



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 4300-4306

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Penerapan *K-Means Clustering* untuk Klasifikasi Jenis Beras Berdasarkan Tren Harga di Pasar Induk Beras Cipinang (PIBC)

Muhamad Rizal Ubaidillah¹, Nugroho Adhi Santoso², Erni Ungguk Sedyta Utami³
STMIK YMI Tegal

¹muhamadrizalubaidillah@gmail.com*, ²nugrohadhisantoso@stmik-tegal.ac.id, ³eunggulsu@gmail.com

Abstrak

Harga beras yang fluktuatif di Pasar Induk Beras Cipinang (PIBC) memerlukan analisis berbasis data untuk memahami pola pergerakan harga dan mendukung kebijakan distribusi yang tepat. Penelitian ini bertujuan untuk mengelompokkan jenis-jenis beras berdasarkan kesamaan tren harga tahunan selama periode 2020–2024. Metode yang digunakan adalah algoritma *K-Means Clustering* dengan tahapan pra-pemrosesan data menggunakan *Min-Max Scaling*, penentuan jumlah kluster optimal melalui metode *Elbow*, dan evaluasi hasil menggunakan *Silhouette Score*. Hasil penelitian menunjukkan bahwa 14 jenis beras berhasil dikelompokkan ke dalam tiga kluster utama, yaitu kluster harga rendah (seperti IR-64 I, IR-64 II, dan Muncul III), kluster harga menengah (seperti Cianjur Kepala, Setra Ramos, dan Saigon), serta kluster harga tinggi (Ketan Hitam dan Ketan Putih Paris), dengan nilai *Silhouette Score* sebesar 0,507 yang menunjukkan kualitas pengelompokan cukup baik. Implikasi penelitian ini adalah metode data mining dapat digunakan untuk memetakan segmentasi pasar beras secara efektif, sehingga pihak terkait seperti pedagang, distributor, dan pembuat kebijakan dapat lebih mudah merumuskan strategi distribusi dan intervensi harga yang tepat sasaran.

Kata kunci: Data Mining, Elbow Method, *K-Means Clustering*, Pasar Induk Beras Cipinang, *Silhouette Score*.

1. Latar Belakang

Harga beras di Indonesia memegang peran yang sangat penting dalam menjaga stabilitas ekonomi dan ketahanan pangan nasional [1]. Sebagai makanan pokok mayoritas penduduk Indonesia, setiap perubahan harga beras memiliki dampak langsung terhadap daya beli masyarakat dan inflasi nasional. Faktor-faktor seperti musim panen, distribusi, kebijakan pemerintah, kondisi cuaca ekstrem, hingga dinamika pasar global dapat memengaruhi kestabilan harga beras. Oleh karena itu, pemantauan dan analisis harga beras secara berkala menjadi aspek yang sangat krusial bagi berbagai pihak, terutama pemerintah, distributor, dan pelaku usaha di sektor pangan [2].

Pasar Induk Beras Cipinang (PIBC) merupakan pusat distribusi beras terbesar di Indonesia yang mencerminkan kondisi pasokan dan permintaan beras nasional. Pergerakan harga di pasar ini sering kali menunjukkan pola fluktuasi serentak pada berbagai jenis beras, meskipun berasal dari varietas berbeda [3]. Kondisi ini menandakan adanya keterkaitan pola tren harga antarjenis beras. Namun, ketiadaan analisis berbasis data yang dapat mengelompokkan jenis beras berdasarkan tren pergerakan harganya berpotensi menimbulkan masalah, seperti distribusi yang tidak tepat sasaran, risiko kelangkaan, serta kenaikan harga yang membebani konsumen [4].

