

# Penerapan *K-Means* pada Produksi Padi dengan Evaluasi *DBI*

## *K-Means Application to Rice Production with DBI Evaluation*

<sup>1</sup>Badruttamam\*, <sup>2</sup>Mukti Qamal, <sup>3</sup>Nunsina

<sup>1,2,3</sup>Universitas Malikussaleh

<sup>1,2,3</sup>Jl. Batam, Bukit Indah - Lhokseumawe, Aceh, Indonesia

\*e-mail: [badruttamam.200170207@mhs.unimal.ac.id](mailto:badruttamam.200170207@mhs.unimal.ac.id), [mukti@unimal.ac.id](mailto:mukti@unimal.ac.id), [nunsina@unimal.ac.id](mailto:nunsina@unimal.ac.id)

(received: 31 May 2026, revised: 27 June 2026, accepted: 3 July 2026)

### Abstrak

Produksi padi merupakan salah satu indikator penting dalam mendukung ketahanan pangan daerah sehingga diperlukan analisis untuk mengetahui pola dan perbedaan tingkat produksi antar kecamatan. Penelitian ini bertujuan menerapkan algoritma *K-Means* dalam mengelompokkan data produksi padi di Kabupaten Bireuen serta mengevaluasi kualitas kluster menggunakan metode Davies-Bouldin Index (DBI). Data yang digunakan merupakan data sekunder periode 2020–2024 yang mencakup 17 kecamatan dengan variabel luas tanam, luas panen, produktivitas, dan total produksi. Metode *K-Means* digunakan untuk membagi data ke dalam tiga kluster, yaitu produksi tinggi, sedang, dan rendah berdasarkan kemiripan karakteristik data. Hasil penelitian menunjukkan bahwa proses klustering mampu mengelompokkan kecamatan secara konsisten pada setiap tahun pengamatan. Evaluasi menggunakan DBI menghasilkan nilai sebesar 0,9147 pada tahun 2020, 0,8292 pada tahun 2021, 0,6119 pada tahun 2022, 0,8821 pada tahun 2023, dan 0,8597 pada tahun 2024. Nilai terbaik diperoleh pada tahun 2022 karena memiliki nilai DBI paling rendah yang menunjukkan kualitas kluster yang lebih optimal. Penelitian ini diharapkan dapat menjadi dasar pertimbangan dalam penyusunan strategi peningkatan produksi dan pengambilan kebijakan pertanian di Kabupaten Bireuen.

**Kata kunci:** *davies-bouldin index, data mining, K-Means, clustering, produksi padi.*

### Abstract

Rice production is a key indicator of regional food security, making it essential to analyze production patterns and identify differences in productivity across districts. This study aims to apply the *K-Means* clustering algorithm to classify rice production data in Bireuen Regency and evaluate the quality of the resulting clusters using the Davies–Bouldin Index (DBI). The study employed secondary data from 2020 to 2024 covering 17 districts, with four variables: cultivated area, harvested area, productivity, and total rice production. The *K-Means* algorithm was used to partition the data into three clusters representing high, medium, and low production levels based on similarities in their characteristics. The results indicate that the clustering process consistently classified the districts across the five-year observation period. Cluster quality evaluation using the Davies–Bouldin Index yielded values of 0.9147 in 2020, 0.8292 in 2021, 0.6119 in 2022, 0.8821 in 2023, and 0.8597 in 2024. The best clustering performance was achieved in 2022, as indicated by the lowest DBI value, reflecting a more compact and well-separated clustering structure. These findings provide valuable insights for supporting agricultural planning, improving rice production strategies, and informing policy decisions related to agricultural development in Bireuen Regency.

**Keywords:** *Davies-Bouldin Index, data mining, K-Means, clustering, rice production.*

## 1 Pendahuluan

Di era digital saat ini, teknologi informasi telah menjadi pendorong utama dalam berbagai sektor kehidupan, termasuk sektor pertanian. Inovasi teknologi memungkinkan pengelolaan data dalam skala besar yang lebih efisien dan efektif, serta mendukung pengambilan keputusan berbasis data. Salah

<http://sistemasi.ftik.unisi.ac.id>

satu penerapan teknologi tersebut adalah dalam bidang analisis data menggunakan pendekatan data mining [1].

