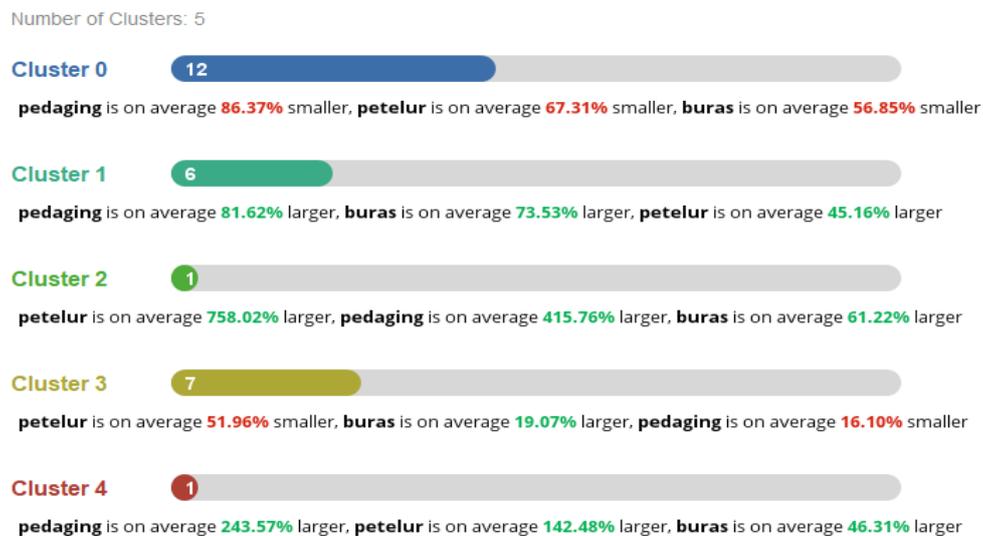
<i>Cluster</i>	<i>Jumlah Elemen</i>
0	12
1	6
2	1
3	7
4	1

Dari perbandingan nilai DBI dapat dilihat bahwa nilai DBI terkecil terdapat pada *cluster* 5 menggunakan algoritma *K-Means* yaitu 0,273, hal ini menunjukkan kinerja algoritma *clustering* terbaik yang dipilih. Jumlah elemen setiap kelompok hasil proses *clustering* dengan 5 *cluster* menggunakan algoritma *K-Means* dapat dilihat pada Tabel 3.

Visualisasi data jumlah elemen dan ikhtisar setiap kelompok hasil proses *clustering* dapat dilihat pada Gambar 7.

Tujuan dari *K-Means Clustering* adalah untuk menemukan sebuah *centroid* untuk setiap *cluster*, semua objek data kemudian diarahkan ke *centroid* terdekat, yang kemudian membentuk sebuah *cluster*. *Centroid* adalah rata-rata dari semua objek data dalam sebuah *cluster* atau objek data yang paling representatif (Umargono *et al.*, 2020). Nilai *centroid* dari setiap *cluster* yang terbentuk hasil proses data dapat dilihat pada Tabel 4.

Rata-rata *cluster distance* dihitung untuk menilai kinerja algoritma *K-Means* yang digunakan. Hasil rata-rata *cluster distance* dapat dilihat pada Gambar 8.



Gambar 7. Grafik Jumlah Elemen *Cluster*

Tabel 4. Nilai Centroid

<i>Cluster</i>	<i>Buras</i>	<i>Pedaging</i>	<i>Petelur</i>
0	2.319.022,75	20.565.533,25	912.549,17
1	9.023.025,33	260.790.364,00	4.050.057,83
2	8.389.804,00	738.605.071,00	23.935.843,00
3	6.222.500,86	121.054.220,40	1.340.753,29
4	7.623.485,00	492.373.644,00	6.764.967,00

```
PerformanceVector:
Avg. within centroid distance: 682499603923448.200
Avg. within centroid distance_cluster_0: 327865706165116.940
Avg. within centroid distance_cluster_1: 1475966683558771.800
Avg. within centroid distance_cluster_2: 0.000
Avg. within centroid distance_cluster_3: 805328675799866.900
Avg. within centroid distance_cluster_4: 0.000
Davies Bouldin: 0.273
```