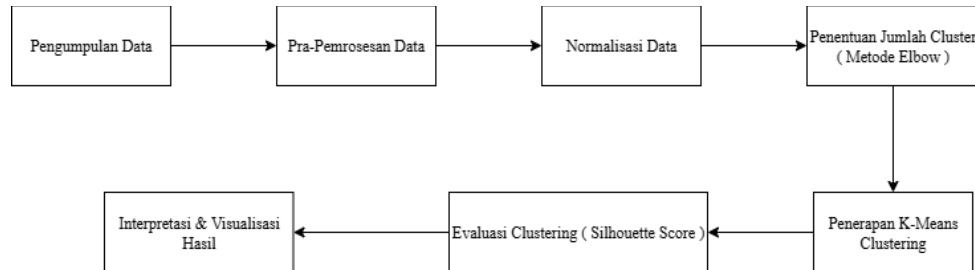
Penelitian sebelumnya telah menunjukkan efektivitas metode data mining dalam berbagai kasus pengelompokan. Penerapan algoritma *K-Means* pada stok barang di sebuah toko menghasilkan tiga kelompok: barang terlaris, sedang, dan kurang laris [5]. Algoritma yang sama juga digunakan pada data penjualan komoditas pertanian untuk membagi barang ke dalam tiga *cluster* berdasarkan tingkat penjualan [6]. Studi lain memanfaatkan *K-Means* untuk memetakan pasar tradisional berdasarkan harga komoditas pangan dan membentuk *cluster* harga tinggi, sedang, dan rendah [7]. Penerapan *K-Means* juga dilakukan pada evaluasi kinerja dosen, menghasilkan empat kelompok penilaian dengan kualitas *Clustering* yang cukup baik [8]. Selain itu, *K-Means* digunakan dalam analisis hasil belajar peserta didik, yang menghasilkan tiga *cluster* capaian belajar dengan kategori rendah, sedang, dan tinggi [9].

Berdasarkan hal tersebut, penelitian ini dilakukan untuk mengidentifikasi pola tren harga tahunan berbagai jenis beras di PIBC periode 2020–2024, menerapkan algoritma *K-Means Clustering* untuk mengelompokkan jenis beras berdasarkan kesamaan pergerakan harga, serta menentukan jumlah kluster optimal dengan pendekatan *Elbow Method* dan evaluasi *Silhouette Score*. Hasil penelitian diharapkan dapat memberikan segmentasi harga beras yang

representatif, sehingga dapat dijadikan dasar dalam sistem pendukung keputusan untuk menjaga stabilitas harga, kelancaran distribusi, serta keberlanjutan pasokan beras di Indonesia.

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif terapan yang berfokus pada pemanfaatan teknik unsupervised learning, khususnya algoritma *K-Means Clustering*, untuk mengelompokkan jenis-jenis beras berdasarkan kemiripan tren harga dari tahun ke tahun. Pendekatan ini dipilih karena sesuai dengan karakteristik data yang tidak memiliki label kelas, namun memiliki pola tersembunyi yang dapat diungkap melalui pengelompokan [10].



Gambar 1. Proses Analisis

Gambar 1. Menunjukkan proses analisis data dalam penelitian ini mencakup beberapa tahapan utama, dimulai dari pengumpulan data harga beras tahunan periode 2020–2024 dari Pasar Induk Beras Cipinang (PIBC) sebagai sumber sekunder. Selanjutnya dilakukan pra-pemrosesan berupa pembersihan nilai kosong, pengecekan konsistensi antar tahun, serta pengklasifikasian per jenis beras dalam format deret waktu. Data kemudian dinormalisasi menggunakan metode *Min-Max Scaling* agar seluruh fitur berada pada skala seragam. Jumlah *cluster* optimal ditentukan dengan metode *Elbow* melalui pemetaan nilai *Sum of Squared Error* (SSE), sebelum algoritma *K-Means* diterapkan secara iteratif untuk menentukan centroid dan mengelompokkan data berdasarkan jarak Euclidean terdekat. Evaluasi hasil *Clustering* dilakukan menggunakan *Silhouette Score* untuk menilai kualitas pemisahan antar *cluster*, dan tahap akhir berupa interpretasi serta visualisasi hasil yang menampilkan karakteristik masing-masing *cluster* dengan grafik atau plot untuk mempermudah pemahaman

2.1. Pengumpulan Data

Data harga beras dikumpulkan dari sumber resmi yaitu Pasar Induk Beras Cipinang (PIBC) yang mencakup harga tahunan berbagai jenis beras dari tahun 2020 hingga 2024. Data ini bersifat sekunder dan diambil dalam bentuk tabel untuk memudahkan pengolahan.