Sektor pertanian memiliki peran penting dalam perekonomian Indonesia, terutama dalam penyediaan pangan dan kesejahteraan petani. Padi sebagai komoditas utama menjadi fokus penting dalam upaya peningkatan produksi guna memenuhi kebutuhan lokal dan menjaga ketahanan pangan nasional. Kabupaten Bireuen, sebagai salah satu daerah agraris di Provinsi Aceh, memiliki kontribusi signifikan dalam produksi padi. Namun, variasi produksi padi di berbagai kecamatan menunjukkan adanya perbedaan pola yang dapat dipengaruhi oleh berbagai faktor, seperti kondisi tanah, pola tanam, curah hujan, dan penggunaan teknologi pertanian [2].

Permasalahan variasi hasil produksi padi di Kabupaten Bireuen masih menjadi tantangan bagi petani dan pemerintah daerah. Setiap kecamatan sering kali memiliki hasil panen yang berbeda-beda, sehingga menimbulkan kesenjangan antar wilayah. Tanpa adanya data dan pemetaan yang jelas, upaya untuk mengembangkan sektor pertanian bisa jadi tidak tepat sasaran. Karena itu, penting dilakukan pengelompokan data produksi padi secara ilmiah, agar pemerintah dan pihak terkait dapat mengambil keputusan yang lebih akurat [3].

Algoritma K-Means merupakan salah satu teknik clustering yang sering digunakan karena kemampuannya dalam mengelompokkan data dalam jumlah besar secara efektif. Namun, algoritma ini memiliki kelemahan dalam menentukan kluster yang optimal dan sangat rentan terhadap titik awal (centroid) yang dipilih secara acak [4]. Untuk mengevaluasi kualitas hasil clustering, digunakan Davies-Bouldin Index (DBI) yang merupakan metode evaluasi internal yang mengukur tingkat pemisahan antar kluster dan kekompakan dalam kluster. Semakin kecil nilai DBI, semakin baik kualitas clustering yang dihasilkan [5].

Penelitian ini bertujuan untuk menerapkan algoritma K-Means dalam mengelompokkan data produksi padi di 17 kecamatan Kabupaten Bireuen berdasarkan variabel luas tanam, luas panen, produktivitas, dan total produksi selama periode 2020–2024, serta mengevaluasi kualitas pengelompokan menggunakan DBI. Hasil penelitian diharapkan dapat memberikan informasi yang bermanfaat bagi pihak terkait dalam merumuskan kebijakan dan perencanaan pertanian di Kabupaten Bireuen.

## **2 Tinjauan Literatur**

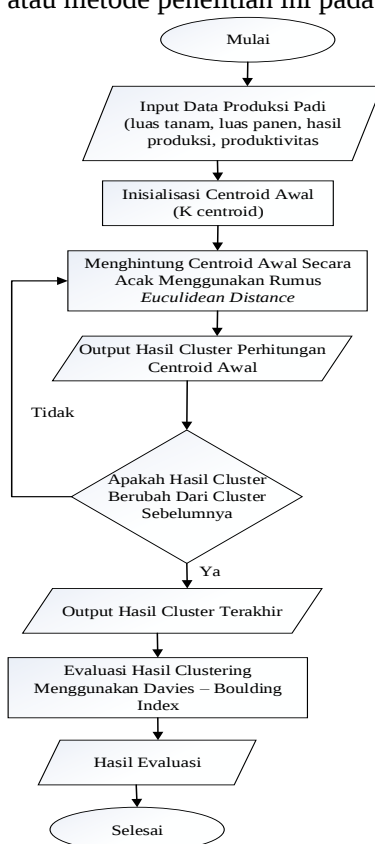
Penelitian tentang penerapan algoritma K-Means dalam bidang pertanian telah banyak dilakukan. Wijayanto dan Fathoni [1] menerapkan K-Means untuk mengelompokkan produktivitas tanaman padi di Jawa Tengah dan berhasil mengidentifikasi tiga kelompok wilayah produksi. Wahyudin dkk. [6] melakukan pengelompokan produktivitas padi di Jawa Barat menggunakan K-Means dan menghasilkan empat kluster dengan karakteristik produksi yang berbeda. Al Masykur dkk. [7] menerapkan K-Means untuk pemetaan lahan produksi jagung di Kabupaten Pasuruan dan menunjukkan bahwa algoritma ini mampu mengelompokkan wilayah berdasarkan tingkat produktivitas. Putra dan Qamal [8] membandingkan K-Means dan Fuzzy C-Means pada data produksi padi di Aceh Utara serta menunjukkan bahwa K-Means menghasilkan pengelompokan yang lebih optimal. Pramesti dan Fajar [9] menerapkan K-Means untuk prediksi hasil panen padi, dan Zulkarnain dan Bustami [10] melakukan pemetaan daerah produksi padi di Aceh Timur menggunakan algoritma yang sama.

