ANALISIS POPULASI AYAM RAS PEDAGING DI INDONESIA DENGAN PENERAPAN DATA MINING K-MEDOIDS CLUSTERING

Nurul Ainisa, Nana Suarna, Willy Prihartono
Teknik Informatika, STMIK IKMI Cirebon
Jalan Perjuangan No. 10 B Majasem Kec. Kesambi Kota Cirebon
Nurulainisa5@gmail.com

ABSTRAK

Dalam era digital ini, data menjadi salah satu aset paling berharga bagi berbagai sektor, termasuk peternakan. Populasi ayam ras pedaging di Indonesia memiliki dampak besar pada industri peternakan dan pasokan daging ayam. Namun, masalah yang dihadapi dalam pengelolaan populasi ayam ras pedaging di seluruh provinsi Indonesia adalah kurangnya pemahaman yang mendalam tentang karakteristik populasi ini. Hal ini membuat pengambilan keputusan yang efisien menjadi sulit, terutama dalam perencanaan pengembangan peternakan yang lebih efisien dan berkelanjutan. Tujuan penelitian ini yaitu untuk pengelolaan populasi ayam ras pedaging di seluruh provinsi Indonesia yang efisien dan pemahaman yang mendalam tentang karakteristik populasi ayam ras pedaging, dengan penerapan metode data mining khususnya *K-medoids Clustering*. Metode penelitian yang digunakan melibatkan pengumpulan data populasi ayam ras pedaging di berbagai provinsi di Indonesia yang berasal dari Badan Pusat Statistik Nasional dengan rentang waktu dari tahun 2018 hingga 2022, serta penerapan metode *K-medoids Clustering* untuk mengelompokkan provinsi-provinsi berdasarkan karakteristik populasi ayam ras pedagingnya. Hasil analisis menunjukkan bahwa terdapat 2 *cluster* yaitu *cluster* 0 terdapat 31 provinsi sedangkan *cluster* 1 terdapat 3 provinsi, dengan nilai *Davies Bouldin* sebesar 0.011. Penelitian ini memiliki potensi untuk memberikan pemahaman yang lebih baik tentang populasi ayam ras pedaging di tingkat provinsi, yang dapat membantu dalam perencanaan pengembangan peternakan yang lebih efisien dan berkelanjutan.

Kata kunci : *Data mining, ayam ras pedaging, K-medoids Clustering*

1. PENDAHULUAN

Dalam era perkembangan pesat di bidang Informatika, data mining telah menjadi elemen kunci dalam pengambilan keputusan yang efektif di berbagai sektor, termasuk teknologi, bisnis, dan peternakan. Data mining adalah proses penggalian informasi berharga dari kumpulan data yang besar dan kompleks [1]. Dalam konteks peternakan, data mining dapat memberikan wawasan yang berharga untuk meningkatkan produktivitas dan efisiensi. Salah satu sektor penting dalam peternakan adalah produksi ayam ras pedaging. Indonesia memiliki populasi ayam ras pedaging yang signifikan di setiap provinsinya. Namun, pengelolaan populasi ini memerlukan pemahaman yang baik tentang karakteristik ayam ras pedaging di seluruh wilayah. Penerapan data mining dapat membantu dalam analisis dan pengelompokan populasi ayam ras pedaging berdasarkan wilayah provinsi, yang dapat menjadi dasar untuk pengambilan keputusan yang lebih efisien dalam pengembangan industri peternakan ayam ras pedaging.

Permasalahan yang ditemukan dalam penelitian ini adalah kompleksitas dalam pengelolaan populasi ayam ras pedaging di seluruh provinsi Indonesia. Populasi ayam ras pedaging di Indonesia memiliki dampak besar pada industri peternakan dan pasokan daging ayam. Namun, pengelolaan data populasi ayam ras pedaging yang efisien dan pemahaman yang mendalam tentang distribusi geografis populasi tersebut masih merupakan tantangan. Permasalahan ini menjadi relevan karena informasi ini sangat penting untuk perencanaan dan pengambilan keputusan dalam industri peternakan ayam ras pedaging. Dalam skala yang besar, sulit untuk menganalisis dan memahami pola populasi ini tanpa bantuan alat analisis yang tepat. Selain itu, permasalahan juga melibatkan identifikasi pola distribusi populasi ayam ras pedaging yang dapat memberikan wawasan strategis untuk pengembangan lebih lanjut di sektor peternakan. Relevansi permasalahan ini sangat penting mengingat kontribusi sektor peternakan ayam ras pedaging terhadap pasokan daging hewan yang merupakan salah satu kebutuhan pokok dalam konsumsi masyarakat Indonesia. Oleh karena itu, penelitian ini akan fokus pada penggunaan metode data mining, khususnya metode *K-medoids Clustering*, untuk mengatasi permasalahan tersebut.