Tabel 1. Data Harga Jenis Beras Dari Tahun 2020-2024

Jenis Beras	2020	2021	2022	2023	2024*
Cianjur Kepala (Pandan Wangi Premium)	13547	13330	13474	15485	18432
Cianjur Slyp (Pandan Wangi)	12634	12302	12668	14769	17346
Setra	13293	13188	13234	14166	15766
Saigon	11916	11751	11759	13090	14966
Muncul I	11338	11076	10730	12138	14865
Muncul II	10422	10022	9928	11469	14130
Muncul III	9758	9422	9343	11003	13537
IR-64 I	10684	10486	10489	12529	14585
IR-64 II	9980	9467	9810	11687	13650
IR-64 III	8700	8523	9002	10740	11479
IR-42 (Pera)	12266	11076	10767	14644	14788
Ketan Putih (Biasa)	12503	11418	11927	12702	18112
Ketan Putih (Paris)	18541	16605	16149	16667	25031
Ketan Hitam	19571	17991	17034	19825	20734

2.2. Pra Pemrosesan Data

Tahapan pra-pemrosesan untuk memastikan bahwa data yang dianalisis bersih, konsisten, dan sesuai dengan format yang dibutuhkan oleh algoritma. Pada tahap ini, dilakukan pembersihan data dengan cara menyaring jenis beras yang memiliki data lengkap untuk periode tahun 2020 hingga 2024.

2.3. Normalisasi Data

Data harga kemudian dinormalisasi menggunakan metode *Min-Max Scaling* agar seluruh fitur memiliki skala yang seragam. Normalisasi sangat penting untuk mencegah distorsi hasil *Clustering* akibat perbedaan skala antar fitur. Rumus *Min-Max Scaling* dituliskan pada persamaan (1)

$$x^1 = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

x merepresentasikan nilai asli dari suatu data, sedangkan x^1 adalah nilai data setelah dinormalisasi. Nilai x_{min} merupakan nilai terkecil dari seluruh data dalam fitur yang sama, dan x_{max} adalah nilai terbesarnya [11].

2.4. Penentuan Jumlah Cluster (*Elbow Method*)

Sebelum menjalankan *K-Means*, jumlah kluster optimal (k) ditentukan menggunakan metode *Elbow*, dengan cara memplot nilai *Sum of Squared Error* (SSE) terhadap jumlah kluster, lalu memilih titik “tekukan” sebagai jumlah *cluster* terbaik. Titik ini disebut sebagai titik *Elbow* dan digunakan sebagai dasar dalam menentukan jumlah kluster yang ideal untuk menghasilkan pengelompokan yang efisien dan bermakna [12]. Rumus *Sum of Squared Error* (SSE) dituliskan di persamaan (2)

$$SSE = \sum_{i=1}^k \sum_{x \in c_i} \|x - \mu_i\|^2 \quad (2)$$

k adalah jumlah kluster, c_i menyatakan kluster ke- i , x adalah titik data dalam kluster tersebut μ_i adalah *centroid* kluster ke- i [13].

2.5. Penerapan *K-Means Clustering*

Algoritma *K-Means* diterapkan pada data yang telah dinormalisasi. Proses ini dilakukan secara iteratif untuk menentukan pusat *cluster* (*centroid*) dan mengelompokkan data berdasarkan jarak *Euclidean* terdekat ke *centroid*. Fungsi tersebut dituliskan pada persamaan (3)

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (3)$$

k adalah jumlah *cluster*, C_i adalah *cluster* ke- i , x adalah titik data dalam *cluster* C_i , dan μ_i adalah *centroid* atau rata-rata dari *cluster* tersebut. Notasi $\|x - \mu_i\|^2$ menyatakan jarak *Euclidean* kuadrat antara data x dan *centroid* μ_i .

2.6. Evaluasi Hasil *Clustering* (*Silhouette Score*)

mengukur seberapa baik hasil pengelompokan, digunakan metrik *Silhouette Score*, yang mengevaluasi tingkat kemiripan data dalam satu *cluster* dibandingkan dengan *cluster* lain. Nilai yang mendekati 1 menunjukkan hasil *Clustering* yang optimal. *Silhouette Score* sangat berguna untuk menilai kualitas pengelompokan secara objektif tanpa perlu mengetahui label data yang sebenarnya [14]. Rumus *Silhouette Score* ada di persamaan (4)

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (4)$$

$a(i)$ adalah selisih antara rata-rata jarak data terhadap seluruh titik dalam *cluster* yang sama, $b(i)$ adalah rata-rata jarak terkecil terhadap titik-titik dalam *cluster* terdekat yang bukan *cluster*-nya [15].