Evaluasi kualitas kluster merupakan aspek penting dalam klustering untuk memastikan validitas hasil pengelompokan. Davies-Bouldin Index (DBI) merupakan salah satu metode evaluasi internal yang banyak digunakan karena tidak memerlukan data label eksternal. Prasetyo dan Fahmi [11] membandingkan K-Means dan K-Medoids clustering pada data produksi tanaman pangan serta menunjukkan bahwa K-Means menghasilkan pengelompokan yang lebih baik. Suryani dan Wahyuni [12] mengevaluasi penerapan Davies-Bouldin Index pada clustering data tanaman pangan di Indonesia dan menyimpulkan bahwa DBI efektif mengukur kualitas kluster. Rahmawati dan Hartati [13] mengoptimasi centroid pada clustering data pertanian dan menunjukkan bahwa pemilihan centroid awal berpengaruh terhadap kualitas kluster. Sinaga dan Yang [14] mengembangkan algoritma Unsupervised K-Means yang mampu menentukan jumlah kluster secara otomatis.

Meskipun berbagai penelitian telah menerapkan K-Means dan evaluasi DBI di berbagai domain, masih terdapat celah penelitian dalam analisis produksi padi di Kabupaten Bireuen menggunakan pendekatan clustering. Penelitian sebelumnya cenderung berfokus pada wilayah lain seperti Jawa Tengah [15]. Selain itu, evaluasi menggunakan DBI pada data produksi padi di tingkat kecamatan masih jarang dilakukan secara longitudinal. Penelitian di Aceh seperti Putra dan Qamal [8] di Aceh Utara serta Zulkarnain dan Bustami [10] di Aceh Timur telah menerapkan K-Means pada data padi, namun belum mengevaluasi hasil clustering menggunakan DBI secara longitudinal. Penelitian ini memberikan kontribusi dalam beberapa aspek. Pertama, penelitian ini menyajikan analisis clustering produksi padi di Kabupaten Bireuen menggunakan K-Means dengan evaluasi DBI. Kedua, penelitian ini menggunakan data lima tahun (2020-2024) sehingga mampu mengidentifikasi konsistensi pola clustering antar tahun. Ketiga, hasil penelitian ini dapat menjadi acuan bagi pemerintah daerah dalam merumuskan kebijakan produksi padi yang lebih tepat sasaran.

### 3 Metode Penelitian

Berikut ini merupakan tahapan atau metode penelitian ini pada Gambar 1.



Gambar 1 Metode penelitian

#### 3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder produksi padi di Kabupaten Bireuen periode 2020-2024 yang diperoleh dari Dinas Pertanian dan Perkebunan Kabupaten Bireuen serta Badan Pusat Statistik (BPS) Kabupaten Bireuen. Data mencakup 17 kecamatan dengan empat variabel utama, yaitu luas tanam (hektar), luas panen (hektar), produktivitas (ton/ha), dan total produksi (ton). Seluruh variabel dipilih karena mampu menggambarkan kondisi produksi padi secara komprehensif di setiap kecamatan.

#### 3.2 Normalisasi Data

Sebelum dilakukan proses clustering, data terlebih dahulu dinormalisasi menggunakan metode Z-Score. Normalisasi diperlukan karena setiap variabel memiliki satuan dan rentang nilai yang berbeda,

<http://sistemasi.ftik.unisi.ac.id>

sehingga perlu disetarakan skalanya agar tidak ada variabel yang mendominasi perhitungan jarak. Rumus normalisasi Z-Score adalah  $Z = (X - \mu) / \sigma$ , di mana X adalah nilai data asli,  $\mu$  adalah rata-rata, dan  $\sigma$  adalah standar deviasi. Dengan normalisasi ini, seluruh data akan memiliki rata-rata 0 dan standar deviasi 1.

### 3.3 Algoritma K-Means

K-Means merupakan algoritma clustering berbasis partisi yang membagi data ke dalam k buah cluster berdasarkan kemiripan karakteristik antar data. Algoritma ini bekerja secara iteratif dengan meminimalkan jarak antara data dengan centroid clusternya. Dalam penelitian ini, nilai k ditentukan sebesar 3 cluster, yaitu klaster produksi tinggi, sedang, dan rendah. Inisialisasi centroid awal dilakukan secara acak menggunakan parameter `random_state = 2` untuk memastikan reproduisibilitas hasil.

