

## PENERAPAN DATA MINING PENGELOMPOKAN PRODUKTIVITAS PADI MENGGUNAKAN ALGORITMA K-MEANS PADA PROVINSI JAWA BARAT

Edi Wahyudin<sup>1</sup>, Rizki Amir Rudin<sup>2</sup>, Kaslani<sup>3</sup>, Sandy Eka Permana<sup>4</sup>

<sup>1,3</sup> Program Studi Rekayasa Perangkat Lunak S1, STMIK IKMI Cirebon

<sup>2</sup> Teknik Informatika S1, STMIK IKMI Cirebon

<sup>4</sup> Program Studi Manajemen Informatika D3, STMIK IKMI Cirebon

Jln Perjuangan No 10 B Karyamulya Kesambi Kota Cirebon Jawa Barat Indonesia

ediwahyudin@gmail.com

### ABSTRAK

Padi, sebagai makanan pokok di Indonesia, terus menghadapi peningkatan kebutuhan yang tidak sejalan dengan ketersediaan lahan pertanian di Jawa Barat. Belum adanya pemetaan produktivitas panen padi di setiap kabupaten atau kota menjadi tantangan bagi pemerintah provinsi. Penelitian ini menggunakan algoritma *K-Means clustering* dan aplikasi *RapidMiner 10.2* yang mengacu metode penelitian *Knowledge Discovery in Database (KDD)* untuk mengelompokkan produktivitas padi di Jawa Barat, menghasilkan 18 *cluster*. *Cluster 0* hingga 6 menunjukkan tingkat produktivitas tinggi, sementara *cluster 9* adalah sedang. Informasi ini menjadi kontribusi berharga bagi pemerintah provinsi dalam merancang kebijakan pertanian dan pangan yang optimal untuk meningkatkan produktivitas padi di setiap daerah. Dengan demikian, hasil dari penelitian ini adalah diharapkan memberikan wawasan mendalam tentang potensi pertanian di kabupaten atau kota, membantu dalam perencanaan strategi tanam, dan mendorong optimalisasi luas area pertanian di Jawa Barat agar semakin maju lagi.

**Kata kunci :** padi, lahan, produktivitas, clustering, k-means

### 1. PENDAHULUAN

Kebutuhan pangan yang terus meningkat sejalan dengan pertumbuhan populasi menuntut peningkatan produktivitas tanaman padi, komoditas pangan krusial di seluruh dunia [1]. Namun, peningkatan ini terhambat oleh konversi lahan pertanian, khususnya lahan sawah, yang sulit dihindari terutama di kota-kota besar [2]. Di Provinsi Jawa Barat, perbedaan produksi padi antar kabupaten atau kota menjadi masalah, terutama karena ketidaksetaraan ketersediaan lahan pertanian [3].

Pemerintah Provinsi Jawa Barat perlu mengatasi ketidaksetaraan tersebut dengan melakukan pemetaan dan analisis produktivitas panen padi. Penelitian ini menerapkan teknik data mining dengan algoritma *K-Means clustering* untuk menganalisis tingkat produktivitas panen padi di tiap kabupaten atau kota. Melalui pendekatan *Knowledge Discovery In Databases (KDD)*, *clustering* secara otomatis mengelompokkan data berdasarkan karakteristik yang sama [4]. Algoritma *K-Means* dipilih karena waktu komputasinya relatif cepat dan hasilnya mudah diinterpretasikan. Hasil pengelompokan memberikan panduan bagi pemerintah provinsi dalam mengidentifikasi wilayah-wilayah yang memerlukan peningkatan produktivitas padi, memastikan keberlanjutan pasokan pangan sesuai dengan kebutuhan masyarakat.

Penelitian ini merujuk pada metodologi sebelumnya, khususnya studi oleh Sena Wijayanto dan M. Yoka Fathoni, yang menggunakan teknik dan algoritma yang sama untuk mengelompokkan produktivitas tanaman padi di Jawa Tengah dengan metode *K-Means clustering*. [5] berhasil mengimplementasikan algoritma ini, menghasilkan

tiga *cluster* dengan karakteristik produktivitas yang berbeda.

Sementara itu, penelitian oleh [3] memfokuskan pada pengelompokan kabupaten/kota di Provinsi Jawa Barat berdasarkan faktor produksi padi, seperti luas tanam, luas panen, dan produktivitas. Hasilnya menunjukkan tiga *cluster* dengan perbedaan produksi padi, memberikan landasan bagi pemerintah Jawa Barat untuk kebijakan peningkatan produksi padi.

Sejalan dengan itu, penelitian oleh [6] mengaplikasikan *K-Means Clustering* untuk mengelompokkan kecamatan di Kabupaten Bima berdasarkan produksi dan luas panen bawang merah. Hasil analisis menyajikan tiga kelompok dengan karakteristik produksi yang berbeda.

