

# PENERAPAN *K-MEANS CLUSTERING* UNTUK PEMETAAN PRODUKTIVITAS PADI DAN PREDIKSI PANEN DI KABUPATEN INDRAMAYU

Riyan Farismana

Program Studi Sistem Informasi Kota Cerdas  
Jurusan Teknik Informatika  
Politeknik Negeri Indramayu

riyanfarismana@polindra.ac.id

**Received:** June 26, 2024. **Revised:** July 31, 2024. **Accepted:** August 2, 2024. **Issue Period:** Vol.8 No.3 (2024), Pp. 589-605

**Abstrak:** Indonesia negara mayoritas masyarakatnya mengkonsumsi nasi, menjadikan isu produktivitas panen padi sebagai perhatian banyak pihak. Upaya menjaga ketahanan pangan khususnya ketersediaan padi merupakan tanggung jawab bersama, dimana upaya yang bisa dilakukan adalah menjaga dan memprediksi hasil produksi beras pada daerah yang menjadi lumbung padi nasional. Indramayu, merupakan salah satu kabupaten penghasil beras nasional, yang memiliki 31 kecamatan dengan jumlah produksi padi yang beragam. Perbedaan produksi padi yang jauh antara kecamatan menjadikan permasalahan penting yang mempengaruhi panen beras di Indramayu, sehingga perlu dilakukan perbaikan diantaranya dengan melakukan *clustering*, yaitu pengelompokan kecamatan berdasarkan potensi panen menggunakan *clustering K-Means* dengan melihat *history* data panen 5 tahun kebelakang antara 2019 - 2023, kemudian dibandingkan dengan menggunakan aplikasi *orange data mining*. *Clustering K-Means* baik secara manual maupun menggunakan *orange3* menghasilkan tiga *cluster* dengan hasil *cluster* kecamatan yang sama. data tersebut juga dimanfaatkan untuk prediksi hasil panen yang akan datang menggunakan algoritma *random forest* pada *orange3*. Analisis hasil *clustering K-Means* dan hasil prediksi menunjukkan kecamatan yang berada pada *cluster* rendah memiliki potensi persentase penurunan panen yang lebih besar dari pada *cluster* dengan kategori panen yang lebih tinggi.

**Kata kunci:** *clustering*, padi, prediksi, indramayu, *orange3*

**Abstract:** Indonesia is a country where the majority of people consume rice, making the issue of rice harvest productivity a concern for many parties. Efforts to maintain food security, especially the availability of rice, are a shared responsibility, where efforts that can be made are to maintain and predict rice production results in areas that are National Rice Granaries. Indramayu is one of the national rice producing districts, which has 31 sub-districts with varying amounts of rice production. The large difference in rice production between sub-districts is an important problem that affects rice harvests in Indramayu, so improvements need to be made, including clustering, namely grouping sub-districts based on harvest potential using *K-Means clustering* by looking at historical harvest data for the past 5 years between 2019–2023. Then compared using the *orange data mining* application. *K-Means clustering* either manually or using *orange3* produces three clusters with the same sub-district cluster results. This data is also used to predict future harvest results using the



DOI: 10.52362/jisamar.v8i3.1572

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

*random forest algorithm in Orange3. Analysis of K-Means clustering results and prediction results shows that sub-districts in the lower cluster have the potential for a greater percentage reduction in harvest than clusters with a higher harvest category.*

**Keywords:** clustering, rice, prediction, indramayu, orange3

## I. PENDAHULUAN

Indonesia merupakan negara agraris dengan potensi pertanian yang besar dan berperan penting dalam mencapai swasembada pangan serta meningkatkan perekonomian Indonesia. Swasembada pangan sendiri sering dikaitkan dengan hasil produksi padi. Hal tersebut terlihat dimana Indonesia menyumbang 9% produksi padi dunia [1]. Padi yang diolah menjadi beras juga merupakan makanan pokok masyarakat dimana rata-rata konsumsi beras tahun 2019 adalah 87,03% [2]. Selain itu faktor-faktor seperti luas lahan, iklim, dan pola tanam juga mempengaruhi jumlah produksi padi di Indonesia [3].

Produksi padi merupakan topik yang selalu menjadi perhatian banyak pihak, hal ini terjadi karena tinggi atau rendahnya produksi padi menjadi tolak ukur ketahanan pangan di Indonesia. Ketahanan pangan juga merupakan isu yang selalu menjadi prioritas pembangunan nasional. Berkurangnya produksi padi akan menjadikan krisis ketahanan pangan yang berdampak pada ketergantungan impor beras dan ekonomi nasional [4]. Untuk menanggulangi permasalahan tersebut salah satu usaha yang bisa dilakukan adalah dengan menjaga dan menambah produksi beras dalam negeri dengan memperhatikan daerah-daerah penghasil beras nasional.