2.7. Interpretasi dan Visualisasi Hasil

Hasil *Clustering* dianalisis untuk melihat karakteristik masing-masing kluster, seperti jenis beras yang cenderung memiliki harga stabil atau *fluktuatif*. Visualisasi menggunakan grafik atau *plot* kluster membantu mempermudah interpretasi hasil.

3. Hasil dan Diskusi

Hasil penelitian yang diperoleh melalui penerapan algoritma *K-Means Clustering* pada data harga beras di PIBC periode 2020–2024.

3.1. Pra-Pemrosesan Data dan Normalisasi

Pada tahap ini, dilakukan pembersihan data dengan cara menyaring jenis beras yang memiliki data lengkap untuk periode tahun 2020 hingga 2024. Data harga beras yang digunakan merupakan data harga rata-rata tahunan untuk masing-masing jenis beras.

Tabel 2. Data Sebelum Normalisasi

Jenis Beras	2020	2021	2022	2023	2024*
Cianjur Kepala (Pandan Wangi Premium)	13547	13330	13474	15485	18432
Cianjur Slyp (Pandan Wangi)	12634	12302	12668	14769	17346
Setra	13293	13188	13234	14166	15766
Saigon	11916	11751	11759	13090	14966
Muncul I	11338	11076	10730	12138	14865
Muncul II	10422	10022	9928	11469	14130
Muncul III	9758	9422	9343	11003	13537
IR-64 I	10684	10486	10489	12529	14585
IR-64 II	9980	9467	9810	11687	13650
IR-64 III	8700	8523	9002	10740	11479
IR-42 (Pera)	12266	11076	10767	14644	14788
Ketan Putih (Biasa)	12503	11418	11927	12702	18112
Ketan Putih (Paris)	18541	16605	16149	16667	25031
Ketan Hitam	19571	17991	17034	19825	20734

Setelah data di kumpulkan maka data di normalisasikan terlebih dahulu menggunakan *min-max scaling* berikut perhitungan manualnya sebagai sampel :

$$x^1 = \frac{x - x_{min}}{x_{max} - x_{min}}$$

$$x^1 = \frac{13547 - 8700}{19571 - 8700}$$

$$x^1 = \frac{4847}{10.871}$$

$$x^1 = 0.445865146$$

```

    DATA SETELAH NORMALISASI:
    Jenis Beras    2020    2021    2022    2023    2024
    Cianjur Kepala 0.445865 0.507710 0.556773 0.522289 0.513061
    Cianjur Slyp   0.361880 0.399134 0.456424 0.443478 0.432925
    Setra          0.422500 0.492712 0.526892 0.377105 0.316337
    Saigon         0.295833 0.340938 0.343252 0.258668 0.257305
    Muncul I       0.242664 0.269645 0.215139 0.153880 0.249852
    Muncul II      0.158403 0.158323 0.115289 0.080242 0.195617
    Muncul III     0.097323 0.094951 0.042455 0.028949 0.151860
    IR-64 I        0.182504 0.207330 0.185134 0.196918 0.229191
    IR-64 II       0.117744 0.099704 0.100598 0.104238 0.160198
    IR-64 III      0.000000 0.000000 0.000000 0.000000 0.000000
    IR-42 (Pera)   0.328029 0.269645 0.219746 0.429719 0.244171
    Ketan Putih (Biasa) 0.349830 0.305767 0.364168 0.215960 0.489448
    Ketan Putih (Paris) 0.905253 0.853612 0.889816 0.652394 1.000000
    Ketan Hitam    1.000000 1.000000 1.000000 1.000000 0.682925
    
```