### 3.4 Perhitungan Euclidean Distance

Setelah centroid awal ditentukan, setiap data dihitung jaraknya terhadap masing-masing centroid menggunakan Euclidean Distance. Data akan ditempatkan ke dalam klaster dengan jarak terdekat terhadap centroid. Setelah seluruh data dikelompokkan, centroid diperbarui dengan menghitung rata-rata dari seluruh anggota dalam masing-masing klaster. Proses ini diulang secara iteratif hingga tidak terjadi lagi perubahan keanggotaan klaster atau mencapai jumlah iterasi maksimum.

### 3.5 Evaluasi Davies-Bouldin Index

Davies-Bouldin Index (DBI) merupakan metode evaluasi internal yang digunakan untuk mengukur kualitas hasil clustering. DBI mengevaluasi dua aspek utama, yaitu kohesi (kedekatan antar anggota dalam satu cluster) dan separasi (jarak antar klaster). Perhitungan DBI terdiri dari tiga tahap. Pertama, menghitung Sum of Squares Within Cluster (SSW) yang mengukur kohesi internal cluster. Kedua, menghitung jarak antar centroid untuk mengukur separasi antar klaster. Ketiga, menghitung rasio similarity ( $R_{ij}$ ) antara setiap pasangan cluster. Nilai DBI diperoleh dari rata-rata nilai rasio maksimum setiap cluster. Semakin kecil nilai DBI mendekati nol, semakin baik kualitas clustering yang dihasilkan.

## 4 Hasil dan Pembahasan

Penelitian ini dilaksanakan di Kantor Dinas Pertanian dan Perkebunan Bireuen serta di Kantor Badan Pusat Statistik (BPS) Kabupaten Bireuen. Data yang dikumpulkan berupa data produksi padi di Kabupaten Bireuen selama periode tahun 2020 hingga 2024. Pelaksanaan penelitian berlangsung pada bulan Juni 2025 hingga Juli 2025.

Dataset yang digunakan merupakan data produksi padi pada tingkat kecamatan di Kabupaten Bireuen selama periode 2020–2024, dengan empat variabel utama: luas tanam (hektar), luas panen (hektar), produktivitas (ton/ha), dan total produksi (ton). Seluruh variabel dipilih karena mampu menggambarkan kondisi produksi padi secara kuantitatif dan komprehensif. Pembagian klaster menggunakan tiga kategori yaitu produksi tinggi (Klaster 1), produksi sedang (Klaster 2), dan produksi rendah (Klaster 3).

**Tabel 1 Data produksi padi tahun 2020**

| Kecamatan   | Luas Tanam (ha) | Luas Panen (ha) | Produktivitas (ton/ha) | Produksi (ton) |
|-------------|-----------------|-----------------|------------------------|----------------|
| Samalanga   | 2.897           | 2.759           | 6,79                   | 18.750         |
| SP. Mamplam | 3.723           | 3.546           | 6,36                   | 22.560         |
| Pandrah     | 2.011           | 1.915           | 6,55                   | 12.536         |
| Jeunib      | 2.644           | 2.518           | 6,99                   | 17.600         |
| Peulimbang  | 1.620           | 1.543           | 6,56                   | 10.120         |
| Peudada     | 3.225           | 3.071           | 6,12                   | 18.780         |

|                   |       |       |      |        |
|-------------------|-------|-------|------|--------|
| Juli              | 1.041 | 991   | 5,99 | 5.930  |
| Jeumpa            | 2.576 | 2.453 | 6,38 | 15.647 |
| Kota Juang        | 896   | 853   | 7,10 | 6.235  |
| Kuala             | 2.042 | 1.945 | 7,14 | 13.890 |
| Jangka            | 3.171 | 3.020 | 6,81 | 20.568 |
| Peusangan         | 4.058 | 3.865 | 7,24 | 27.975 |
| Peusangan Selatan | 1.351 | 1.287 | 5,39 | 7.625  |
| Ps. Siblah Krueng | 2.077 | 1.978 | 5,95 | 11.762 |
| Makmur            | 2.353 | 2.241 | 5,79 | 12.982 |
| Gandapura         | 2.148 | 2.046 | 5,86 | 11.985 |
| Kuta Blang        | 2.949 | 2.809 | 6,40 | 17.982 |

#### 4.1 Normalisasi Data

Sebelum dilakukan perhitungan jarak, data terlebih dahulu dinormalisasi menggunakan metode Z-Score dengan rumus  $Z = (X - \mu) / \sigma$ . Proses normalisasi bertujuan untuk menyetarakan skala antar variabel, mengingat setiap atribut memiliki satuan dan rentang nilai yang berbeda. Apabila normalisasi tidak dilakukan, variabel dengan nilai yang lebih besar dapat mendominasi hasil perhitungan jarak. Pada data tahun 2020, diperoleh parameter normalisasi sebagai berikut.