Salah satu penghasil padi terbesar di Indonesia adalah Kabupaten Indramayu. Luas panen tanaman padi di Indramayu dalam satuan (ha/Hektar) tahun 2023 seluas 231354 Ha [5]. Indramayu juga menempati posisi teratas sebagai lumbung padi baik di tingkat provinsi maupun Nasional. Cakupan wilayah Indramayu saat ini memiliki 31 kecamatan dengan jumlah hasil produksi padi yang bervariasi dimana pada tahun 2020 panen padi tertinggi adalah kecamatan gantar menghasilkan 163,397 Ton padi, dan panen terendah di kecamatan pasekan dengan 12,247 Ton [6].

Variasi dan disparitas jumlah produksi padi antar kecamatan di Indramayu merupakan permasalahan kompleks yang mempengaruhi produksi padi secara keseluruhan di Indramayu. Beberapa kecamatan memiliki hasil panen yang tinggi dan konsisten, sementara kecamatan lainnya menghadapi tantangan dalam meningkatkan produktivitas panen padi. Untuk mengatasi permasalahan tersebut perlu dilakukan upaya-upaya perbaikan diantaranya dengan melakukan pengelompokan potensi padi untuk mendukung keputusan pemerintah daerah dalam menentukan daerah-daerah yang menjadi sentra padi [7]. Selain itu pengelompokan juga berguna untuk memetakan kecamatan mana saja yang perlu ditingkatkan produksi padi agar ketahanan pangan terjaga.

*Clustering* merupakan salah satu model dalam teknik *data mining* yang merupakan inti dari *knowledge discovery in database (KDD)* [8]. *K-Means clustering* merupakan metode pengelompokan data ke dalam *cluster* berdasarkan variabel, dengan tujuan utama untuk meminimalkan jarak antara titik data dan *cluster* terkait [9]. Teknik lain yang bisa digunakan dengan memanfaatkan variasi dari data adalah dengan membuat prediksi hasil panen padi dimasa yang akan datang. Prediksi adalah strategi yang tepat untuk mengetahui hasil pertanian dengan mempertimbangkan berbagai macam faktor [10].

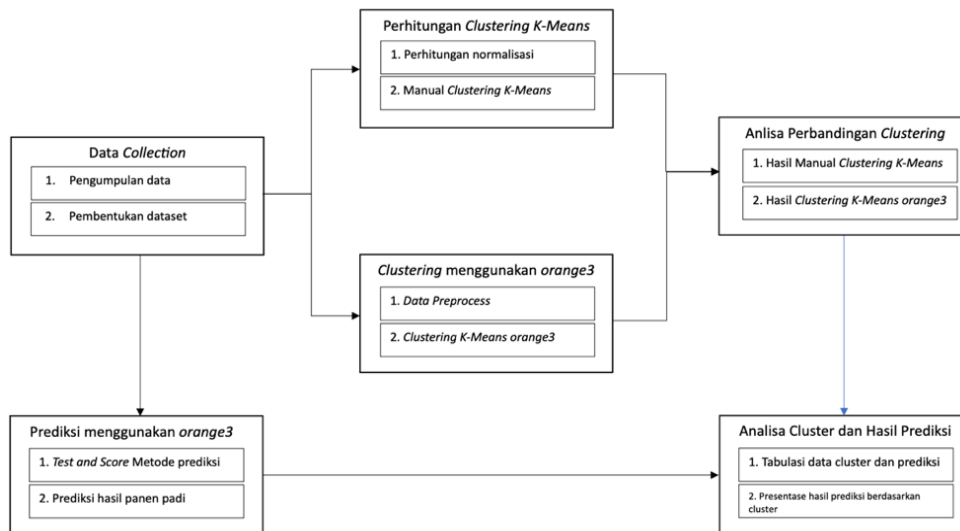
Berdasarkan penjelasan tersebut, penelitian ini dimaksudkan untuk menggali data historis hasil panen padi di Kabupaten Indramayu berdasarkan kecamatan selama 5 tahun kebelakang, kemudian mengelompokkan kecamatan kedalam beberapa cluster tinggi, sedang, dan rendah menggunakan perhitungan manual *clustering K-Means*. Data *cluster* yang diperoleh dari perhitungan *clustering K-Means* tersebut kemudian dilakukan perbandingan perhitungan *K-Means* pada aplikasi *orange data mining* untuk melihat kesamaan hasil *clustering*, dan setelah itu data panen tersebut dianalisis untuk selanjutnya digunakan sebagai dataset pada proses peramalan atau prediksi menggunakan aplikasi *orange data mining* berdasarkan nilai error terkecil.