Penelitian ini mengadopsi pendekatan serupa untuk mengelompokkan produktivitas padi di Jawa Barat, dengan fokus pada potensi produktivitas kabupaten/kota. Informasi yang dihasilkan diharapkan dapat menjadi acuan bagi pemerintah provinsi dalam mengoptimalkan strategi pertanian, memastikan pasokan pangan yang memadai, dan meningkatkan produktivitas padi di seluruh daerah. Teknik *clustering*, khususnya algoritma *K-Means*, digunakan untuk menganalisis data produktivitas padi dengan memanfaatkan *data mining* [7], [8].

Dalam konteks keberlanjutan pangan, produktivitas padi memegang peranan vital sebagai sumber pangan utama. Penting bagi pemerintah untuk meningkatkan produktivitas panen padi secara merata di berbagai daerah, memastikan kebutuhan pangan masyarakat terpenuhi. Oleh karena itu, penekanan pada pengelompokan produktivitas padi berdasarkan kabupaten/kota di Jawa Barat menjadi krusial. Tujuannya adalah memberikan wawasan terkait

potensi dan tantangan produktivitas padi di setiap daerah. Informasi ini tidak hanya bermanfaat bagi pemerintah dalam perumusan kebijakan pertanian yang efektif, tetapi juga sebagai landasan bagi penelitian masa depan yang memperhatikan potensi produktivitas padi. Oleh karena itu, skripsi ini diajukan dengan judul "Penerapan *Data Mining* Pengelompokan Produktivitas Padi Menggunakan Algoritma *K-Means* pada Provinsi Jawa Barat".

## 2. LITERATURE REVIEW

### 2.1. Metode Literature Review

Penelitian ini menggunakan metode *literature review* dengan fokus pada penerapan *data mining*, khususnya algoritma *K-means*, dalam pengelompokan produktivitas padi berdasarkan kabupaten/kota di Jawa Barat [9]. Tahapan *literature review* melibatkan formulasi permasalahan, pencarian literatur menggunakan kata kunci "Algoritma *K-means*" dan "Pengelompokan produktivitas padi," evaluasi data dengan memilih 15 artikel relevan dari lima tahun terakhir (2019-2023), dan analisis interpretatif. Pelaksanaan *literature review* menggabungkan teknik perbandingan, kontras, kritik, sintesis, dan ringkasan untuk mendapatkan wawasan mendalam terkait penerapan *K-means* dalam konteks pengelompokan produktivitas padi.

### 2.2. Hasil Literature Review

Hasil *literature review* menyoroti beberapa penelitian terkait penerapan metode *K-means* dalam pengelompokan produktivitas padi dan tanaman pangan lainnya. Dalam penelitian oleh [5], dilakukan pengelompokan produktivitas padi di Jawa Tengah, menghasilkan tiga *cluster* dengan karakteristik sedang, rendah, dan tinggi. Begitu juga dalam penelitian [10] terkait kelapa sawit di Riau, dan [3] terkait faktor produksi padi di Jawa Barat, keduanya menggunakan *K-means* dan menghasilkan *cluster* dengan karakteristik produktivitas yang berbeda.

Penelitian lainnya, seperti yang dilakukan [6], mengenai produksi bawang merah di Bima, serta [2], mengenai analisis potensi lahan pangan di Kalimantan Selatan, juga memanfaatkan *K-means* dan menghasilkan kelompok dengan karakteristik produksi yang beragam [11].

Dalam konteks ketahanan pangan, penelitian di Jawa Timur oleh Indonesia (2023) menunjukkan bahwa *K-means* dapat digunakan untuk mengelompokkan ketahanan pangan menjadi tiga kelompok, yaitu tinggi, sedang, dan rendah. Sementara itu, [12] menggunakan *K-means* untuk mengelompokkan produksi padi di Jawa Barat, menghasilkan tiga *cluster* dengan karakteristik produksi yang berbeda.

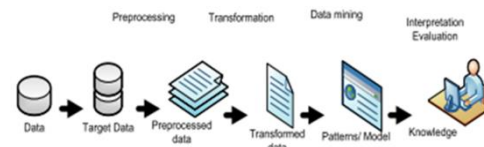
Terakhir, penelitian yang melibatkan klusterisasi potensi desa di Bojonegoro oleh [13], serta klusterisasi hasil pertanian di Jambi oleh [14], menunjukkan penerapan *K-means* dalam menghasilkan kelompok dengan karakteristik potensi atau hasil pertanian yang

berbeda [15]. Dalam semua penelitian ini, *K-means* terbukti menjadi alat yang efektif untuk menganalisis dan mengelompokkan data produktivitas tanaman dengan hasil yang beragam dan informatif.