**Tabel 2 Parameter normalisasi data tahun 2020**

| Fitur         | Mean ( $\mu$ ) | Standar Deviasi ( $\sigma$ ) |
|---------------|----------------|------------------------------|
| Luas Tanam    | 2.398,9412     | 863,6614                     |
| Luas Panen    | 2.284,7059     | 822,6120                     |
| Produktivitas | 6,4365         | 0,5186                       |
| Produksi      | 14.141,3845    | 6.752,7530                   |

Sebagai contoh, normalisasi untuk Kecamatan Samalanga pada variabel Luas Tanam dihitung sebagai berikut:  $Z = (2.897 - 2.398,94) / 863,66 = 0,5767$ . Proses yang sama diterapkan pada seluruh variabel dan seluruh kecamatan hingga diperoleh data ternormalisasi lengkap.

#### 4.2 Inisialisasi Centroid dan Iterasi K-Means (Tahun 2020)

Setelah normalisasi, centroid awal ditentukan menggunakan parameter `random_state = 2` untuk memastikan konsistensi reproduksi hasil. Dari proses inisialisasi diperoleh tiga centroid awal berikut.

**Tabel 3 Centroid awal data padi tahun 2020**

| Klaster | Luas Tanam | Luas Panen | Produktivitas | Produksi |
|---------|------------|------------|---------------|----------|
| C1      | 0,5767     | 0,5766     | 0,6817        | 0,6825   |
| C2      | 1,5331     | 1,5333     | -0,1475       | 1,2467   |
| C3      | 0,9565     | 0,9559     | -0,6103       | 0,6869   |

Pada Iterasi 1, dihitung jarak setiap kecamatan terhadap masing-masing centroid menggunakan Euclidean Distance. Sebagai contoh, perhitungan jarak Kecamatan Samalanga terhadap C1:

$$d(\text{Samalanga}, C1) = \sqrt{((0,5767-0,5767)^2 + (0,5766-0,5766)^2 + (0,6817-0,6817)^2 + (0,6825-0,6825)^2)} = 0$$

Terhadap C2:

$$d(\text{Samalanga}, C2) = \sqrt{((-0,9564)^2 + (-0,9567)^2 + (0,8292)^2 + (-0,5642)^2)} = \sqrt{2,8358} = 1,6840$$

Terhadap C3:

$$d(\text{Samalanga}, C3) = \sqrt{((-0,3798)^2 + (-0,3793)^2 + (1,2920)^2 + (-0,0044)^2)} = \sqrt{1,9574} = 1,3990$$

Karena jarak terkecil berada pada C1 (= 0), maka Kecamatan Samalanga dikelompokkan ke Klaster 1 (Rendah) pada Iterasi 1. Seluruh hasil iterasi 1 ditampilkan pada tabel berikut.