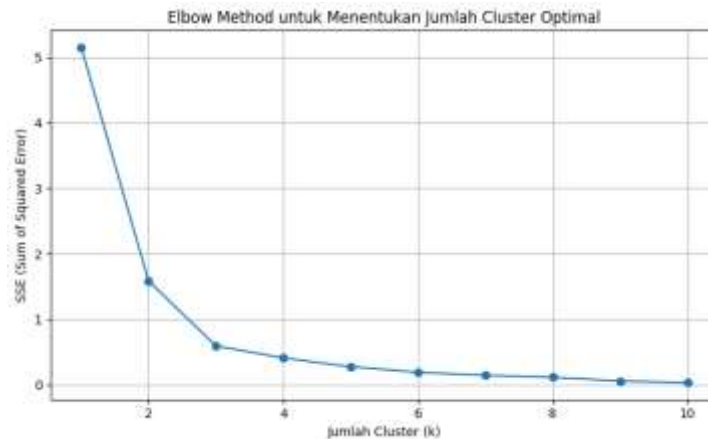
Gambar 2. Data Setelah Normalisasi

Gambar 2. menunjukkan data harga beras yang telah melalui proses normalisasi menggunakan metode *Min-Max Scaling*. Proses ini mengubah skala data asli ke dalam rentang antara 0 hingga 1, sehingga seluruh fitur (tahun) memiliki bobot yang seimbang dalam proses analisis. Nilai 1 menunjukkan harga tertinggi pada tahun tertentu, sedangkan nilai mendekati 0 menunjukkan harga terendah. Misalnya, Ketan Hitam memiliki nilai 1 di sebagian

besar tahun karena harganya selalu tertinggi dibanding jenis beras lain, sedangkan IR-64 III memiliki nilai yang rendah karena berada pada kisaran harga paling murah. Proses normalisasi ini penting untuk memastikan bahwa algoritma *K-Means* tidak bias terhadap fitur dengan nilai besar, dan menghasilkan pengelompokan yang lebih objektif.

3.2. Penentuan Jumlah *Cluster*

Penentuan jumlah *cluster* optimal dilakukan dengan metode *Elbow* melalui perhitungan nilai SSE pada berbagai k , di mana titik siku pada grafik menunjukkan jumlah *cluster* terbaik untuk data harga beras.



Gambar 3. Penentuan Jumlah *Cluster*

Gambar 3. merupakan hasil visualisasi metode *Elbow* yang digunakan untuk menentukan jumlah *cluster* optimal dalam proses *K-Means Clustering*. Grafik ini memetakan nilai Sum of Squared Error (SSE) terhadap jumlah *cluster* (k) dari 1 hingga 10. Terlihat bahwa penurunan nilai SSE paling tajam terjadi antara $k = 1$ hingga $k = 3$, kemudian setelah $k = 3$, penurunan nilai SSE mulai melandai dan tidak signifikan. Titik tekuk (*Elbow*) yang jelas terlihat pada $k = 3$, yang mengindikasikan bahwa jumlah *cluster* optimal untuk data ini adalah tiga. Dengan demikian, pemilihan $k = 3$ sebagai jumlah kluster merupakan pilihan yang tepat karena memberikan keseimbangan antara akurasi pengelompokan dan kompleksitas model.

3.3. Menentukan titik *Centroid*

Setelah jumlah *cluster* optimal diperoleh, langkah selanjutnya adalah menentukan titik centroid awal secara acak, kemudian memperbaruinya secara iteratif hingga posisi centroid konvergen dan menghasilkan pembagian data yang stabil.

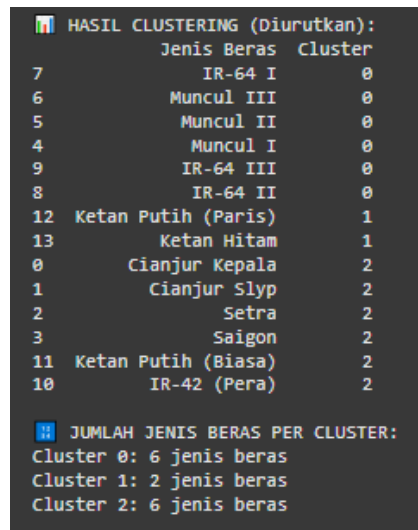
TITIK CENTROID PER CLUSTER (Skala Normalisasi):					
	2020	2021	2022	2023	2024
Cluster 0	0.133106	0.138326	0.109769	0.094038	0.164453
Cluster 1	0.952626	0.926806	0.944908	0.826197	0.841463
Cluster 2	0.367323	0.385984	0.411209	0.374537	0.375541