Gambar 8. Rata-rata *Cluster Distance*

Tabel 5. *Cluster* Wilayah Produksi Daging Ayam

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Kab. Garut	Kab. Sukabumi	Kab. Bogor	Kab. Bandung	Kab. Ciamis
Kab. Bekasi	Kab. Cianjur		Kab. Kuningan	
Kab. Pangandaran	Kab. Tasikmalaya		Kab. Cirebon	
Kota Bogor	Kab. Subang		Kab. Majalengka	
Kota Sukabumi	Kab. Purwakarta		Kab. Sumedang	
Kota Bandung	Kab. Karawang		Kab. Indramayu	
Kota Cirebon			Kab. Bandung Barat	
Kota Bekasi				
Kota Depok				
Kota Cimahi				
Kota Tasikmalaya				
Kota Banjar				

Berdasarkan hasil proses *data mining* menunjukkan bahwa elemen pada kelompok yang terbentuk sudah terpisah pada masing-masing kelompok, dengan demikian tidak terdapat elemen kelompok yang masuk dalam dua kelompok. Suatu *cluster* dapat dinyatakan konvergen jika tidak terjadi perpindahan atau perubahan elemen dari satu *cluster* ke *cluster* yang lain. Hal ini membuktikan bahwa elemen hasil *clustering* bisa mewakili setiap *cluster*. Nilai pusat *cluster* menunjukkan hasil analisis pengelompokan menggunakan algoritma *K-Means* yaitu:

- Cluster 0 berisi produksi ayam buras yang sangat rendah, produksi ayam pedaging yang sangat rendah dan produksi ayam petelur yang sangat rendah.
- Cluster 1 berisi produksi ayam buras yang sangat tinggi, produksi ayam pedaging yang sedang dan produksi ayam petelur yang sedang.
- Cluster 2 berisi produksi ayam buras yang tinggi, produksi ayam pedaging yang sangat tinggi dan produksi ayam petelur yang sangat tinggi.
- Cluster 3 berisi produksi ayam buras yang rendah, produksi ayam pedaging yang rendah dan produksi ayam petelur yang rendah.
- Cluster 4 berisi produksi ayam buras yang sedang, produksi ayam pedaging yang tinggi dan produksi ayam petelur yang tinggi.

Hasil pengelompokan wilayah produksi daging ayam bisa dilihat pada Tabel 5.

Cluster wilayah produksi daging ayam menunjukkan pengelompokan wilayah produksi daging ayam di provinsi Jawa Barat menjadi 5 *cluster*. *Cluster* 0 merupakan *cluster* yang memiliki anggota kelompok paling banyak dibandingkan dengan *cluster* yang lain. Hasil *clustering* dapat dimanfaatkan dalam proses bisnis terkait informasi jumlah produksi daging ayam di wilayah Jawa Barat sebagai acuan dalam pola pembinaan untuk meningkatkan produksi pangan

hewani, pengembangan potensi budidaya ayam, dan pengembangan potensi distribusi pakan ternak.

KESIMPULAN DAN SARAN

Kesimpulan

Algoritma *K-Means* dengan 5 *cluster* mampu mengelompokkan wilayah potensial produksi daging ayam di provinsi Jawa Barat dengan optimal. Pengukuran *Davies Bouldin Index* (DBI) dilakukan untuk memilih algoritma paling baik antara penggunaan *K-Means* dan *K-Medoids*. Nilai DBI yang diperoleh menunjukkan bahwa jumlah *k* optimal adalah 5 *cluster* dengan hasil DBI pada algoritma *K-Means* = 0,273 dan DBI pada algoritma *K-Medoids* = 0,598. Hasil proses *clustering* menunjukkan wilayah produksi daging ayam di provinsi Jawa Barat tersebar di 5 *cluster*. Adapun jumlah wilayah kabupaten/kota pada setiap *cluster* yaitu 12 wilayah termasuk di *cluster* 0; 6 wilayah termasuk di *cluster* 1; 1 wilayah termasuk di *cluster* 2; 7 wilayah termasuk di *cluster* 3 dan 1 wilayah termasuk di *cluster* 4.