Penelitian terdahulu menurut Halimatusakdiah Pohan dkk dengan judul Penerapan Algoritma *K-medoids* dalam Pengelompokan Balita Stunting di Indonesia, menyimpulkan dalam penelitian tersebut metode *K-Medoid* mampu untuk mengelompokkan balita yang mengalami stunting dengan menggunakan 2 *cluster* yaitu tinggi dan rendah. Dengan hasilnya terdapat 28 provinsi dengan *cluster* tertinggi dan 6 provinsi dengan *cluster* terendah [2]. Penelitian yang dilakukan Alia Ahadi Argasah & Dudih Gustian, dengan judul *Data mining analysis to determine employee salaries according to needs based on the k-medoids clustering algorithm*, menyimpulkan metode *k-medoids* dalam data mining sangat berguna sebagai tolak ukur penggajian, akan sangat membantu dalam proses penggajian karyawan oleh perusahan [3]. Menurut Indri Fatma dkk yang berjudul Analisis metode *k-medoids cluster* dalam mengelompokkan

siswa yang berprestasi, menyimpulkan bahwa metode *K-Medoid Clustering* dapat mempermudah untuk mencari siswa berprestasi [4].

Tujuan penelitian ini adalah penerapan teknik data mining, khususnya *K-medoids Clustering*, untuk mengelompokkan populasi ayam ras pedaging di Indonesia berdasarkan wilayah provinsi. Penelitian ini akan memberikan pemahaman yang lebih mendalam tentang karakteristik populasi ayam ras pedaging masing-masing klaster. Signifikansi penelitian ini terletak pada kontribusi pemahaman yang lebih baik tentang populasi ayam ras pedaging di tingkat provinsi, yang dapat membantu dalam perencanaan pengembangan peternakan yang lebih efisien dan berkelanjutan. Selain itu, penelitian ini juga dapat memberikan dasar untuk analisis lebih lanjut dalam sektor pertanian Indonesia. Dari penelitian ini dapat membantu meningkatkan produktivitas dan kesejahteraan peternak, serta menyumbang pada ketahanan pangan nasional.

Metode penelitian ini mulai mengumpulkan data populasi ayam ras pedaging dari berbagai wilayah provinsi di Indonesia yang berasal dari Badan Pusat Statistik Nasional melalui situs https://www.bps.go.id/, dengan rentang data dari tahun 2018 hingga 2022 yang mencakup 34 provinsi. Data yang diperoleh akan diolah menggunakan metode *K-medoids Clustering* sebagai pendekatan utama untuk mengelompokkan data populasi ayam ras pedaging. *K-medoids Clustering* adalah metode data mining yang mampu mengidentifikasi pola dalam data dan mengelompokkannya ke dalam klaster-klaster berdasarkan kemiripan. Pendekatan ini akan digunakan untuk memahami bagaimana populasi ayam ras pedaging di Indonesia dapat dikelompokkan berdasarkan karakteristiknya, khususnya berdasarkan wilayah provinsi. Penelitian ini akan memberikan gambaran tentang bagaimana populasi ayam ras pedaging terdistribusi secara geografis dan karakteristik apa yang mempengaruhi pengelompokkan tersebut.

Hasil dari penelitian ini dapat memberikan panduan praktis bagi pemangku kepentingan di sektor peternakan dalam merencanakan pengelolaan populasi ayam ras pedaging secara lebih efektif. Informasi tentang pola distribusi dan karakteristik populasi dapat digunakan untuk mengoptimalkan alokasi sumber daya, merumuskan strategi pemasaran yang lebih sesuai, dan meningkatkan koordinasi antara berbagai wilayah provinsi. Selain itu, penelitian ini dapat menjadi landasan bagi penelitian lebih lanjut dalam mengembangkan model prediktif untuk pertumbuhan populasi ayam ras pedaging di masa depan.

2. TINJAUAN PUSTAKA

2.1. Ayam Ras Pedaging

Ayam ras pedaging merupakan salah satu jenis hewan yang dikembangkan untuk memenuhi kebutuhan protein hewani, dagingnya empuk, badannya besar, efisiensi nutrisinya tinggi, dan pertambahan berat badannya sangat cepat. Dan mudah untuk dipelihara, memiliki rasa yang enak tetapi mudah rusak jika dimasak dalam waktu lama. Keunggulan ayam ini adalah pertumbuhannya yang cepat yaitu 1-5 minggu. Peranan ayam dalam kehidupan manusia tidak dapat dipungkiri karena selain menghasilkan daging ayam, juga dapat menghasilkan produk samping berupa kotoran ayam yang dapat digunakan sebagai pupuk dan bulu ayam untuk kebutuhan komersial lainnya [5].

2.2. Teori Clustering

Clustering merupakan suatu teknik data mining yang bertujuan untuk mengidentifikasi kelompok objek yang memiliki kesamaan karakteristik tertentu yang dapat dipisahkan dari kelompok objek lainnya. Hal ini membuat objek yang termasuk dalam kelompok yang sama relatif lebih homogen dibandingkan objek di kelompok yang berbeda.Tujuan pengelompokan sekumpulan objek data ke dalam beberapa kelompok yang mempunyai ciri-ciri tertentu dan dapat dibedakan satu sama lain adalah untuk analisis dan interpretasi lebih lanjut tergantung pada tujuan penelitian yang dilakukan. Model yang diambil mengasumsikan bahwa data yang dapat digunakan adalah data berupa data interval, frekuensi dan biner [6].