Dalam beberapa penelitian, fokus utama adalah meningkatkan produktivitas dan luas lahan pertanian, khususnya tanaman padi. Menyusutnya produktivitas dan luas lahan merupakan masalah serius. Solusi yang diusulkan melibatkan klusterisasi berdasarkan produktivitas dan luas panen. Proses klusterisasi menggunakan algoritma *K-Means* dan algoritma lainnya, dengan *K-Means* menjadi pilihan utama karena kemampuannya menghasilkan kluster optimal. Hasil klusterisasi mengidentifikasi kelompok seperti tinggi dan rendah, dengan kelompok rendah menunjukkan daerah yang memerlukan perhatian khusus untuk peningkatan.

## 3. METODE PENELITIAN

### 3.1. Metode Penelitian



Gambar 1. Tahapan alur penelitian

Gambar 1 menggambarkan tahapan *Knowledge Discovery in Database (KDD)*, yang mencakup algoritma untuk eksplorasi data, pembangunan model, dan penemuan pola yang belum diketahui [16]. Tahap perancangan pada skripsi ini menggunakan metode KDD, yang merupakan proses untuk menggali informasi berharga dari basis data besar dengan tujuan menghasilkan pengetahuan bagi para pemangku kepentingan [17].

Tabel 1. Alur penelitian

Tahapan	Aktivitas	Deskripsi Aktivitas
<i>Data selection</i>	Melakukan pengumpulan data produktivitas padi pada kabupaten atau kota di Provinsi Jawa Barat.	Data diperoleh dari dokumentasi website resmi pemerintah Provinsi Jawa Barat, <a href="http://opendata.jabarprov.go.id">opendata.jabarprov.go.id</a> , dan akan diolah menggunakan RapidMiner.
<i>Data Pre-Processing</i>	Melakukan pembersihan data	Proses <i>preprocessing</i> melibatkan pembersihan data kosong, kesalahan penginputan, duplikat, dan data tidak relevan, yang berkontribusi pada akurasi hasil <i>data mining</i> .
<i>Data Transformation</i>	Melakukan penyesuaian	Transformasi data dilakukan untuk mengubah format data

Tahapan	Aktivitas	Deskripsi Aktivitas
	bentuk atau format data.	agar dapat diaplikasikan pada proses <i>data mining</i> , khususnya algoritma <i>K-Means</i> , yang memerlukan data dalam bentuk numerik.
<i>Data Mining</i>	Melakukan penerapan algoritma clustering	Pada fase ini, diterapkan algoritma atau metode kecerdasan untuk mengekstrak pola informasi potensial dari data yang dipilih.
<i>Data Interpretation</i>	Melakukan penyimpulan data menjadi pola informasi yang mudah dipahami	Tahap evaluasi menyimpulkan pola informasi dari proses <i>data mining</i> , kemudian hasilnya diinterpretasikan untuk memudahkan pemahaman.

Tabel 1 menyajikan dengan rinci aktivitas yang dilakukan pada setiap tahap penelitian, memberikan gambaran komprehensif tentang langkah-langkah yang diambil selama proses penelitian.

### 3.2. Sumber Data

Dalam penelitian mengenai produktivitas padi di kabupaten atau kota di Jawa Barat, data diperoleh dari sumber resmi pemerintah Provinsi Jawa Barat di <https://opendata.jabarprov.go.id/id/dataset/produktivitas-padi-kerangka-sampel-area-ksa-berdasarkan-bulan-di-jawa-barat>. Dataset ini mencakup 1620 *record* dengan 11 atribut, seperti id, Kode Provinsi, Nama Provinsi, Kode Kabupaten Kota, Nama Kabupaten Kota, Bulan, Luas Panen, Satuan, Produktivitas Padi, Satuan, dan Tahun. Rentang waktu data mencakup periode tahun 2018 hingga 2022, diambil pada tanggal 23 November 2023 pukul 17.53 WIB. Dalam penelitian ini, digunakan data sekunder, yang merupakan data yang diambil secara tidak langsung, dapat berasal dari berbagai sumber seperti jurnal, buku, atau *website* resmi pemerintah.

### 3.3. Teknik Pengumpulan Data

Data pada penelitian ini dikumpulkan menggunakan teknik data sekunder melalui dokumentasi dan riset pustaka. Dokumentasi mencakup data produktivitas padi di Jawa Barat tahun 2018-2022 yang diunduh dari *website* resmi pemerintah melalui *Open Data* Jabar. Riset pustaka dilakukan dengan merujuk pada jurnal, penelitian sebelumnya, dan literatur terkait.