Diharapkan dengan membagi *cluster* pada masing-masing kecamatan yang ada di Indramayu, serta mengetahui prediksi hasil panen yang akan datang pada setiap kecamatan tersebut, dapat dijadikan landasan bagi pemerintah daerah Kabupaten Indramayu khususnya Dinas Pertanian dan Ketahanan Pangan untuk membuat berbagai kebijakan yang dapat dijadikan landasan untuk mempertahankan ketahanan pangan khususnya di Indramayu dan Umumnya secara Nasional dikarenakan Kabupaten Indramayu merupakan lumbung padi Nasional.



## II. METODE DAN MATERI

Pada penelitian ini memiliki beberapa Langkah yang secara garis besar terdiri dari *data collection*, perhitungan *clustering* menggunakan metode *K-Means*, dan prediksi data produksi padi pada tiap-tiap kecamatan. Adapun alur penelitian dapat dijelaskan sebagai berikut.



Gambar 1. Desain Penelitian

### 2.1. Data Collection

Keberhasilan suatu penelitian tidak bisa dilepaskan dari pengumpulan data sebagai sumber penelitian [11]. Data yang dikumpulkan dalam penelitian ini merupakan data kuantitatif dimana mencari data hasil panen padi pada kecamatan di Indramayu dengan kurun waktu 5 tahun terakhir.

#### 1. Pengumpulan data panen

Pengumpulan data panen padi diperoleh dari data yang terdapat pada *open data* Indramayu (<https://opendata.indramayukab.go.id/>) khususnya pada topik ekonomi, dimana terdapat data jumlah produksi padi berdasarkan kecamatan. Selain itu untuk validasi data yang lebih akurat, penelitian ini juga didasarkan pada dokumen yang terdapat pada Dinas Pertanian dan Ketahanan Pangan Indramayu.

#### 2. Pembentukan *dataset*

*Dataset* merupakan sekumpulan data yang dapat dijadikan bahan di dalam sebuah penelitian [12]. Dari data yang diperoleh pada tahap pengumpulan data panen terutama data yang didapat dari Dinas Pertanian dan Ketahanan Pangan masih beragam dan belum rapih, sehingga perlu dilakukan pembentukan *dataset* baru yang berisi data kecamatan dan hasil panen dari tahun 2019 – 2023.

### 2.2. Perhitungan *Clustering K-means*

Setelah proses pembentukan dataset baru selesai. Tahapan selanjutnya adalah melakukan perhitungan *clustering K-Means* secara manual. Perhitungan ini memiliki dua tahapan utama yaitu perhitungan normalisasi dan manual *clustering K-Means*. Hasil dari tahapan ini adalah pembagian kecamatan berdasarkan *cluster* tinggi, sedang dan rendah.

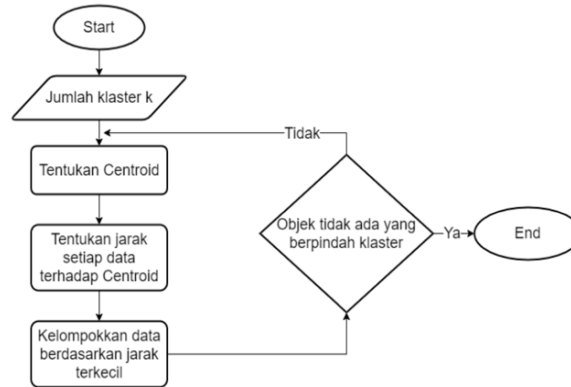
#### 1. Perhitungan normalisasi

*K-Means* merupakan metode yang *sensitif* terhadap data *outlier*, sehingga perlu dilakukan normalisasi data pada suatu *dataset* untuk membentuk data tersebut ke dalam rentang posisi nilai yang sama [13]. untuk memudahkan perhitungan data normalisasi, sebelumnya dilakukan pencarian nilai minimum dan maksimum data panen padi pada tiap kecamatan berdasarkan tahun 2019 – 2023.



## 2. Manual clustering K-Means

Algoritma *K-means* beroperasi sebagai algoritma partisi, karena bergantung pada penetapan jumlah awal grup dengan menentukan nilai *centroid* awal. Melalui proses berulang, algoritma *K-means* secara efektif menghasilkan *database cluster* [14]. Berikut merupakan diagram alir dari pembentukan *clustering K-Means*.