2.3. Teori K-Medoids

K-medoids adalah algoritma yang digunakan untuk menemukan *medoid* dalam sebuah *cluster*. *K-medoids* lebih kuat daripada *K-means* ketika mencari k sebagai objek representatif, sehingga mengurangi jumlah ketidaksamaan dalam objek data dan mengurangi gangguan dan pencilan. Strategi utama algoritma ini adalah menemukan k *clustering* pada n objek yang dilakukan secara acak. Setiap objek diberi nama *Medoid* yang paling mirip. Algoritma *K-medoids* berfungsi sebagai titik representatif untuk mengambil nilai objek rata-rata pada setiap cluster [7]. *Medoid* adalah data yang dianggap sebagai representasi atau pusat dari suatu *cluster* dan dihitung menggunakan jarak rata-rata atau nilai kesalahan dari seluruh data dalam cluster tersebut. Berbeda dengan *Centroid*, *Medoid* adalah bagian data dalam cluster yang sudah ada. Dalam metode *k-medoid*, jumlah *cluster* yang dihasilkan harus ditentukan terlebih dahulu, dan titik *medoid* diinisialisasi secara acak. Data *cluster* yang ada kemudian diperbarui dengan memperbarui *medoid* yang mewakili setiap *cluster*. Iterasi ini berjalan hingga tidak ada lagi perubahan *medoid* atau batas iterasi tertentu tercapai [8]

2.4. Davies bouldin index

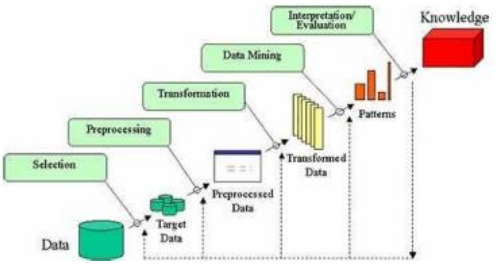
Metode yang diperkenalkan oleh David L. Davies dan Donald W. Bouldin dan metode ini menggunakan nama tersebut dan Davies-Bouldin (DBI) digunakan untuk evaluasi *cluster*. Evaluasi *Davies bouldin index* memiliki skema penilaian klaster internal yang menentukan apakah hasil klaster baik

atau buruk berdasarkan jumlah dan kedekatan antar hasil klaster. *Davies bouldin index* adalah teknik untuk mengukur validitas hasil *cluster* dari suatu metode pengelompokan. Cohesion didefinisikan sebagai jumlah hubungan data antara pusat *cluster* dan *cluster* berikutnya. Pemisahan, sebaliknya, didasarkan pada jarak antara pusat *cluster* dan *cluster*. pengukuran menggunakan *Davies bouldin index*. Hal ini memaksimalkan jarak antara *cluster* Ci dan Cj, sementara mencoba meminimalkan jarak antar titik dalam *cluster*. Jika jarak antar *cluster* kurang dari atau sama dengan berarti fitur antar *cluster* kurang sama, sehingga perbedaan antar *cluster* lebih terlihat. Jarak minimum dalam suatu *cluster* berarti terdapat tingkat kemiripan yang tinggi pada fitur setiap objek dalam *cluster* tersebut [9]

3. METODE PENELITIAN

Metode penelitian yang digunakan yaitu menggunakan pendekatan kuantitatif. Pendekatan kuantitatif adalah pendekatan yang biasanya menekankan pada analisis data atau angka numerik yang digunakan untuk mempelajari populasi dan sampel tertentu sesuai dengan jumlah populasi atau sampel yang diuji.

Data yang dikumpulkan akan dianalisis dengan penerapan data mining menggunakan proses tahapan *Knowledge discovery in databases* (KDD) yang terdiri dari *data, data selection, pre-processing, transformation, data mining, dan Interpretation/Evaluation.*



Gambar 1. Proses *Knowledge discovery in databases* (KDD)[10]

Berdasarkan gambar 1 berikut penjelasan tahapan *Knowledge discovery in databases* (KDD)

3.1. Data

Representasi awal dari dataset yang terdiri dari informasi yang dikumpulkan dari berbagai sumber. Pada tahap ini, data awal yang diperoleh dari berbagai sumber diidentifikasi dan dianggap sebagai potensi sumber pengetahuan yang dapat diolah lebih lanjut.

3.2. Data Selection

Sebelum memasuki langkah penambangan informasi dalam proses *Knowledge discovery in databases* (KDD), data harus dipilih dari sekumpulan data operasional. Data ini disimpan dalam file yang berbeda dari database operasi.

3.3. Pre-processing/Cleaning

Knowledge discovery in databases (KDD) berfokus pada pembersihan data sebelum memulai proses data mining. Selama proses pembersihan, data duplikat harus dihilangkan, data yang tidak konsisten harus diperiksa, dan ketidakakuratan harus diperbaiki.

3.4. Transformation

Sebelum data dapat digunakan, beberapa teknik penambangan data memerlukan format data tertentu. Modifikasi dan seleksi data ini juga mempengaruhi kualitas hasil penambangan data, yang pada gilirannya bergantung pada beberapa pendekatan penambangan data tertentu.