Gambar 4. Menentukan titik *Centroid*

Gambar 4. menunjukkan nilai titik pusat (centroid) dari masing-masing *cluster* yang dihasilkan melalui algoritma *K-Means* berdasarkan data harga beras yang telah dinormalisasi. Nilai centroid merepresentasikan rata-rata posisi dari anggota-anggota dalam setiap *cluster* terhadap masing-masing fitur (harga tahunan), dengan skala antara 0 hingga 1. Berdasarkan hasil yang ditampilkan, *Cluster 1* memiliki nilai centroid tertinggi secara konsisten pada seluruh tahun, yaitu 0.9526 pada tahun 2020 hingga 0.8415 pada tahun 2024. Hal ini menunjukkan bahwa *cluster* tersebut terdiri dari jenis-jenis beras dengan harga tertinggi. Sebaliknya, *Cluster 0* menunjukkan nilai centroid paling rendah, dimulai dari 0.1331 pada tahun 2020 hingga 0.1645 pada tahun 2024, yang mengindikasikan bahwa *cluster* ini mencerminkan kelompok beras dengan harga paling rendah. Adapun *Cluster 2* memiliki nilai centroid yang berada di antara keduanya, menggambarkan kelompok beras dengan harga menengah. Nilai-nilai centroid ini digunakan sebagai acuan utama dalam pengelompokan dan menjadi dasar dalam mengidentifikasi karakteristik harga masing-masing *cluster*.

3.4. Hasil *Clustering*

Hasil *Clustering* menunjukkan pengelompokan jenis beras ke dalam beberapa *cluster* sesuai kesamaan pola tren harga tahunan, sehingga tiap *cluster* merepresentasikan karakteristik pergerakan harga yang berbeda.



```
HASIL CLUSTERING (diurutkan):
Jenis Beras Cluster
7 IR-64 I 0
6 Muncul III 0
5 Muncul II 0
4 Muncul I 0
9 IR-64 III 0
8 IR-64 II 0
12 Ketan Putih (Paris) 1
13 Ketan Hitam 1
0 Cianjur Kepala 2
1 Cianjur Slyp 2
2 Setra 2
3 Saigon 2
11 Ketan Putih (Biasa) 2
10 IR-42 (Pera) 2

JUMLAH JENIS BERAS PER CLUSTER:
Cluster 0: 6 jenis beras
Cluster 1: 2 jenis beras
Cluster 2: 6 jenis beras
```

Gambar 5. Hasil *Clustering*

Gambar 5. menunjukkan hasil pengelompokan jenis beras menggunakan algoritma *K-Means* dengan jumlah *cluster* sebanyak tiga. Setiap jenis beras diklasifikasikan ke dalam salah satu *cluster* berdasarkan kemiripan pola harga rata-rata tahunan dari tahun 2020 hingga 2024. Berdasarkan hasil tersebut, *Cluster 0* terdiri dari 6 jenis beras dengan kecenderungan harga yang lebih rendah dan stabil, seperti IR-64 I, IR-64 II, dan Muncul III. *Cluster 1* berisi 2 jenis beras, yaitu Ketan Hitam dan Ketan Putih (Paris), yang secara konsisten memiliki harga paling tinggi sepanjang periode pengamatan. Sementara itu, *Cluster 2* mencakup 6 jenis beras dengan karakteristik harga menengah dan tren yang relatif meningkat, di antaranya Cianjur Kepala, Setra, dan Ketan Putih (Biasa). Hasil pengelompokan ini menggambarkan struktur harga beras di Pasar Induk Beras Cipinang, serta memberikan gambaran segmentasi pasar berdasarkan tren harga dalam lima tahun terakhir. Setelah proses *Clustering* dilakukan, proses selanjutnya adalah mengevaluasi sejauh mana efektivitas pengelompokan, digunakan metrik *Silhouette Score*.

3.4. Evaluasi Metrik

Evaluasi hasil *Clustering* dilakukan menggunakan *Silhouette Score* untuk menilai kualitas pemisahan antar *cluster*, di mana nilai yang semakin mendekati 1 menunjukkan pengelompokan yang semakin baik.