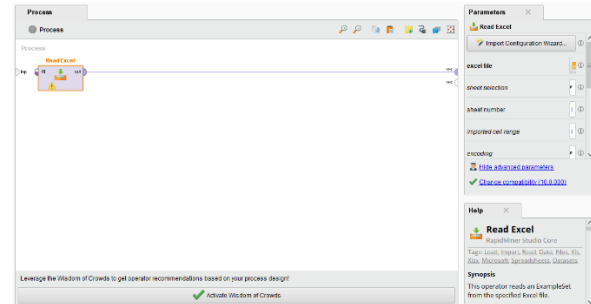
### 3.4. Teknik Analisis Data

Dalam penelitian ini, analisis data dilakukan melalui pendekatan deskriptif dengan menggunakan teknik analisis *K-means* sebagai algoritma utama.

Pendekatan deskriptif membantu memberikan gambaran rinci tentang karakteristik data, sementara *K-means* digunakan untuk mengelompokkan data secara optimal. Dengan kombinasi kedua teknik ini, penelitian dapat memberikan wawasan mendalam mengenai pola dan struktur yang terdapat dalam dataset produktivitas padi di Jawa Barat.

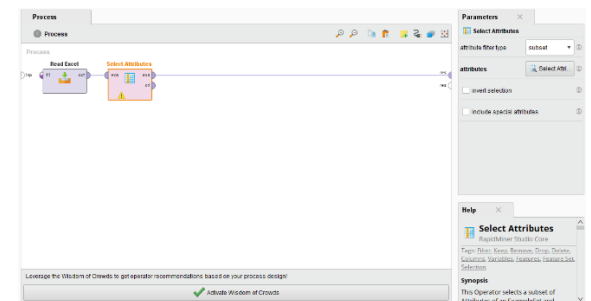
## 4. HASIL DAN PEMBAHASAN

### 4.1. Seleksi Data



Gambar 2. Operator *Read Excel*

Gambar 2 menggambarkan operator *Read Excel* yang terlibat dalam proses. Dalam tahapan seleksi data, terdapat dua langkah utama, yakni membaca *file excel* dan memilih atribut. Pada langkah pertama, yaitu membaca *file excel*, operator *Read Excel* bertugas untuk mengakses data produktivitas padi di Kabupaten atau Kota di Jawa Barat. Data yang diambil pada tahap ini bersifat murni tanpa melalui proses *editing* atau penghapusan data.



Gambar 3. Operator *Select Attributes*

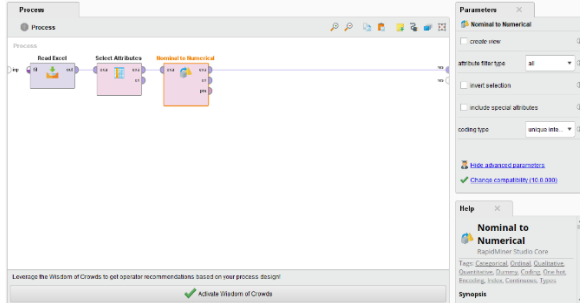
Gambar 3 memperlihatkan tahapan kedua dalam proses seleksi data, yaitu tahap "*select attributes*." Operator *select attributes* bertujuan untuk memilih atribut mana yang akan digunakan dalam proses data mining. Dalam langkah seleksi ini, terjadi penghapusan atribut yang tidak relevan untuk proses data mining. Data yang akan digunakan dalam analisis merupakan data yang sesuai dengan variabel perhitungan yang relevan dalam proses analisis. Peneliti menggunakan 6 atribut pada proses data mining, yaitu Bulan, Luas Panen, Nama Kabupaten atau Kota, Produksi Padi, Produktivitas Padi, dan Tahun.

### 4.2. Data Pre-processing

Langkah selanjutnya adalah tahap *data pre-processing*, yang bertujuan untuk menghilangkan atau

membersihkan nilai-nilai yang hilang agar hasil data mining optimal. Data yang digunakan dalam penelitian ini sudah lengkap dan tidak mengandung nilai kosong, sehingga prosesnya dapat dilanjutkan ke tahap berikutnya.

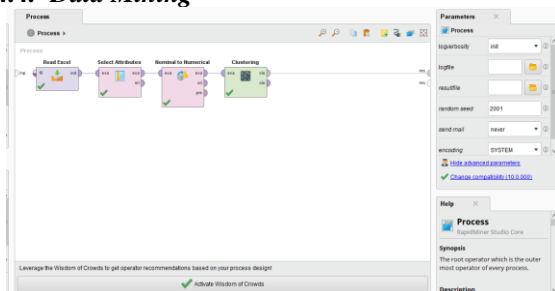
#### 4.3. Transformasi Data



Gambar 4. Proses *data transformation*

Gambar 4 menunjukkan tahap transformasi data, di mana terjadi perubahan tipe data untuk mendukung algoritma *K-means*. Pada proses *data mining*, *K-means* optimal dengan data numerik karena mengukur jarak antar data. Tahapan transformasi ini menggunakan operator *nominal to numerical* untuk mengubah data nominal menjadi format numerik dengan tipe kode unik berupa *integer*.