3.5. Data Mining

Data mining adalah teknik yang digunakan untuk menemukan pola atau informasi yang menarik dalam kumpulan data. Tujuan dan penemuan pengetahuan dalam proses *Knowledge discovery in databases* (KDD) menentukan metode, teknik, dan algoritma penambangan data terbaik. Dalam data mining, model *cluster K-medoids* adalah teknik *clustering*; model *cluster* lain adalah *K-means Cluster*. Pusat *cluster* kedua metode *clustering* ini berbeda. *K-medoids* memanfaatkan objek sebagai representasi pusat klaster, atau *medoid*, sementara *K-means* menggunakan nilai rata-rata, atau *mean*, sebagai pusat klaster. Berikut adalah prosedur penyelesaian untuk *K-medoids Clustering* [4]

1. Membentuk k pusat klaster sesuai dengan jumlah klaster yang diinginkan.
2. Mengalokasikan setiap data (objek) ke klaster terdekat dengan menggunakan persamaan berikut untuk mengukur *Euclidean Distance:*

$$d(i,j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2} \quad (1)$$

3. Memilih secara acak objek yang akan berfungsi sebagai *medoid* baru di setiap klaster.
4. Menghitung jarak setiap objek yang berada di masing-masing klaster terhadap *medoid* baru.
5. Menghitung total simpangan (S) dengan mengurangi nilai total jarak baru dengan total jarak sebelumnya. Jika S kurang dari 0, maka mengganti objek dengan data klaster untuk membentuk sekelompok k objek baru sebagai *medoid*.
6. Untuk mendapatkan *cluster* dengan masing-masing anggota, ulangi tahap 3 hingga 5 hingga tidak ada perubahan *medoid*. Nilai k dari data *clustering K-medoids* yang ada di dalam proses *clustering* dapat ditemukan dengan menggunakan nikai DBI terkecil, yang merupakan *Indeks Davies Bouldin*.

3.6. Interpretation /Evaluation

Pada tahap ini, interpretasi dan evaluasi hasil dari proses penambangan data (*Knowledge discovery in* databases/KDD) sangat penting untuk memahami

signifikansi dan implikasi temuan. Proses ini mencakup pemaparan *Davies bouldin index* (DBI) menggunakan *Euclidean distance* sebagai salah satu metrik evaluasi utama. DBI merupakan metrik evaluasi yang digunakan untuk mengukur kualitas klaster yang dihasilkan oleh algoritma *K-medoids*. Nilai DBI yang lebih rendah menunjukkan pembentukan klaster yang lebih baik.

4. HASIL DAN PEMBAHASAN

Hasil penelitian yang dilakukan dalam pembahasan ini yaitu akan menguraikan proses bagaimana pengelompokkan dataset populasi ayam ras pedaging dengan proses pengujian menggunakan *Rapidminer* versi 10.2:

4.1. Data

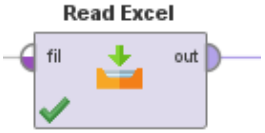
Dataset yang digunakan dalam penelitian ini sebanyak 170 *record* dan 6 atribut yang terdiri dari atribut provinsi, jumlah populasi ayam ras pedaging (ekor) dari 2018, 2019, 2020, 2021, 2022.yang sebanding dengan data dalam tabel 4.1

Tabel 4. 1 Populasi Ayam Ras Pedaging

Provinsi	**Populasi Ayam Ras Pedaging menurut Provinsi (Ekor)**				
	2018	**2019**	**2020**	**2021**	**2022**
Aceh	16821377	33328202	32590982	34075587	42624010
Sumatera Utara	17418 0412	13748 6712	13944 7786	14704 4203	16249 5132
Sumatera Barat	65436 217	57893 566	54364 507	46715 100	36835 762
Riau	83691 805	96875 647	84743 269	81658 751	87783 728
….	….	….	….	….	….
Papua Barat	62449 0	10010 02	89582 2	92664 2	74923 8
Papua	77934 68	64311 56	54653 18	52800 03	32829 17

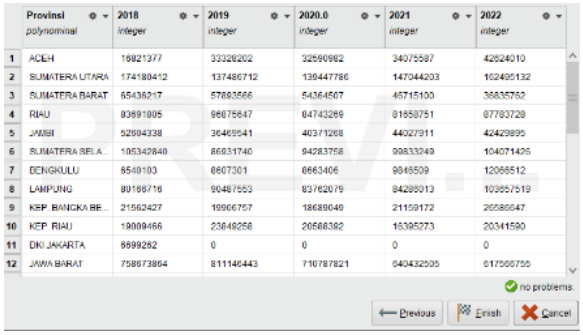
4.2. Data Selection

Tahap data selection merupakan langkah pertama dalam proses *Knowledge Discovery in Database* (KDD). Dari kumpulan data yang diperoleh, dilakukan seleksi data menggunakan file *Microsoft Excel* dalam format xlsx. Data tersebut diimpor ke aplikasi *Rapidminer* menggunakan operator *Read Excel*, sebagaimana ditunjukkan pada Gambar 2.



Gambar 2. Read Excel

Parameter pada operator *Read Excel* menggunakan parameter default. Data yang berhasil diimpor muncul dalam *preview data*, sebagaimana terlihat pada Gambar 3.