Saran

Penelitian tentang penerapan *data mining* produksi daging ayam dapat dilanjutkan dengan menerapkan beragam teknik *clustering* untuk pengelompokan berbagai jenis produksi di berbagai wilayah atau penggunaan beragam teknik *data mining* yang dipadukan sistem *data warehouse* dan intelijensia bisnis.

DAFTAR PUSTAKA

- Aggarwal CC dan Reddy CK. 2014. *Data Clustering Algorithms and Applications*. CRC Press.
- Ahmar AS, Napitupulu D, Rahim R, Hidayat R, Sonatha Y, Azmi M. 2018. Using K-Means Clustering to Cluster Provinces in Indonesia. *Journal of Physics: Conference Series*,

- 1028(1).
- Arbin N, Suhaimi NS, Mokhtar NZ, Othman Z. 2015. Comparative Analysis between K-Means and K-Medoids for Statistical Clustering. *2015 Third International Conference on Artificial Intelligence, Modelling and Simulation*, 117–121. <https://doi.org/10.1109/AIMS.2015.82>
- Aryuni M, Madyatmadja ED, dan Miranda E. 2018. Customer Segmentation in XYZ Bank using K-Means and K-Medoids Clustering. *2018 International Conference on Information Management and Technology (ICIMTech), September*, 1–9. <https://doi.org/10.1109/ICIMTech.2018.8528086>
- Badan Pusat Statistik. 2022. *Produksi Daging Ayam Ras Pedaging menurut Provinsi (Ton), 2019-2021*. <https://www.bps.go.id/indicator/24/488/1/produksi-daging-ayam-ras-pedaging-menurut-provinsi.html>
- Chikohora TT. 2014. A Study of the Factors Considered when Choosing an Appropriate Data Mining Algorithm. *International Journal of Soft Computing and Engineering*, 4(3), 42–45.
- Dinas Ketahanan Pangan dan Peternakan. 2022. *Produksi Daging Ayam Berdasarkan Kabupaten/Kota di Jawa Barat*. <https://opendata.jabarprov.go.id/id/hasil-pencarian?q=produksi+daging+ayam&suggestion=on>
- Djouzi K dan Beghdad-Bey K. 2019. A Review of Clustering Algorithms for Big Data. *Proceedings - ICNAS 2019: 4th International Conference on Networking and Advanced Systems*, 1–6. <https://doi.org/10.1109/ICNAS.2019.8807822>
- Fitriana R, Saragih J, dan Luthfiana N. 2017. Model business intelligence system design of quality products by using data mining in R Bakery Company. *IOP Conference Series: Materials Science and Engineering*, 277(1).
- Fitriana R, Saragih J, dan Hasyati BA. 2018. Perancangan Model Sistem Intelijensia Bisnis Untuk Menganalisis Pemasaran Produk Roti di Pabrik Roti Menggunakan Metode Data Mining dan Cube. *Jurnal Teknologi Industri Pertanian*, 28(1), 113–126.
- Ghany KKA, Aziz AMA, Soliman THA, Sewisy AAEM. 2022. A hybrid modified step Whale Optimization Algorithm with Tabu Search for data clustering. *Journal of King Saud University - Computer and Information Sciences*, 34(3), 832–839. <https://doi.org/10.1016/j.jksuci.2020.01.015>
- Hidayat M K, Fitriana R. 2022. Penerapan Sistem Intelijensia Bisnis Dan K-Means Clustering Untuk Memantau Produksi Tanaman Obat. *Jurnal Teknologi Industri Pertanian*, 32(2), 204–219
- Indraputra RA dan Fitriana R. 2020. K-Means Clustering Data COVID-19. *Jurnal Teknik Industri*, 10(3), 275–282. <https://doi.org/10.25105/jti.v10i3.8428>
- Jassim MA dan Abdulwahid SN. 2021. Data Mining preparation : Process , Techniques and Major Issues in Data Analysis. *IOP Conf. Series: Materials Science and Engineering*, 1090(1). <https://doi.org/10.1088/1757-899X/1090/1/012053>
- Joshi AP dan Patel BV. 2020. Data Preprocessing : the Techniques for Preparing Clean and Quality Data for Data Analytics Process. *Oriental Journal of Computer Science and Technology*, 13(2), 0–3. <https://doi.org/10.13005/ojcs13.0203.03>
- Maheshwari A. 2015. *Data Analytics Made Accessible*.
- Marlina D, Putri NF, Fernando A, Ramadhan A. 2018. Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak. *CoreIT*, 4(2), 64–71. <https://doi.org/10.24014/coreit.v4i2.4498>
- Maryani I, Riana D, Astuti RD, Ishaq A, Pratama EA. 2018. Customer Segmentation based on RFM model and Clustering Techniques With K-Means Algorithm. *2018 Third International Conference on Informatics and Computing (ICIC)*, 1–6. <https://doi.org/10.1109/IAC.2018.8780570>
- Nurhayati, Marimin, Djatna T, Permana, IG. 2016. Kinerja Rantai Pasok Dan Nilai Tambah Dengan Internalisasi Aspek Lingkungan Pada Agroindustri Ayam Ras Pedaging. *Jurnal Teknologi Industri Pertanian*, 26(3), 311–320.
- Shamrat FMJM, Tasnim Z, Mahmud I, Jahan MN, Nobel NI. 2020. Application Of K-Means Clustering Algorithm To Determine The Density Of Demand Of Different Kinds Of Jobs. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, 9(02), 2550–2557.
- Shihab SH, Afroge S, dan Mishu SZ. 2019. RFM Based Market Segmentation Approach Using Advanced K-means and Agglomerative Clustering: A Comparative Study. *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 1–4. <https://doi.org/10.1109/ECACE.2019.8679376>
- Shukur BK dan Alrashid SZN. 2014. Evaluation of Clustering Image Using Steady State Genetic and Hybrid K-Harmonic Clustering Algorithms. *IJCCCE*, 14(1), 10–20.
- Sindi S, Ningse WRO, Sihombing IA, RHZer FI, Hartama D. 2020. Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 166–173.