Gambar 2. Diagram alir algoritma *K-Means* [15]

## 2.3. Clustering Menggunakan Orange3

*Orange data mining* adalah platform berbasis komponen dengan pendekatan komprehensif berdasarkan pembelajaran mesin dan penambahan data. *Orange* digunakan dalam sains, industri dan Pendidikan. *Orange* memberikan gambaran dengan visualisasi data, klasifikasi, evaluasi, *unsupervised learning*, hingga implementasi *prototype* [16]. Sebelum data di *Clustering* menggunakan algoritma *K-Means* dalam penelitian ini dilakukan *data preprocess*.

### 1. Data preprocess

Tahapan ini dilakukan untuk menormalkan data yang terdapat dalam dataset hasil panen padi, dengan *normalize to interval* [0,1]. Hal ini dilakukan untuk akurasi proses *clustering* menggunakan *K-Means*.

### 2. K-Means clustering orange3

Setelah data di *praprocess* selanjutnya dilakukan proses *clustering* dengan menghubungkan widget *K-Means* dan *praprocess* pada aplikasi *orange3*.

## 2.4. Analisa Perbandingan Clustering

Analisis perbandingan dilakukan untuk melihat apakah hasil clustering menggunakan perhitungan manual *K-Means* dan perhitungan dengan bantuan aplikasi *orange3* terdapat perbedaan atau kesamaan hasil *cluster*. Dalam proses ini data hasil *clustering* dari masing-masing perhitungan disajikan dan dianalisis hasilnya.

## 2.5. Prediksi Menggunakan Orange3

Tindakan prediksi melibatkan perkiraan kejadian di masa depan secara sistematis dengan memanfaatkan informasi masa lalu dan masa kini, dengan tujuan meminimalkan perbedaan antara hasil aktual dan hasil perkiraan[17]. Untuk melakukan prediksi menggunakan aplikasi *orange3*, pada penelitian ini tahapannya terdiri dari.

### 1. Test and score metode prediksi

Tahapan ini dilakukan untuk mencari nilai error terkecil dari beberapa algoritma prediksi yang disediakan didalam aplikasi *orange3*. Hasil penghitungan *error* dilihat dari beberapa metode yaitu *Mean Absolute Error (MAE)*, *Mean Square Error (MSE)*, *Mean Absolute Percentage Error (MAPE)*, dan *Root Mean Squared Error (RMSE)*. Dimana Semakin kecil tingkat *error* maka semakin baik prediksi yang dilakukan [18].

### 2. Dari hasil test and score, metode prediksi yang memiliki nilai error terkecil yang akan digunakan dalam melakukan prediksi hasil panen.



### 2.6. Analisa Cluster dan Hasil Prediksi

Tahapan ini digunakan untuk melihat berapa banyak kecamatan yang mengalami penurunan dan peningkatan hasil panen berdasarkan prediksi yang telah dilakukan berdasarkan pembagian *cluster* yang telah dilakukan sebelumnya, dengan membuat tabulasi data dari hasil *cluster* dan prediksi. Kemudian dibuat persentase hasil prediksi berdasarkan *cluster* untuk menarik kesimpulan berapa banyak kecamatan yang mengalami penurunan sesuai dengan *cluster* yang telah dihitung.

## III. PEMBAHASA DAN HASIL

### 3.1. Data Collection dan Pembentukan Dataset.

Data yang diperoleh dari Dinas Pertanian dan Ketahanan Pangan Indramayu, secara rinci masa panen selama satu tahun dibagi menjadi dua periode mengikuti periode masa tanam (MT). dari dua periode tersebut Dinas Pertanian dan Ketahanan Pangan pada akhir MT akan menggabungkan hasil panen periode pertama dan kedua MT menjadi jumlah panen di tahun tersebut dalam satuan Ton/GKB atau Gabah Kering Giling. Selain dari Dinas Pertanian dan Ketahanan Pangan, data yang diambil untuk melihat konsistensi, data juga diperoleh dari <https://opendata.indramayukab.go.id/>. Data dari situs ini sudah merupakan data akhir panen dalam satu tahun, akan tetapi *update* data masih sampai tahun 2022, sehingga untuk tahun 2023 masih menggunakan dari Dinas Pertanian dan Ketahanan Pangan Kab. Indramayu.

Berdasarkan kedua sumber data tersebut, pembentukan *dataset* dimulai dari mengelompokkan jumlah hasil panen akhir tahun yang dikategorikan berdasarkan kecamatan dalam kurun waktu historis lima tahun dari 2019 sampai dengan 2023 dalam satuan Ton/GKB.