	Provinsi polynominal	2018 integer	2019 integer	2020.0 integer	2021 integer	2022 integer
1	ACEH	16821377	33328202	32590982	34075587	42624010
2	SUMATERA UTARA	174180412	137486712	139447786	147044203	162495132
3	SUMATERA BARAT	65436217	57893566	54364507	46715100	36835762
4	RIAU	83691805	96875647	84743269	81658751	87783728
5	JAMBI	52694338	36469541	40371268	44027911	42429895
6	SUMATERA SELA...	105342840	86931740	94283758	99833249	104071426
7	BENGKULU	6540103	6607301	6663406	9946509	12066512
8	LAMPUNG	80166716	90487553	93762079	84289013	103657519
9	KEP. BANGKA BE...	21562427	19966767	18689049	21159172	26686647
10	KEP. RIAU	19009466	23648256	20568382	16395273	20341590
11	DKI JAKARTA	6899262	0	0	0	0
12	JAWA BARAT	758673864	811146443	710787821	640432505	617566755

no problems.

Previous Finish Cancel

Gambar 3. *Review Data*

Langkah berikutnya melibatkan operator *Select Attributes*, yang berfungsi untuk memilih atribut yang akan digunakan dalam analisis, sebagaimana terlihat pada Gambar 4.

Gambar 4. Select Attributes

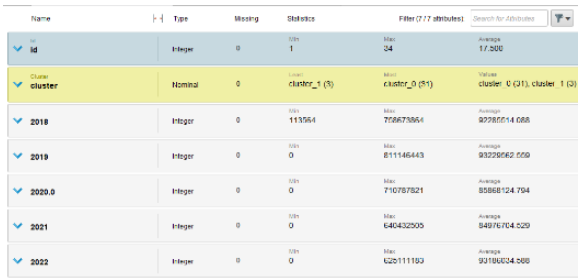
Parameter pada operator *Select Attributes* hanya mengganti tipe filter atribut menjadi *subset*, seperti yang terlihat pada Tabel 4.2

Tabel 4. 2 Parameter Select Attributes

No.	**Parameter**	***Value***
1	*type*	*include attributes*
2	*attribute filter type*	*a subset*
3	*select subset*	*select attributes:*
		2018
		2019
		2020
		2021
		2022
		provinsi

4.3. Pre-Processing/Cleaning

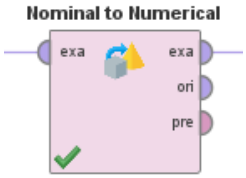
Langkah selanjutnya adalah melakukan pra-pemrosesan pada data yang mungkin mengandung nilai yang hilang (*Missing Value*). Pra-pemrosesan data dilakukan sebelum data dipilih dan digunakan untuk menghilangkan duplikat, mengisi celah, dan memperbaiki kesalahan. Pembersihan data diperlukan untuk mengolah dan melakukan proses data mining. Pada penelitian ini, tidak terdapat *Missing Value* pada data yang telah dipilih, sehingga tahap pra-pemrosesan tidak diperlukan



Gambar 5. *Missing Value*

4.4. Transformation

Tahap transformasi dilakukan menggunakan satu operator dengan tujuan menyiapkan data untuk proses Data Mining. Ini melibatkan pengubahan data yang telah dipilih menjadi bentuk yang sesuai untuk proses Data Mining. Operasi transformasi menggunakan operator *Nominal to Numerical*, seperti yang terlihat pada Gambar 5.



Gambar 6. Nominal to Numerical

Operator *Nominal to Numerical* membantu proses data untuk dapat diproses dalam Data Mining dengan mengubah data nominal menjadi numerik. Tahap ini dilakukan untuk memudahkan proses *Clustering*, terutama pada atribut Provinsi yang merupakan atribut nominal dan tidak dapat diproses dengan operator *k-medoids*.

Tabel 4. 3 Parameter *Nominal to Numerical*

No.	Parameter	*Value*
1.	*attribute filter type*	*subset*
2.	*attributes*	*select attributes:*
		Provinsi

4.5. Data Mining

Data mining merupakan proses mencari pola atau informasi dalam data yang telah dipilih dan diubah selama proses transformasi. Penelitian ini menggunakan metode *k-medoids* untuk mengelompokkan data populasi ayam pedaging, seperti terlihat pada Gambar 7.

Gambar 7. *Clustering K-Medoids*

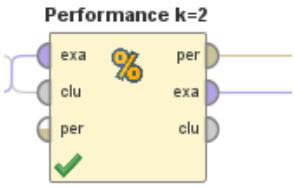
Hasil dari operasi *Clustering K-medoids* menghasilkan 2 *cluster*, yaitu *cluster* 0 dengan 3 item, dan *cluster* 1 dengan 31 item, dengan total 34 item.

Cluster Model

```
Cluster 0: 31 items
Cluster 1: 3 items
Total number of items: 34
```

Gambar 8. *Cluster* Model

Operator *Cluster Distance Performance* ditambahkan dalam proses *Clustering* untuk mengevaluasi kinerja dengan menggunakan *Davies bouldin index* (DBI). *Cluster Distance Performance* digunakan untuk menilai hasil klasterisasi, dengan tujuan mendapatkan nilai DBI terendah.