#### 4.4. Data Mining



Gambar 5. *Clustering* menggunakan algoritma *K-Means*

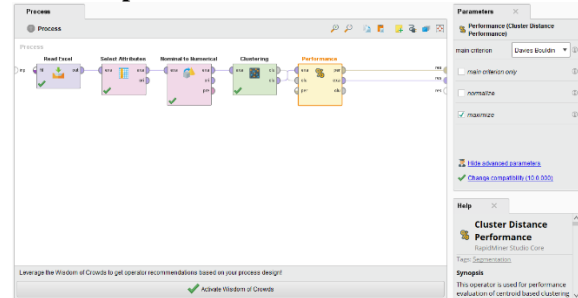
Tahap selanjutnya adalah data mining menggunakan data hasil dari transformasi data. Model *data mining* yang digunakan yaitu pengelompokan menggunakan algoritma *K-means* seperti yang terlihat pada Gambar 5.

Cluster Model	
Cluster 0:	571 items
Cluster 1:	2 items
Cluster 2:	26 items
Cluster 3:	70 items
Cluster 4:	103 items
Cluster 5:	161 items
Cluster 6:	1 items
Cluster 7:	129 items
Cluster 8:	6 items
Cluster 9:	35 items
Cluster 10:	20 items
Cluster 11:	2 items
Cluster 12:	215 items
Cluster 13:	67 items
Cluster 14:	181 items
Cluster 15:	3 items
Cluster 16:	24 items
Cluster 17:	4 items
Total number of items: 1620	

Gambar 6. *Cluster model*

Gambar 6 merupakan *cluster model* pada tahap *data mining*, nilai *k* ditentukan secara acak dengan  $k=18$ ,  $max\ run=10$ , menggunakan *euclidean distance*, dan  $max\ optimization\ steps\ 100$ . Hasilnya, terbentuk 18 kluster dengan jumlah items atau kabupaten/kota yang berbeda-beda di setiap kluster. Kluster 0 memiliki 571 items, sementara kluster lain memiliki jumlah items yang beragam, mulai dari 1 hingga 215 items.

#### 4.5. Interpretasi



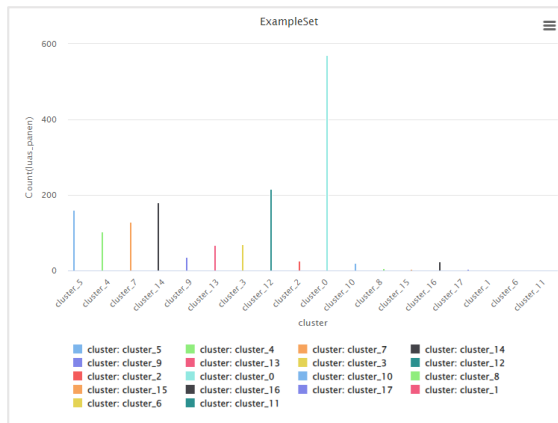
Gambar 7. Operator *Cluster Distance Performance*

Untuk menentukan nilai *k* optimal, digunakan operator *Cluster Distance Performance*, seperti yang terlihat pada Gambar 7. Operator ini berguna untuk melihat nilai *Davies Bouldin Index* (DBI). Nilai *k* optimal adalah yang memiliki DBI mendekati 0. Pada tahap ini, nilai *k* yang ditentukan secara acak dimasukkan ke dalam operator *K-means clustering*, dan kemudian nilai DBI-nya diperiksa.

Tabel 2. Nilai DBI

No	Kelompok (K)	DBI
1	2	0,537
2	3	0,544
3	4	0,514
4	5	0,478
5	6	0,478
6	7	0,470
7	8	0,479
8	9	0,477
9	10	0,474
10	11	0,472
11	12	0,457
12	13	0,462
13	14	0,458
14	15	0,464
15	16	0,445
16	17	0,455
17	18	0,442
18	19	0,478
19	20	0,481

Tabel 2 menunjukkan nilai DBI dari 19 percobaan, dan hasilnya menunjukkan bahwa nilai *k* yang paling optimal adalah  $k=18$  dengan nilai DBI sebesar 0,442.