- <https://doi.org/10.36294/jurti.v4i1.1296>
- Sitompul BJD, Sitompul OS, dan Sihombing P. 2019. Enhancement Clustering Evaluation Result of Davies-Bouldin Index with Determining Initial Centroid of K-Means Algorithm. *IOP Conf. Series: Journal of Physics*, 1235(1), 6–12. <https://doi.org/10.1088/1742-6596/1235/1/012015>
- Sivaguru M dan Punniyamoorthy M. 2021. Performance-enhanced rough k-means clustering algorithm. *Soft Computing*, 25(2), 1595–1616. <https://doi.org/10.1007/s00500-020-05247-2>
- Suarna N, Wijaya YA, Mulyawan, Hartati T, Suprapti T. 2021. Comparison K-Medoids Algorithm and K-Means Algorithm for Clustering Fish Cooking Menu from Fish Dataset. *IOP Conference Series: Materials Science and Engineering*, 1088(1). <https://doi.org/10.1088/1757-899X/1088/1/012034>
- Sucipto S, Al-Mubarak MS, dan Setiyawan DT. 2021. Hierarchical Clustering Bahan Menu di Kantin Universitas untuk Menunjang Implementasi Sistem Jaminan Halal. *Jurnal Teknologi Industri Pertanian*, 31(1), 103–124. <https://doi.org/10.1201/b19706>
- Syakur MA, Khotimah BK, dan Rochman EMS. 2017. Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. *IOP Conf. Series: Materials Science and Engineering*, 336(1). <https://doi.org/10.1088/1757-899X/336/1/012017>
- Tan PN, Steinbach M, Karpatne A, Kumar V. 2019. *Introduction to Data Mining Second Edition*. Pearson Education Limited.
- Thakare YS dan Bagal SB. 2015. Performance Evaluation of K-means Clustering Algorithm with Various Distance Metrics. *International Journal of Computer Applications*, 110(11), 12–16. <https://doi.org/10.5120/19360-0929>
- Umargono E, Suseno JE, dan SK VG. 2020. K-Means Clustering Optimization using the Elbow Method and Early Centroid Determination Based-on Mean and Median. *International Conferences on Information System and Technology, Conrist 2019*, 234–240. <https://doi.org/10.5220/0009908402340240>