Gambar 9. *Cluster Distance Performance*

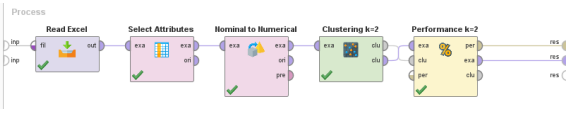
Dari hasil pengklusteran *cluster* 2 maka menghasilkan nilai *davies bouldin indeks* sebesar 0.011.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 824951375399671.800
Avg. within centroid distance_cluster_0: 485079295601827.300
Avg. within centroid distance_cluster_1: 4336962866644064.000
Davies Bouldin: 0.011
```

Gambar 10. *Davies Bouldin Index*

Model proses pada *Rapidminer* dapat dilihat pada Gambar 4.10, menampilkan semua langkah dalam proses.



Gambar 4. 1 Proses *K-medoids* pada *Rapidminer*

Tahapan-tahapan dalam pengelompokkan jumlah populasi ayam ras pedaging di Indonesia, yang digunakan untuk perhitungan algoritma *k-medoids clustering*, dapat dijelaskan sebagai berikut:

Langkah 1:

Memulai dengan menetapkan pusat klaster sejumlah k (jumlah klaster), di mana dalam kasus ini akan menggunakan 2 klaster. Pemilihan *medoid* dilakukan secara acak, dengan DI Yogyakarta dan Jawa Timur diasumsikan sebagai medoid awal.

Tabel 4. 4 *Medoid* Awal

C1	24372273	51245533	51674388	50039073	63702420
C2	44201347 3	45957007 8	38539359 1	39338764 1	49364783 3

Langkah 2:

Menghitung nilai *cost* atau jarak terdekat menggunakan persamaan *euclidean distance*. Oleh karena itu, perhitungan untuk menentukan jarak setiap objek dengan *medoid* awal dapat dijabarkan sebagai berikut:

$$Aceh_{(C1)} = \sqrt{(16821377-24372273)^2+(33328202-51245533)^2+(32590982-51674388)^2+(34075587-50039073)^2+(42624010-63702420)^2} = 37965186{,}93$$

$$Aceh_{(C2)} = \sqrt{(16821377-442013473)^2+(33328202-459570078)^2+(32590982-385393591)^2+(34075587-393387641)^2+(42624010-493647833)^2} = 905244595{,}3{,}3$$

Hasil keseluruhan dapat dilihat dari tabel 4.5.

Tabel 4. 5 Hasil Perhitungan Algoritma *K-medoids* Iterasi Ke-1

Provinsi	Jarak ke *Medoid*	
	C1	C2
ACEH	37965186,9	905244595,3
SUMATERA UTARA	238231367	637423550,7
SUMATERA BARAT	49704594,2	861158419,4
RIAU	90962511,5	782453101,9
JAMBI	40395474,3	880536081,4
SUMATERA SELATAN	117279210	758218529,8
BENGKULU	90926667,6	956009666,9
LAMPUNG	91933882,5	778314085,9
KEP. BANGKA BELITUNG	65495618,1	928038427,9
KEP. RIAU	68976239,3	931670489,7
DKI JAKARTA	110320395	973485160,6
JAWA BARAT	1485398547	637297585,6
JAWA TENGAH	1171873300	318239599,6
DI YOGYAKARTA	0	869058634,2
JAWA TIMUR	869058634	0
BANTEN	353342088	540861325,3
BALI	86419301	801988868,1
NUSA TENGGARA BARAT	47534251	908893043,7
NUSA TENGGARA TIMUR	83104806,5	946877348,7
KALIMANTAN BARAT	36360478,1	861201376,6
KALIMANTAN TENGAH	49941182,6	912110922,1
KALIMANTAN SELATAN	109751114	764000412,2
KALIMANTAN TIMUR	44502180,6	847089055,8
KALIMANTAN UTARA	104086664	969341109,8
SULAWESI UTARA	88781814	952358658,3
SULAWESI TENGAH	95092816,8	959403269
SULAWESI SELATAN	107272609	769948315,1
SULAWESI TENGGARA	97807156,9	962967061,7
GORONTALO	102076713	967187080,2
SULAWESI BARAT	103069819	966959068,5
MALUKU	110542508	975465111,4
MALUKU UTARA	111393959	976289359
PAPUA BARAT	109736352	974639947,3
PAPUA	100361118	964008650,1
Jumlah	6469698557	28238737912
TotalCost	34708436470	

Langkah 3:

Setelah mendapatkan hasil jarak atau *cost* dari setiap objek pada iterasi pertama, langkah selanjutnya adalah melanjutkan ke iterasi kedua. Pada iterasi ini, kandidat *medoid* baru *(non-medoid)* dapat ditemukan dalam Tabel 4.6.

Tabel 4. 6 *Medoid* Baru *(Non-Medoid)* Iterasi Ke-2

c1	6699262	0	0	0	0
c2	758673864	811146443	710787821	640432505	617566755

Langkah 4:

Hitung jarak terdekat dari iterasi ke-2.

$$Aceh_{(C1)} = \sqrt{(16821377-6699262)^2+(33328202-0)^2+(32590982-0)^2+(34075587-0)^2+(42624010-0)^2} = 72479998{,}8$$

$$Aceh_{(C2)} = \sqrt{(16821377-758673864)^2+(33328202-811146443)^2+(32590982-710787821)^2+(34075587-640432505)^2+(42624010-617566755)^2} = 1521027663$$