Tabel 1. *Dataset* hasil panen padi Kab. Indramayu

KECAMATAN	2023	2022	2021	2000	1999
Haurgeulis	57,47	60,138	54,772	69,076	83,889
Gantar	112	143,691	126,485	163,397	148,5
Kroya	76,685	102,356	138,447	119,239	113,705
Gabuswetan	76,755	93,844	83,667	85,632	74,946
Cikedung	75,576	94,104	94,318	95,942	93,477
Terisi	133,924	111,162	114,9	90,23	90,65
Lelea	60,293	73,711	75,493	77,065	75,512
Bangodua	48,286	44,719	46,492	47,409	53,654
Tukdana	53,851	55,174	56,861	59,164	65,313
Widasari	44,203	44,251	37,902	42,291	43,583
Kertasemaya	45,855	46,875	42,871	50,552	51,163
Sukagumiwang	37	38,805	37,664	38,903	39,754
Krangkeng	57,971	65,618	61,337	55,942	60,951
Karangampel	32,486	29,994	32,182	32,487	31,376
Kedokan Bunder	32,14	31,997	32,977	34,059	32,67
Juntinyuat	63,763	64,468	62,081	61,325	63,44
Sliyeg	64,843	62,716	63,655	66,016	69,906
Jatibarang	45,086	43,474	43,667	43,408	47,034
Balongan	26,465	27,012	28,193	28,588	28,049
Indramayu	26,143	28,275	24,379	24,488	24,274



Sindang	33,28	23,538	29,845	31,705	27,753
Cantigi	21,341	29,178	27,305	27,817	23,28
Pasekan	13,235	12,7	12,183	12,247	12,802
Lohbener	33,126	37,688	37,344	38,895	37,208
Arahan	28,381	36,16	36,903	36,136	35,023
Losarang	60,92	72,475	63,832	74,987	58,718
Kandanghaur	66,244	83,413	62,225	68,544	52,567
Bongas	55,561	54,627	58,064	55,413	60,241
Anjatan	88,775	85,4	86,795	87,796	97,374
Sukra	50,214	50,782	50,936	51,36	48,973
Patrol	45,835	48,097	44,466	46,827	44,786

### 3.2. Perhitungan Clustering K-Means

#### 1. Normalisasi Data

Sebelum melakukan normalisasi, tahap pertama yang dilakukan adalah mencari nilai minimal dan maksimal dari hasil panen di setiap kecamatan pada masing-masing periode waktu 2019 sampai 2023

Tabel 2. Nilai minimal dan maksimal hasil panen

2023		2022		2021		2000		1999	
MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX
13,235	133,924	12,7	143,691	12,183	138,447	12,247	163,397	12,802	148,5

Nilai minimal dan maksimal yang telah diketahui, digunakan untuk melakukan normalisasi data. Perbedaan hasil panen yang jauh antar kecamatan menyebabkan normalisasi data hasil panen dilakukan. Untuk menghitung nilai normalisasi menggunakan persamaan [19].

$$\text{Nilai Normalisasi} = \frac{(\text{Nilai Awal} - \text{Nilai Minimal})}{(\text{Nilai Maksimal} - \text{Nilai Minimal})} \quad (1)$$

Sebagai contoh normalisasi data untuk Kecamatan Haurgeulis pada tahun 2023 memiliki hasil panen sebesar 57,47 Ton, sehingga normalisasi yang dilakukan adalah Nilai Normalisasi =  $(57,47 - 13,235) / (133,924 - 13,235)$ . Sehingga hasilnya adalah 0,366537133.

Tabel 3. Nilai normalisasi data hasil panen

Kecamatan	2023	2022	2021	2000	1999
Haurgeulis	0,366537133	0,362147018	0,33730121	0,375977506	0,523861811
Gantar	0,818235299	1	0,905261991	1	1
Kroya	0,525731425	0,684443969	1	0,707853126	0,74358502
Gabuswetan	0,526311429	0,619462406	0,56614712	0,485511082	0,457958113
Cikedung	0,516542518	0,621447275	0,650502123	0,553721469	0,594518711
Terisi	1	0,751669962	0,813509789	0,515931194	0,573685684
Lelea	0,38991126	0,465764824	0,501409745	0,428832286	0,46212914
Bangodua	0,290424148	0,244436641	0,271724324	0,232629838	0,301050863