Gambar 6. Hasil Evaluasi Metrik

Gambar 6. menunjukkan hasil evaluasi pengelompokan menggunakan metrik *Silhouette Score*, yang digunakan untuk menilai kualitas hasil *Clustering*. Nilai *Silhouette Score* berada pada rentang -1 hingga 1, di mana nilai mendekati 1 menunjukkan bahwa objek-objek dalam satu *cluster* sangat mirip satu sama lain dan berbeda jelas dari *cluster* lainnya. Berdasarkan hasil perhitungan, diperoleh nilai *Silhouette Score* sebesar 0.507, yang menunjukkan bahwa hasil pengelompokan menggunakan algoritma *K-Means* berada pada kategori cukup baik. Artinya, sebagian besar jenis beras telah berhasil dikelompokkan ke dalam *cluster* yang sesuai dengan pola kemiripan harga, meskipun masih terdapat kemungkinan perbaikan pada struktur *cluster* agar hasilnya lebih optimal.

Berdasarkan hasil penelitian, penerapan algoritma *K-Means Clustering* berhasil mengelompokkan 14 jenis beras di Pasar Induk Beras Cipinang (PIBC) menjadi tiga *cluster* utama berdasarkan pola tren harga rata-rata tahunan periode 2020–2024. *Cluster* pertama berisi enam jenis beras dengan harga rendah dan stabil, seperti IR-64 I, IR-

64 II, dan Muncul III; *cluster* kedua terdiri dari enam jenis beras dengan harga menengah dan tren meningkat, seperti Cianjur Kepala, Setra Ramos, dan Saigon; sedangkan *cluster* ketiga mencakup dua jenis beras premium, yaitu Ketan Hitam dan Ketan Putih (Paris), yang memiliki harga tertinggi. Pemilihan jumlah *cluster* optimal ditentukan melalui metode *Elbow* dengan $k = 3$, serta dievaluasi menggunakan *Silhouette Score* sebesar 0,507 yang menunjukkan kualitas *Clustering* cukup baik. Hasil ini menegaskan bahwa metode *K-Means* efektif untuk memetakan segmentasi pasar beras di PIBC, memberikan gambaran jelas tentang struktur harga, dan dapat mendukung pengambilan keputusan terkait distribusi maupun kebijakan harga

4. Kesimpulan

Penelitian ini membuktikan bahwa algoritma *K-Means Clustering* dapat secara efektif mengelompokkan 14 jenis beras di Pasar Induk Beras Cipinang (PIBC) menjadi tiga segmen harga rendah, menengah, dan tinggi dengan jumlah *cluster* optimal $k = 3$ yang ditentukan melalui metode *Elbow* dan kualitas *Clustering* yang cukup baik ditunjukkan oleh *Silhouette Score* sebesar 0,507. Hasil pengelompokan tersebut mencerminkan struktur pasar beras yang nyata, mempermudah analisis pola harga, serta memberikan dasar bagi pengambilan keputusan distribusi dan kebijakan harga yang lebih tepat. Untuk pengembangan ke depan, penelitian disarankan menggunakan data harga dengan interval waktu lebih rinci, mengeksplorasi metode *Clustering* lain seperti *Hierarchical Clustering* atau DBSCAN, serta menambahkan variabel pendukung seperti volume distribusi atau asal produksi beras agar pemetaan pasar menjadi lebih akurat dan mendalam.