Maka didapatkan hasil keseluruhan iterasi ke-2 pada tabel 4.7:

Tabel 4. 7 Perhitungan Algoritma *K-medoids* Iterasi Ke-2

Provinsi	Jarak ke *Medoid*	
	C1	C2
ACEH	72479998,8	1521027663
SUMATERA UTARA	338267471	1253954410
SUMATERA BARAT	115309880	1472619687
RIAU	192011195	1396643654
JAMBI	93839733,8	1494857077
SUMATERA SELATAN	216736359	1373801525
BENGKULU	19791918	1570675716
LAMPUNG	196088151	1395092861
KEP. BANGKA BELITUNG	46054680,1	1543097913
KEP. RIAU	42741026,8	1545773559
DKI JAKARTA	0	1587509424
JAWA BARAT	1587509424	0
JAWA TENGAH	1278518002	369390527
DI YOGYAKARTA	110320395	1485398547
JAWA TIMUR	973485161	637297585,6
BANTEN	448019001	1143966103
BALI	173499341	1415127983
NUSA TENGGARA BARAT	65825264,4	1523363262
NUSA TENGGARA TIMUR	28960930,4	1560508357
KALIMANTAN BARAT	112943586	1474853891
KALIMANTAN TENGAH	63470703,5	1528277551
KALIMANTAN SELATAN	214000864	1384616169
KALIMANTAN TIMUR	128219187	1463883223
KALIMANTAN UTARA	9706198,85	1583443572
SULAWESI UTARA	21692442,9	1566749552
SULAWESI TENGAH	15445613,7	1573886704
SULAWESI SELATAN	205364297	1387955671
SULAWESI TENGGARA	13873186,1	1577615159
GORONTALO	10148027,9	1581580689
SULAWESI BARAT	7450163,52	1581274099
MALUKU	6523314,95	1589631847
MALUKU UTARA	6588371,62	1590477319
PAPUA BARAT	6334617,58	1588821179
PAPUA	10539579,3	1577856530
	6831758085	47341029009
TotalCost	54172787095	

Langkah 5: Hitung simpangan total (S) dengan menghitung total *cost* lama dan baru. Apabila nilai S lebih besar dari 0, maka proses pengelompokan akan dihentikan.

S = total cost baru – total cost lama

S = 54172787095 – 34708436470

S = 19464350625

Karena nilai S lebih besar dari 0 maka proses iterasi dihentikan. Sehingga diperoleh anggota tiap *cluster* pada tabel 4.8

Tabel 4. 8 Hasil Pengklasteran dengan *K-medoids* Clustering

Provinsi	Jarak ke *Medoids*		Terdekat	Klaster
	C1	C2		
ACEH	72479998,8	1521027663	72479998,8	1
SUMATERA UTARA	338267471	1253954410	338267471	1
SUMATERA BARAT	115309880	1472619687	115309880	1
RIAU	192011195	1396643654	192011195	1
JAMBI	93839733,8	1494857077	93839733,8	1
SUMATERA SELATAN	216736359	1373801525	216736359	1
BENGKULU	19791918	1570675716	19791918	1
LAMPUNG	196088151	1395092861	196088151	1
KEP. BANGKA BELITUNG	46054680,1	1543097913	46054680,1	1
KEP. RIAU	42741026,8	1545773559	42741026,8	1
DKI JAKARTA	0	1587509424	0	1
JAWA BARAT	1587509424	0	0	2
JAWA TENGAH	1278518002	369390527	369390527	2
DI YOGYAKARTA	110320395	1485398547	110320395	1

Provinsi	Jarak ke *Medoids*		Terdekat	Klaster
	C1	C2		
JAWA TIMUR	973485161	637297585,6	637297586	2
BANTEN	448019001	1143966103	448019001	1
BALI	173499341	1415127983	173499341	1
NUSA TENGGARA BARAT	65825264,4	1523363262	65825264,4	1
NUSA TENGGARA TIMUR	28960930,4	1560508357	28960930,4	1
KALIMANTAN BARAT	112943586	1474853891	112943586	1
KALIMANTAN TENGAH	63470703,5	1528277551	63470703,5	1
KALIMANTAN SELATAN	214000864	1384616169	214000864	1
KALIMANTAN TIMUR	128219187	1463883223	128219187	1
KALIMANTAN UTARA	9706198,85	1583443572	9706198,85	1
SULAWESI UTARA	21692442,9	1566749552	21692442,9	1
SULAWESI TENGAH	15445613,7	1573886704	15445613,7	1
SULAWESI SELATAN	205364297	1387955671	205364297	1
SULAWESI TENGGARA	13873186,1	1577615159	13873186,1	1
GORONTALO	10148027,9	1581580689	10148027,9	1
SULAWESI BARAT	7450163,52	1581274099	7450163,52	1
MALUKU	6523314,95	1589631847	6523314,95	1
MALUKU UTARA	6588371,62	1590477319	6588371,62	1
PAPUA BARAT	6334617,58	1588821179	6334617,58	1
PAPUA	10539579,3	1577856530	10539579,3	1

4.6. Interpretation/ Evaluation