Tukdana	0,336534398	0,324251284	0,35384591	0,310400265	0,386969594
Widasari	0,256593393	0,240863876	0,203692264	0,198769434	0,226834589
Kertasemaya	0,270281467	0,260895787	0,243046316	0,253423751	0,282693923
Sukagumiwang	0,194922487	0,199288501	0,201807324	0,176354615	0,198617518
Krangkeng	0,370671727	0,403981953	0,389295444	0,289083692	0,354824684
Karangampel	0,159509152	0,132024338	0,158390357	0,133906715	0,136877478
Kedokan Bunder	0,156642279	0,147315464	0,164686688	0,14430698	0,146413359
Juntinyuat	0,418662844	0,395202724	0,39518786	0,324697321	0,373166885
Sliyeg	0,427611464	0,381827759	0,407653805	0,355732716	0,420816814
Jatibarang	0,263909718	0,234932171	0,249350567	0,206159444	0,252266061
Balongan	0,109620595	0,109259415	0,12679782	0,108111148	0,112359799
Indramayu	0,106952581	0,118901299	0,096591269	0,080985776	0,084540671
Sindang	0,166088044	0,082738509	0,139881518	0,128733047	0,110178485
Cantigi	0,067164365	0,125794902	0,119764937	0,103010255	0,077215582
Pasekan	0	0	0	0	0
Lohbener	0,164812038	0,190761197	0,199272952	0,176301687	0,179855267
Arahan	0,12549611	0,179096274	0,19578027	0,158048296	0,163753335
Losarang	0,395106431	0,456329061	0,409055629	0,415084353	0,338369025
Kandanghaur	0,439219813	0,539830981	0,396328328	0,372457823	0,293040428
Bongas	0,350703047	0,320075425	0,363373566	0,285583857	0,349592477
Anjatan	0,625906255	0,554999962	0,590920611	0,499827985	0,623236894
Sukra	0,306399092	0,290722263	0,306920421	0,258769434	0,266555145
Patrol	0,270115752	0,270224672	0,255678578	0,228779358	0,235699863

Variabel hasil panen per tahun akan dinormalisasi antara rentang 0 sampai dengan 1. Proses normalisasi data ini digunakan sebelum melakukan perhitungan nilai *centroid*.

## 2. Manual Clustering K-Means

Jumlah *cluster* yang akan dibentuk dalam penelitian ini sebanyak tiga cluster berdasarkan status produksi tinggi, sedang dan rendah. Untuk penentuan initial cluster center dapat menggunakan metode acak atau *random* [20]. Status produksi tinggi nilai yang diberikan adalah 1 pada masing-masing tahun, status produksi sedang nilai didapat dari kecamatan Sliyeg, dan status produksi rendah dari kecamatan Lohbener.

Tabel 4. *Initial cluster center* iterasi 1

Cluster	Status Produksi	2023	2022	2021	2000	2019
3	Tinggi	1	1	1	1	1
2	Sedang	0,43	0,38	0,41	0,36	0,42
1	Rendah	0,16	0,19	0,19	0,17	0,17



Untuk menentukan jarak antara data pada setiap *centroid*, Langkah yang dilakukan adalah dengan menghitung data hasil normalisasi dengan titik *centroid* yang sudah ditentukan. Untuk menghitungnya Rumus *Euclidean Distance* [8].

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \quad (2)$$

Dimana  $D_e$  adalah *Euclidean Distance*,  $i$  mempresentasikan banyaknya data atau objek,  $(x,y)$  merupakan koordinat objek, dan  $(s,t)$  koordinat *centroid*. Dari persamaan tersebut, dilakukan perhitungan iterasi 1 untuk mencari data pada tiap *cluster*, dan menentukan jarak terkecil yang akan mempresentasikan *cluster* tersebut. Berdasarkan perhitungan didapat hasil.

Tabel 5. Hasil iterasi 1

Kecamatan	Jarak C3	Jarak C2	Jarak C1
Haurgeulis	1,37	0,14	0,51
Gantar	0,20	1,24	1,73
Kroya	0,69	0,82	1,29
Gabuswetan	1,06	0,33	0,80
Cikedung	0,93	0,44	0,92
Terisi	0,72	0,82	1,30
Lelea	1,23	0,16	0,61
Bangodua	1,64	0,29	0,22
Tukdana	1,47	0,13	0,38
Widasari	1,73	0,39	0,13
Kertasemaya	1,65	0,31	0,20
Sukagumiwang	1,80	0,46	0,05
Krangkeng	1,43	0,11	0,42
Karangampel	1,91	0,57	0,08
Kedokan Bunder	1,90	0,55	0,06
Juntinyuat	1,39	0,06	0,47
Sliyeg	1,35	0,00	0,50
Jatibarang	1,70	0,35	0,16
Balongan	1,98	0,64	0,14
Indramayu	2,02	0,68	0,18
Sindang	1,96	0,61	0,14
Cantigi	2,02	0,68	0,18
Pasekan	2,24	0,89	0,39
Lohbener	1,83	0,49	0,02
Arahan	1,87	0,53	0,04
Losarang	1,34	0,13	0,51
Kandanghaur	1,34	0,20	0,55
Bongas	1,49	0,15	0,36
Anjatan	0,95	0,41	0,91