Gambar 8. Cluster hasil pengelompokan

Gambar 8 menggambarkan diagram grafik dari hasil pengelompokan kluster. Diagram ini memberikan visualisasi yang jelas terkait distribusi dan karakteristik dari setiap kluster yang dihasilkan melalui proses pengelompokan data.

#### 4.5. Pembahasan

Dari proses *data mining* produktivitas padi pada kabupaten atau kota di Jawa Barat menghasilkan 18 *cluster* optimal.

Tabel 3. Jumlah anggota *cluster*

Cluster	Luas Panen	Produksi Padi	Produktivitas Padi	Jumlah Anggota	Ket.
Cluster 0	7.572.346	43.138.399	80.263,04	571	Sangat tinggi
Cluster 12	7.830.587	44.658.766	77.375,75	215	Tinggi 7
Cluster 14	8.042.562	45.766.302	79.412,25	181	Tinggi 6
Cluster 5	8.042.562	45.766.302	79.312,25	161	Tinggi 5
Cluster 7	7.965.623	43.678.902	77.985,03	129	Tinggi 4
Cluster 4	6.098.756	43.678.212	76.989,05	103	Tinggi 3
Cluster 3	5.098.543	42.351.234	72.897,78	70	Tinggi 2
Cluster 13	5.876.098	41.675.342	70.872,89	67	Tinggi 1
Cluster 9	4.675.897	40.897.452	69.243,87	35	Sedang
Cluster 2	4.576.982	39.762.123	65.870,54	26	Rendah 8
Cluster 16	4.321.023	38.098.534	63.231,98	24	Rendah 7
Cluster 10	3.987.478	36.987.756	63.899,76	20	Rendah 6
Cluster 8	2.067.654	32.654.098	37.876,09	6	Rendah 5
Cluster 17	2.543.765	33.876.098	36.765,23	4	Rendah 4
Cluster 15	1.789.451	25.897.098	35.987,56	3	Rendah 3
Cluster 1	1.486.762	20.786.765	35.786,98	2	Rendah 2
Cluster 11	1.236.870	12.986.764	34.034,23	2	Rendah 1
Cluster 6	51.009	310.093	60,68	1	Sangat Rendah

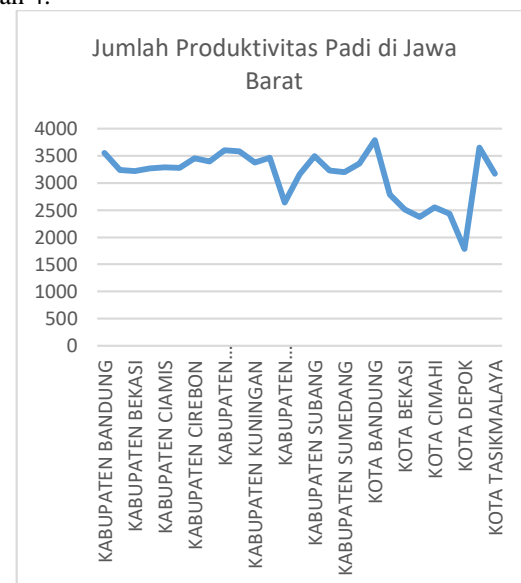
Tabel 3 menunjukkan hasil pengelompokan produktivitas padi berdasarkan cluster di Jawa Barat dari yang sangat tinggi hingga sangat rendah.

Pada *cluster* 12, 14 dan 5 produktivitas padi kisaran 12.00-8.000 lebih banyak dibandingkan pada *cluster* 7 dan 4.

#### 4.6. Produktivitas Padi

Gambar 9. Diagram produktivitas padi berdasarkan *cluster*

Gambar 9 memvisualisasikan bahwa produktivitas padi paling banyak terjadi pada *cluster* 0 artinya daerah pada *cluster* ini merupakan daerah dengan tingkat produktivitas padinya tinggi. sedangkan paling sedikit yaitu terjadi pada *cluster* 1, 11, 15, 17 dan 6 yaitu hanya di angka 200-100 kuintal.

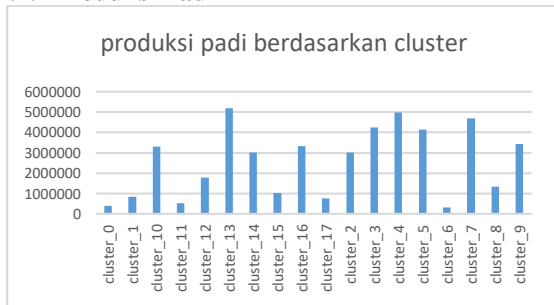


Gambar 10. Grafik produktivitas padi di Jawa Barat



Gambar 10 menampilkan hasil analisis produktivitas padi di Provinsi Jawa Barat tahun 2018-2022. Produktivitas padi tertinggi di Kota Bandung (3788,53 kwintal), Kota Sukabumi (3652,56 kwintal), dan Kabupaten Indramayu (3605,25 kwintal), sedangkan yang terendah terjadi di Kota Depok (1787,6 kwintal) dan Kota Bogor (2370,6 kwintal) per hektar.