**Tabel 4 Hasil jarak euclidean iterasi 1 (tahun 2020)**

| No | Kecamatan         | d(C1) | d(C2) | d(C3) | Klaster     |
|----|-------------------|-------|-------|-------|-------------|
| 1  | Samalanga         | 0     | 1,684 | 1,399 | K1 (Rendah) |
| 2  | SP. Mamplam       | 1,684 | 0     | 1,092 | K2 (Tinggi) |
| 3  | Pandrah           | 3,165 | 4,375 | 3,516 | K1 (Rendah) |
| 4  | Jeunib            | 0,591 | 2,267 | 1,936 | K1 (Rendah) |
| 5  | Peulimbang        | 2,490 | 3,924 | 3,044 | K1 (Rendah) |
| 6  | Peudada           | 1,399 | 1,092 | 0     | K3 (Sedang) |
| 7  | Juli              | 3,902 | 5,086 | 4,059 | K1 (Rendah) |
| 8  | Jeumpa            | 1,055 | 2,140 | 1,263 | K1 (Rendah) |
| 9  | Kota Juang        | 3,812 | 5,414 | 4,644 | K1 (Rendah) |
| 10 | Kuala             | 1,713 | 3,389 | 2,854 | K1 (Rendah) |
| 11 | Jangka            | 0,525 | 1,287 | 1,360 | K1 (Rendah) |
| 12 | Peusangan         | 2,497 | 1,955 | 2,895 | K2 (Tinggi) |
| 13 | Peusangan Selatan | 4,051 | 4,845 | 3,758 | K3 (Sedang) |
| 14 | Ps. Siblah Krueng | 2,345 | 3,232 | 2,173 | K3 (Sedang) |
| 15 | Makmur            | 2,289 | 2,873 | 1,783 | K3 (Sedang) |
| 16 | Gandapura         | 2,392 | 3,167 | 2,091 | K3 (Sedang) |
| 17 | Kuta Blang        | 0,765 | 1,439 | 0,714 | K3 (Sedang) |

Setelah iterasi 1, centroid diperbarui dengan menghitung rata-rata data dalam masing-masing klaster. Proses iterasi berlanjut hingga iterasi ke-5, di mana tidak lagi terjadi perubahan keanggotaan klaster. Nilai perubahan centroid (centroid shift) pada setiap iterasi adalah: Iterasi 1 = 7,834302; Iterasi 2 = 1,865805; Iterasi 3 = 0,623141; Iterasi 4 = 0,142075; Iterasi 5 = 0. Kondisi konvergen tercapai pada iterasi ke-5.

Penentuan label klaster dilakukan berdasarkan nilai centroid akhir pada skala asli. Klaster 1 (Rendah) memiliki rata-rata produksi sebesar 7.237,51 ton yang merupakan nilai terendah, sehingga diberi label "Rendah". Klaster 2 (Tinggi) memiliki rata-rata produksi sebesar 20.602,14 ton yang merupakan nilai tertinggi, sehingga diberi label "Tinggi". Klaster 3 (Sedang) memiliki rata-rata produksi sebesar 12.000,2 ton yang berada di antara keduanya, sehingga diberi label "Sedang".

**Tabel 5 Centroid final (skala asli) data padi tahun 2020**

| Klaster     | Luas Tanam (ha) | Luas Panen (ha) | Produktivitas (ton/ha) | Produksi (ton) |
|-------------|-----------------|-----------------|------------------------|----------------|
| K1 (Rendah) | 1.522           | 1.449,4         | 6,67                   | 7.237,51       |
| K2 (Tinggi) | 3.238,14        | 3.084           | 6,67                   | 20.602,14      |
| K3 (Sedang) | 2.101           | 2.001           | 5,87                   | 12.000,2       |

#### 4.3 Perhitungan Davies-Bouldin Index (DBI) Tahun 2020

Setelah clustering mencapai konvergensi, dilakukan evaluasi kualitas hasil pengelompokan menggunakan DBI. Perhitungan DBI mencakup tiga tahap utama: (1) menghitung nilai Si (rata-rata

jarak dalam klaster), (2) menghitung jarak antar centroid ( $M_{ij}$ ), dan (3) menghitung rasio similarity ( $R_{ij}$ ).

Nilai  $S_i$  dihitung menggunakan rumus rata-rata jarak seluruh anggota klaster terhadap centroidnya. Sebagai contoh detail untuk Klaster 1 (5 anggota):

$$S_1 = (1,3555 + 0,5014 + 1,5387 + 1,3293 + 1,5888) / 5 = 1,2627$$

**Tabel 6 Rata-rata jarak dalam klaster ( $s_i$ ) tahun 2020**

| Klaster            | Jumlah Data | $S_i$    |
|--------------------|-------------|----------|
| Klaster 1 (Rendah) | 5           | 1,262746 |
| Klaster 2 (Tinggi) | 7           | 1,024764 |
| Klaster 3 (Sedang) | 5           | 0,747275 |

Jarak antar centroid dihitung menggunakan Euclidean Distance. Sebagai contoh, jarak antara C1 dan C2:

$$M_{1,2} = \sqrt{((-1,0154 - 0,9717)^2 + (-1,0154 - 0,9717)^2 + (0,4465 - 0,4558)^2 + (-1,0224 - 0,9568)^2)} = 3,4371$$