<i>Sukra</i>	1,60	0,26	0,25
<i>Patrol</i>	1,67	0,33	0,18

Setelah didapat data kecamatan yang memiliki jarak terdekat pada masing-masing *cluster*, selanjutnya pada algoritma *K-Means* dilakukan penentuan atau *update* nilai *centroid* baru untuk melakukan iterasi ke-2. Penentuan nilai *centroid* baru menggunakan nilai rata-rata atau *Mean* nilai normalisasi panen pada masing-masing *cluster* yang telah terbentuk . Sebagai contoh untuk nilai *cluster center* pada C3 tahun 2023 diperoleh dengan menghitung rata-rata dari *cluster C1* yaitu kecamatan Gantar, Kroya, dan Terisi tahun 2023 dengan nilai (0,818235299),( 0,525731425), (1) = (0,78).

Tabel 6. *Initial cluster center* iterasi 2

Cluster	Status Produksi	2023	2022	2021	2000	2019
3	Tinggi	0,78	0,81	0,91	0,74	0,77
2	Sedang	0,43	0,45	0,45	0,39	0,43
1	Rendah	0,18	0,18	0,18	0,16	0,17

Hasil pembentukan *centroid* baru tersebut, dilakukan perhitungan iterasi ke-2 dengan metode dan rumus yang sama yaitu *Euclidean Distance*, sehingga hasilnya adalah.

Tabel 7. Hasil iterasi 2

Kecamatan	Jarak C3	Jarak C2	Jarak C1
Hargeulis	0,95	0,18	0,51
Gantar	0,39	1,16	1,73
Kroya	0,30	0,75	1,29
Gabuswetan	0,62	0,25	0,80
Cikedung	0,49	0,36	0,93
Terisi	0,39	0,76	1,30
Lelea	0,79	0,08	0,62
Bangodua	1,20	0,37	0,21
Tukdana	1,04	0,21	0,38
Widasari	1,30	0,46	0,12
Kertasemaya	1,22	0,38	0,20
Sukagumiwang	1,37	0,53	0,04
Krangkeng	0,99	0,16	0,42
Karangampel	1,48	0,64	0,07
Kedokan Bunder	1,46	0,63	0,05
Juntinyuat	0,95	0,12	0,46
Sliyeg	0,91	0,09	0,50
Jatibarang	1,26	0,43	0,15
Balongan	1,55	0,71	0,14
Indramayu	1,58	0,75	0,18

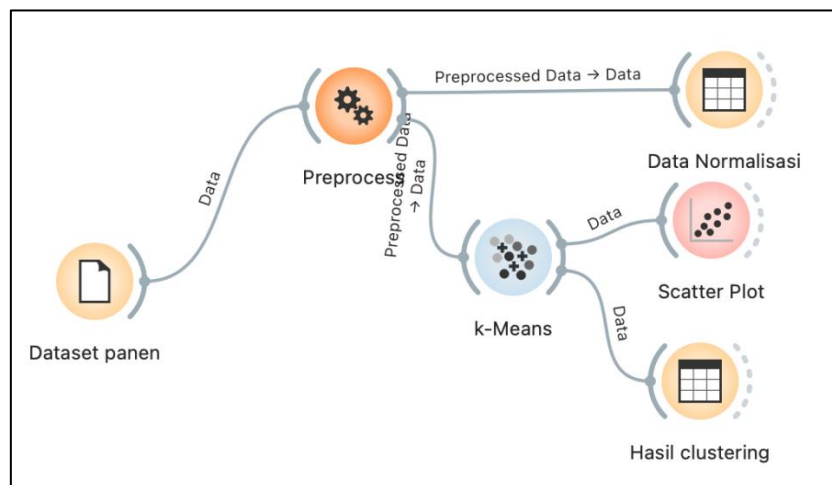


Sindang	1,52	0,69	0,13
Cantigi	1,58	0,75	0,18
Pasekan	1,80	0,96	0,39
Lohbener	1,39	0,56	0,03
Arahan	1,43	0,60	0,06
Losarang	0,90	0,11	0,52
Kandanghaur	0,90	0,17	0,55
Bongas	1,05	0,22	0,36
Anjatan	0,52	0,34	0,91
Sukra	1,16	0,33	0,25
Patrol	1,24	0,40	0,17

Dari hasil perhitungan didapat pembagian cluster C3, C2, dan C1 yang sama sesuai baik perhitungan iterasi 1 dan iterasi 2. Hal tersebut mengindikasikan bahwa perhitungan iterasi *cluster* sudah selesai dikarenakan data yang dihasilkan telah sesuai karena objek tidak ada yang berpindah *cluster*.