#### 4.7. Produksi Padi



Gambar 11. Diagram produksi padi berdasarkan cluster

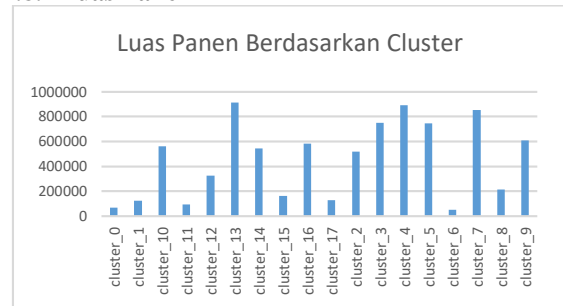
Gambar 11 menunjukkan cluster 13 memiliki produksi padi paling tinggi, mencapai lebih dari 5.000.000 ton, sedangkan cluster 10, 16, dan 9 memiliki produksi sedang, berkisar 3.500.000-3.000.000 ton. Cluster 14 dan 2 memiliki selisih produksi yang kecil, sementara cluster 6 memiliki produksi padi yang rendah.



Gambar 12. Grafik produksi padi di Jawa Barat

Gambar 12 menunjukkan Kabupaten Indramayu memimpin produksi padi di Jawa Barat dengan 6.956.356 ton, diikuti oleh Kabupaten Karawang (5.800.378 ton) dan Kabupaten Subang (4.902.925 ton). Kota Bekasi, Bogor, Cimahi, Cirebon, Depok, Sukabumi, dan Kota Tasikmalaya memiliki produksi padi yang lebih rendah, berkisar 1.000-15.000 ton.

#### 4.8. Luas Panen



Gambar 13. Diagram luas panen berdasarkan cluster

Panen terluas terjadi pada Cluster 13 dengan lebih dari 9.000.000 hektar, diikuti oleh Cluster 4 dan Cluster 7. Cluster ini mencakup daerah dengan luas panen padi yang paling besar. Sementara Cluster 3 dan 5 memiliki luas panen hampir sama, yaitu sekitar 7.000.000 hektar. Sebaliknya, Cluster 6 mencatat luas panen padi yang paling sedikit, sebagaimana terlihat pada Gambar 13.



Gambar 14. Grafik luas panen di Jawa Barat

Gambar 14 menunjukkan Kabupaten Indramayu tercatat sebagai luas panen tertinggi di Jawa Barat, mencapai 1.152.854 hektar, diikuti oleh Kabupaten Karawang dengan 969.445 hektar, dan Kabupaten Subang dengan 841.347 hektar. Sementara Kabupaten Bekasi, Cianjur, dan Majalengka memiliki luas panen sedang (400.000-500.000 hektar), sedangkan luas panen paling kecil terdapat di kota Bandung, Bekasi, Bogor, Cimahi, Cirebon, Depok, Sukabumi, dan kota Tasikmalaya (100-5.000 hektar).

#### 5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis menggunakan algoritma *K-Means*, data produktivitas padi di Jawa Barat berhasil dikelompokkan menjadi 18 cluster dengan berbagai tingkat produktivitas. Cluster 0 menunjukkan produktivitas sangat tinggi, diikuti oleh beberapa cluster tinggi dan rendah. Evaluasi menggunakan *David Bouldin Index* (DBI) menemukan nilai k yang optimal adalah 18. Produktivitas padi paling tinggi terdapat pada cluster 0, dengan Kota Bandung sebagai kabupaten yang paling produktif,

terutama pada bulan September. *Cluster* 13 menunjukkan produksi padi tertinggi, didominasi oleh Kabupaten Indramayu, terutama pada bulan April. Luas panen terbesar terdapat pada *cluster* 13, menandakan keterkaitan erat antara jumlah produksi dan luas panen. Saran untuk penelitian selanjutnya mencakup pembandingan dengan algoritma lain, pengembangan hasil klasterisasi sebagai basis pengetahuan, dan pemanfaatan data pengelompokan untuk klasifikasi tingkat produktivitas padi di kabupaten atau kota di Jawa Barat. Penelitian ini dapat menjadi acuan bagi Pemerintah Provinsi Jawa Barat dalam upaya peningkatan produktivitas padi.