**Tabel 7 Jarak antar centroid dan rasio similarity tahun 2020**

| Pasangan | $M_{ij}$ | $S_i$  | $S_j$  | $R_{ij}$ | $D_i$ (max $R_{ij}$ ) |
|----------|----------|--------|--------|----------|-----------------------|
| C1–C2    | 3,4371   | 1,2627 | 1,0248 | 0,6655   | 1,0393                |
| C1–C3    | 1,9341   | 1,2627 | 0,7473 | 1,0393   | —                     |
| C2–C3    | 2,7318   | 1,0248 | 0,7473 | 0,6487   | 0,6655                |

Nilai DBI final dihitung sebagai rata-rata dari nilai  $D_i$  masing-masing klaster:

$$DBI = (1/3) \times (1,0393 + 0,6655 + 1,0393) = (1/3) \times 2,7441 = 0,9147$$

Nilai DBI tahun 2020 sebesar 0,9147 menunjukkan bahwa struktur klaster yang terbentuk sudah cukup baik karena mendekati nilai 0. Hal ini mengindikasikan bahwa jarak antar klaster relatif cukup terpisah dan variasi data di dalam masing-masing klaster masih dalam batas yang dapat diterima.

#### 4.4 Hasil Clustering dan DBI Lima Tahun (2020–2024)

Setelah seluruh proses clustering dan evaluasi dilakukan untuk setiap tahun, diperoleh rekapitulasi nilai DBI dan distribusi anggota cluster sebagai berikut.

**Tabel 8 Rekapitulasi nilai DBI dan distribusi klaster 2020–2024**

| Tahun | Nilai DBI | K1 (Rendah) | K2 (Tinggi) | K3 (Sedang) | Ket       |
|-------|-----------|-------------|-------------|-------------|-----------|
| 2020  | 0,9147    | 5 kec       | 7 kec       | 5 kec       |           |
| 2021  | 0,8292    | 5 kec       | 6 kec       | 6 kec       |           |
| 2022  | 0,6119    | 4 kec       | 7 kec       | 6 kec       | ★ Terbaik |
| 2023  | 0,8821    | 5 kec       | 6 kec       | 6 kec       |           |
| 2024  | 0,8597    | 4 kec       | 7 kec       | 6 kec       |           |

Berdasarkan hasil evaluasi lima tahun, nilai DBI terbaik diperoleh pada tahun 2022 dengan nilai 0,6119. Hal ini menunjukkan bahwa pada tahun 2022, struktur klaster yang terbentuk memiliki tingkat pemisahan antar klaster yang paling jelas dan kohesi dalam klaster yang paling rapat

dibandingkan tahun lainnya. Sebaliknya, nilai DBI tertinggi terjadi pada tahun 2023 sebesar 0,8821, mengindikasikan bahwa pemisahan klaster pada tahun tersebut kurang optimal.

Secara keseluruhan, seluruh nilai DBI yang diperoleh (0,6119–0,9147) berada dalam rentang yang tergolong baik, karena semua nilai mendekati nol. Hal ini membuktikan bahwa algoritma K-Means berhasil membentuk pengelompokan yang bermakna dan dapat diinterpretasikan untuk keperluan analisis produksi padi di Kabupaten Bireuen.

Hasil pengelompokan menunjukkan adanya pola yang konsisten di mana Kecamatan Peusangan dan SP. Mamplam secara konsisten masuk dalam klaster produksi tinggi atau sedang, sedangkan kecamatan seperti Juli dan Peusangan Selatan cenderung berada di klaster produksi rendah. Informasi ini sangat bermanfaat bagi pemerintah daerah dalam menetapkan prioritas pembinaan dan alokasi sumber daya pertanian.

## 5 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut. Pertama, penerapan algoritma K-Means mampu mengelompokkan data produksi padi ke dalam 3 klaster yaitu klaster produksi tinggi, sedang, dan rendah berdasarkan variabel luas tanam, luas panen, produktivitas, dan total produksi. Proses klastering berjalan dengan baik pada setiap tahun data dan menghasilkan pengelompokan yang konsisten sesuai karakteristik masing-masing kecamatan. Kedua, evaluasi menggunakan Davies-Bouldin Index menunjukkan bahwa kualitas klaster yang dihasilkan tergolong baik dengan rentang nilai DBI antara 0,6119 hingga 0,9147. Nilai DBI terbaik diperoleh pada tahun 2022 sebesar 0,6119, yang berarti struktur klaster pada tahun tersebut memiliki tingkat pemisahan antar klaster yang lebih jelas dan kohesi dalam klaster yang lebih rapat. Ketiga, hasil pengelompokan selama lima tahun menunjukkan adanya perbedaan pola produksi yang signifikan antar kecamatan, sehingga dapat menjadi dasar evaluasi bagi pemerintah daerah dalam menentukan strategi peningkatan produksi. Kecamatan yang berada pada klaster rendah dapat menjadi prioritas pembinaan, sedangkan klaster tinggi dapat dijadikan wilayah percontohan dalam pengelolaan produksi padi di Kabupaten Bireuen.