Untuk menentukan jumlah klaster yang menghasilkan nilai *Davies bouldin index* terkecil dilakukan evaluasi klaster. Rekapitulasi *Davies bouldin index* (DBI) diperoleh dengan menggunakan klaster set (K = 2, 3, 4, 5, 6, 7, 8, 9, 10). Untuk mendapatkan nilai K yang optimal, dilakukan 10 kali percobaan. Hasil percobaan dapat tercantum dalam Tabel 4.9

Tabel 4. 9 Rekapitulasi *Davies Bouldin Index*

No	*Cluster*	DBI
1.	K2	0.011
2.	K3	0.013
3.	K4	0.048
4.	K5	0.042
5.	K6	0.031
6.	K7	0.043
7.	K8	0.044
8.	K9	0.019
9.	K10	0.025

Dari 10 kali percobaan, nilai K yang optimal adalah K = 2 dengan DBI sebesar 0.011.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian analisis populasi ayam ras pedaging di Indonesia dengan penerapan data mining *k-medoids clustering*, maka hasilnya dapat disimpulkan pengelompokan populasi ayam ras pedaging menggunakan algoritma *k-medoids* berhasil dilakukan dengan nilai *davies bouldin index* sebesar 0.011. Kelompok yang dihasilkan ada 2 *cluster* yaitu *cluster* 0 terdapat 31 provinsi sedangkan *cluster* 1 terdapat 3 provinsi. *Cluster* 0 merupakan populasi ayam ras pedaging rendah dan untuk *cluster* 1 merupakan populasi ayam ras pedaging tinggi. Saran untuk penelitian selanjutnya dapat menggunakan data per kabupaten/kota pada setiap provinsi agar dapat dilihat lebih detail

DAFTAR PUSTAKA

[1] R. Rinawati, E. G. Sihombing, and ..., "Analisis algoritma datamining pada kasus daerah pelaku kejahatan pencurian berdasarkan provinsi," *J-SAKTI (Jurnal …*, 2020, [Online]. Available: http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/189

[2] H. Pohan, M. Zarlis, E. Irawan, H. Okprana, and Y. Pranayama, "Penerapan Algoritma K-Medoids dalam Pengelompokan Balita Stunting di Indonesia," *JUKI J. Komput. dan Inform.*, vol. 3, no. 2, pp. 97–104, 2021, doi: 10.53842/juki.v3i2.69.

[3] A. A. Argasah and D. Gustian, "Data Mining Analysis To Determine Employee Salaries According To Needs Based on the K-Medoids Clustering Algorithm," *J. Tek. Inform.*, vol. 3, no. 1, pp. 29–36, 2022, [Online]. Available: https://doi.org/10.20884/1.jutif.2022.3.1.154

[4] I. Fatma, H. S. Tambunan, and F. Rizki, "Analisis Metode K-Medoids Cluster Dalam Mengelompokkan Siswa Yang Berprestasi," *Bull. Informatics Data Sci.*, vol. 1, no. 1, pp. 14–19, 2022, [Online]. Available: https://ejurnal.pdsi.or.id/index.php/bids/index

[5] B. Aditama and C. Bella, "Program Pakan Ayam Otomatis Menggunakan Internet of Things," *Portaldata.org*, vol. 1, no. 3, pp. 1–22, 2021.

[6] N. Arief, I. Sudahri Damanik, E. Irawan, S. Tunas Bangsa, and S. Utara, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Algoritma K-Medoids Dalam Mengelompokkan Tingkat Kasus Kejahatan di Setiap Provinsi," *Media Online*, vol. 2, no. 3, pp. 111–116, 2021, [Online]. Available: https://djournals.com/klik

[7] D. Kurmiati, M. Zakiy Fauzi, and A. Falegas, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Clustering of

Earthquake Prone Areas in Indonesia Using K-Medoids Algorithm Klasterisasi Daerah Rawan Gempa Bumi di Indonesia Menggunakan Algoritma K-Medoids," *Journal.Irpi.or.Id*, vol. 1, no. April, pp. 47–57, 2021.

[8] A. Meiriza, E. Ali, Rahmiati, and Agustin, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Program BPJS Ketenagakerjaan," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 714–728, 2023, doi: 10.33022/ijcs.v12i2.3184.

[9] M. Herviany, S. Putri Delima, T. Nurhidayah, and Kasini, "Comparison of K-Means and K-Medoids Algorithms for Grouping Landslide Prone Areas in West Java Province," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 34–40, 2021.

[10] L. Karlina and O. Nurdiawan, "Penerapan K-Medoids Dalam Klasifikasi Persebaran Lahan Kritis Di Jawa Barat Berdasarkan Kabupaten/Kota," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 527–532, 2023, doi: 10.36040/jati.v7i1.6348.