Referensi

- [1] Fakultas and R. Juliansyah, "Pengaruh Pengadaan Beras dan Operasi Pasar Oleh Perum Bulog Terhadap Stabilitas Harga Beras di Provinsi Aceh," *Jurnal Ilmiah Ekonomi Terpadu (Jimetera)*, vol. 4, no. 1, pp. 38–44, 2024, doi: 10.35308/jimetera.v4i1.8825.
- [2] F. A. Adit Kurniawan, I. Ramadhani, and R. A. Fitri, "Analisis Pengambilan Keputusan Kebijakan Impor Beras dalam Konteks Pemenuhan Pangan Nasional di Indonesia," *Indonesian Journal of Public Administration Review*, vol. 1, no. 2, p. 15, May 2024, doi: 10.47134/par.v1i2.2394.
- [3] Serly. F. Nenoharan, D. R. Nendisa, and I. W. Nampa, "ANALISIS FLUKTUASI HARGA BERAS KUALITAS MEDIUM DAN PREMIUM DI PASAR TRADISIONAL KOTA KUPANG DAN MAUMERE. Studi Kasus: Pasar Inpres Naikoten 1 Kupang dan Pasar Alok Maumere," *Buletin Ilmiah IMPAS*, vol. 22, no. 3, Nov. 2021, doi: 10.35508/impas.v22i3.5695.
- [4] D. Setiadi and Efan, "Analisis Prediksi Harga Beras Berbasis Kualitas Menggunakan Algoritma K-Nearest Neighbor," *Jurnal Ilmiah Informatika Global*, vol. 15, no. 3, Dec. 2024, doi: 10.36982/jiig.v15i3.4810.
- [5] F. Matheos Sarimole and Hakim Lukmanul, "Klasifikasi Barang Menggunakan Metode Clustering K-Means Dalam Penentuan Prediksi Stok Barang," *Jurnal Sains dan Teknologi*, vol. 5, no. 3, p. 2024, doi: 10.55338/saintek.v5i1.2709.
- [6] D. Abid, R. Wirya Adikusuma, A. Mufti, A. Fikri, D. Rinci, and K. Hapsari, "Penerapan Metode K-Means Clustering Untuk Analisa Penjualan Komoditas Toko Tani Indonesia," *Jurnal Riset Inovasi Bidang Informatika Dan Pendidikan Informatika (KERNEL)*, vol. 3, no. 2, 2022, doi: 10.31284/j.kernel.2022.v3i2.4076.
- [7] A. Novita and H. B. Seta, *Pemetaan Pasar Tradisional Berdasarkan Harga Pangan Komoditas Menggunakan Algoritma K-Means*. 2021.
- [8] N. Luh, P. P. Dewi, I. Nyoman Purnama, and N. W. Utami, "Penerapan Data Mining Untuk Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: STMIK Primakara)," *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 16, no. 2, 2022, doi: 10.32815/jitika.v16i2.761.
- [9] M. Rafi Muttaqin, T. Iman Hermanto, and M. Agus Sunandar, "PENERAPAN K-MEANS CLUSTERING DAN CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING (CRISP-DM) UNTUK MENGELOMPOKAN PENJUALAN KUE," vol. 19, no. 1, pp. 38–53, 2022, [Online]. Available: <https://journal.unpak.ac.id/index.php/komputasi>
- [10] A. Maulana Artiarno, P. Setiaji, and F. Nugraha, "K-Means Clustering untuk Segmentasi Pelanggan: Mengungkap Pola Pembelian Strategi Pemasaran pada Sektor Ritel," *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 2, 2025, doi: 10.29408/edumatic.v9i2.30336.
- [11] P. Palinggik Allorerung, A. Erna, M. Bagussahrir, and S. Alam, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," *Jurnal Informatika Sunan Kalijaga*, vol. 9, no. 3, pp. 178–191, Sep. 2024, doi: 10.14421/jiska.2024.9.3.178-191.
- [12] N. A. Maori and E. Evanita, "Metode elbow dalam optimasi jumlah cluster pada k-means clustering," *Jurnal Simetris*, vol. 14, no. 2, pp. 1–11, Nov. 2023.
- [13] F. P. Azizah, S. S. Hilabi, T. Tukino, and A. Hananto, "Perbandingan Algoritma K-Means dan Hierarchical Untuk Klasterisasi Data Kehadiran Karyawan," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 14, no. 1, 2025, doi: 10.35889/jutisi.v14i1.2644.
- [14] N. Hendrastuty, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.
- [15] T. Rahmawati, Y. Wilandari, and P. Kartikasari, "ANALISIS PERBANDINGAN SILHOUETTE COEFFICIENT DAN METODE ELBOW PADA PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN INDIKATOR IPM DENGAN K-MEDOIDS," *Jurnal Gaussian*, vol. 13, no. 1, pp. 13–24, Aug. 2024, doi: 10.14710/j.gauss.13.1.13-24.