## Referensi

- [1] Wijayanto, S., & Yoka Fathoni, M. (2021). Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah menggunakan Metode *Clustering K-Means*. *Jupiter*, 13(2), 212-219.
- [2] Wahyudin, E., Amir Rudin, R., Kaslani, K., & Eka Permana, S. (2024). Penerapan Data Mining Pengelompokan Produktivitas Padi menggunakan Algoritma *K-Means* pada Provinsi Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 522-528.
- [3] Al Masykur, A., Gusti, S. K., Sanjaya, S., Yanto, F., & Syafria, F. (2023). Penerapan Metode *K-Means Clustering* untuk Pemetaan Pengelompokan Lahan Produksi Jagung di Kabupaten Pasuruan. *Jurnal Informatika*, 10(1), 8364-8372.
- [4] Sopyan, Y., Lesmana, A. D., & Juliane, C. (2022). Analisis Algoritma *K-Means* dan *Davies Bouldin Index* dalam mencari *Cluster* Terbaik Kasus Perceraian di Kabupaten Kuningan. *Building of Informatics, Technology and Science (BITS)*, 4(3), 1464-1470.
- [5] Fathurrahman, S. H. (2023). Evaluasi *Clustering K-Means* dan *K-Medoid* pada Persebaran Covid-19 di Indonesia dengan Metode *Davies-Bouldin Index (DBI)*. *Jurnal Ilmiah*, 6(2), 117-128.
- [6] Nur Afifah, I. A., & Nurdianto, H. (2023). *Data Mining Clustering* dalam Pengelompokan Buku Perpustakaan menggunakan *Algoritma K-Means*. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 8(3), 802-814.
- [7] Handoyo, R., & Saputra, R. A. (2021). Penerapan Metode *K-Means Clustering* untuk Pengelompokan Data Produksi Padi di Provinsi Jawa Tengah. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(4), 703-712.
- [8] Putra, A. B., & Qamal, M. (2024). Perbandingan *K-Means* dan *Fuzzy C-Means* untuk *Clustering* Produksi Padi di Kabupaten Aceh Utara. *Jurnal Informatika Kaputama*, 8(2), 210-220.

- [9] Pramesti, D. A., & Fajar, M. (2023). Penerapan *Data Mining* untuk Prediksi Hasil Panen Padi menggunakan Metode *K-Means Clustering*. *Jurnal Sistemasi*, 12(2), 178-188.
- [10] Zulkarnain, Z., & Bustami, B. (2022). Implementasi *Algoritma K-Means* untuk Pemetaan Daerah Produksi Padi di Kabupaten Aceh Timur. *Jurnal Teknologi Informasi dan Komputer*, 8(1), 89-98.
- [11] Prasetyo, E., & Fahmi, H. (2022). Analisis Perbandingan Metode *K-Means* Dan *K-Medoids Clustering* pada Data Produksi Tanaman Pangan. *Jurnal Sains Komputer dan Informatika*, 6(2), 123-135.
- [12] Suryani, D., & Wahyuni, S. (2021). Evaluasi *Davies-Bouldin Index* pada *Clustering* Produksi Tanaman Pangan di Indonesia. *Seminar Nasional Teknologi Informasi dan Komunikasi*, 1(1), 45-52.
- [13] Rahmawati, N., & Hartati, S. (2024). Analisis *Clustering* Data Pertanian menggunakan Metode *K-Means* dengan Optimasi Centroid. *Jurnal Ilmu Komputer dan Agri-Informatika*, 11(1), 33-43.
- [14] Sinaga, K. P., & Yang, M. S. (2024). *Unsupervised K-Means Clustering Algorithm*. *IEEE Access*, 12, 45678-45691.
- [15] Amalia, R., & Fadlisya, F. (2023). Pengelompokan Data Produksi Padi di Provinsi Aceh menggunakan *Algoritma K-Means*. *Techsi - Jurnal Teknik Informatika*, 15(1), 56-67.