#### DAFTAR PUSTAKA

- [1] m. A. R. Siregar, "Peningkatan produktivitas tanaman padi melalui penerapan teknologi pertanian terkini," pp. 1–11, 2023.
- [2] r. P. Harjono and m. A. I. Pakereng, "Penerapan metode k- means clustering untuk analisis potensi lahan pangan pada provinsi kalimantan selatan," vol. 7, pp. 332–338, 2013.
- [3] S. Anisah and L. Wachidah, "Penerapan metode non-hierarchical clustering k-means untuk pengelompokan kabupaten / kota di provinsi jawa barat berdasarkan faktor produksi padi tahun 2021," pp. 323–332, 2021.
- [4] R. R. Khaerullah, N. Suarna, and O. Nurdiawan, "Analisa Pengelompokan Dataset Komputer Menggunakan Algoritma X-Means," *J. Inform. dan Teknol. Inf.*, vol. 1, no. 2, pp. 125–132, 2023, doi: 10.56854/jt.v1i2.135.
- [5] S. Wijayanto and M. yoka fathoni, "Pengelompokan produktivitas tanaman padi di jawa tengah menggunakan metode clustering k-means," *Jupiter*, vol. 13, no. 2, pp. 212–219, 2021.
- [6] a. A. Maulana, a. W. Rafii, y. A. Anjelina, and E. Widodo, "Pengelompokan kecamatan di kabupaten bima berdasarkan jumlah produksi dan luas panen bawang merah tahun 2021 menggunakan k-means clustering," vol. 16, no. 1, pp. 442–451, 2023.
- [7] s. W. Harjono, N. widya utami, I. Gusti, A. Pramesti, and D. Putri, "Klasterisasi tingkat penjualan pada startup panak.id dengan algoritma k-means. Jurnal ilmiah teknologi informasi asia," *J. Ilm. Teknol. Inf. asia*, vol. 17, no. 1, pp. 55–66, 2023.
- [8] O. Nurdiawan, A. Irma Purnamasari, and I. Ali, "Analisa Penjualan Mobil Dengan Menggunakan Algoritma K-Means Di PT. Mulya Putra Kencana," *J. Data Sci. dan Inform.*, vol. 1, no. 2, pp. 32–35, 2021.
- [9] T. Yuniati and m. F. Sidiq, "Legalitas dokumen elektronik menggunakan tanda tangan digital sebagai alternatif pengesahan dokumen di masa pandemi," *J. resti*, vol. 4, no. 10, 2021.
- [10] a. Al Masykur, s. K. Gusti, S. Sanjaya, and F. Yanto, "Penerapan metode k-means clustering untuk pemetaan pengelompokan lahan produksi tandan buah segar," vol. 10, no. 1, 2023.
- [11] M. Walid and F. Halimiyah, "Klasifikasi Kemandirian Siswa SMA/MA Double Track Menggunakan Metode Naive Bayes," *J. ICT Inf. Commun. Technol.*, vol. 22, pp. 190–197, 2022.
- [12] A. Aziz, A. Siregar, and C. Zonyfar, "Penerapan algoritma k-means dan fuzzy c-means untuk pengelompokan kabupaten kota berdasarkan produksi padi di provinsi jawa barat," pp. 1–8, 2022, [Online]. Available: <https://journal.ubpkarawang.ac.id/mahasiswa/index.php/ssj/article/view/411>
- [13] T. Pribadi, R. Irsyada, H. Audytra, and d. A. Fatah, "Implementasi algoritma k-means untuk klasterisasi potensi desa pada sektor produksi pertanian di kabupaten bojonegoro," *J. simantec*, vol. 9, no. 1, pp. 20–28, 2020.
- [14] Y. Pratama, E. Rasywir, B. tio carenina, and D. riski anggraini, "Penerapan algoritma k-means clustering untuk mengelompokkan provinsi berdasarkan banyaknya desa/kelurahan dengan upaya antisipasi/mitigasi bencana alam," *Technol. Sci.*, vol. 4, no. 3, pp. 1232–1240, 2022.
- [15] Putriyana and O. Nurdiawan, "Implementasi Data Mining Untuk Memprediksi Kelulusan Siswa SMK Al Huda Kedungwungu Dengan Menggunakan Algoritma Naïve Bayes Classifier," *Exp. Student Exp.*, vol. 1, no. 1, pp. 1–7, 2023.
- [16] A. Neva and F. Dikananda, "Penerapan metode k-means clustering dalam mengelompokkan jumlah wisatawan asing di jawa barat," p. 2, 2023.
- [17] y. B. Utomo, I. Kurniasari, and I. Yanuartanti, "Penerapan knowledge discovery in database untuk analisa tingkat kecelakaan lalu lintas," *Jtik (jurnal Tek. Inform. kaputama)*, vol. 7, no. 1, pp. 171–180, 2023.