### 3.3. Clustering K-Means Menggunakan Orange3

Perhitungan *clustering K-Means* menggunakan aplikasi *orange3* bertujuan untuk melihat perbandingan hasil perhitungan manual dan menggunakan aplikasi *orange3*. Berikut desain *clustering K-Means* pada *orange3*.



Gambar 3. Desain proses *clustering* menggunakan *orange3*

Untuk mendapatkan hasil yang akurat, proses *clustering K-Means* menggunakan *orange3* mengikuti Langkah-langkah perhitungan manual yaitu.

#### 1. *Preprocess data*

Tahapan ini adalah proses pembentukan data normalisasi yang sama dengan perhitungan manual, dimana data di normalisasi pada interval 0 – 1. Nilai hasil normalisasi dapat dilihat pada *widjet* data normalisasi yang terhubung dengan *preprocess*.



Data normalisasi						
	Kecamatan	2023	2022	2021	2000	1999
1	Haurgeulis	0.366537	0.362147	0.337301	0.375978	0.523862
2	Gantar	0.818235	1	0.905262	1	1
3	Kroya	0.525731	0.684444	1	0.707853	0.743585
4	Gabuswetan	0.526311	0.619462	0.566147	0.485511	0.457958
5	Cikedung	0.516543	0.621447	0.650502	0.553721	0.594519
6	Terisi	1	0.75167	0.81351	0.515931	0.573686
7	Lelea	0.389911	0.465765	0.50141	0.428832	0.462129
8	Bangodua	0.290424	0.244437	0.271724	0.23263	0.301051
9	Tukdana	0.336534	0.324251	0.353846	0.3104	0.38697
10	Widasari	0.256593	0.240864	0.203692	0.198769	0.226835
11	Kertasemaya	0.270281	0.260896	0.243046	0.253424	0.282694
12	Sukagumiwa...	0.194922	0.199289	0.201807	0.176355	0.198618
13	Krangkeng	0.370672	0.403982	0.389295	0.289084	0.354825
14	Karangampel	0.159509	0.132024	0.15839	0.133907	0.136877
15	Kedokan Bu...	0.156642	0.147315	0.164687	0.144307	0.146413
16	Juntinyuat	0.418663	0.395203	0.395188	0.324697	0.373167
17	Sliyeg	0.427611	0.381828	0.407654	0.355733	0.420817
18	Jatibarang	0.26391	0.234932	0.249351	0.206159	0.252266
19	Balongan	0.109621	0.109259	0.126798	0.108111	0.11236
20	Indramayu	0.106953	0.118901	0.0965913	0.0809858	0.0845407
21	Sindang	0.166088	0.0827385	0.139882	0.128733	0.110178
22	Cantigi	0.0671644	0.125795	0.119765	0.10301	0.0772156
23	Pasekan	0	0	0	0	0
24	Lohbener	0.164812	0.190761	0.199273	0.176302	0.179855
25	Arahan	0.125496	0.179096	0.19578	0.158048	0.163753
26	Losarang	0.395106	0.456329	0.409056	0.415084	0.338369
27	Kandanghaur	0.43922	0.539831	0.396328	0.372458	0.29304
28	Bongas	0.350703	0.320075	0.363374	0.285584	0.349592
29	Anjatan	0.625906	0.555	0.590921	0.499828	0.623237
30	Sukra	0.306399	0.290722	0.30692	0.258769	0.266555
31	Patrol	0.270116	0.270225	0.255679	0.228779	0.2357

Gambar 4. Hasil normalisasi data menggunakan *orange3*

Pada hasil normalisasi menggunakan *orange3*, menunjukkan nilai normalisasi yang didapat sama dengan hasil perhitungan normalisasi secara manual.

## 2. K-Means Clustering Orange3

Data yang diperoleh dari nilai hasil *preprocess* selanjutnya digunakan untuk menghitung *clustering* menggunakan *widget K-means*, dimana sebelumnya ditentukan terlebih dahulu sebanyak 3 *cluster* yang akan dibentuk. Untuk melihat hasil *clustering* dapat menggunakan *widget data table* hasil *clustering*.



Hasil Clustering							
Kecamatan	Cluster	Silhouette	2023	2022	2021	2000	1999
Gantar	C3	0.629176	0.818235	1	0.905262	1	1
Kroya	C3	0.564178	0.525731	0.684444	1	0.707853	0.743585
Terisi	C3	0.544019	1	0.75167	0.81351	0.515931	0.573686
Haurgeulis	C2	0.645169	0.366537	0.362147	0.337301	0.375978	0.523862
Gabuswetan	C2	0.656573	0.526311	0.619462	0.566147	0.485511	0.457958
Cikedung	C2	0.592529	0.516543	0.621447	0.650502	0.553721	0.594519
Lelea	C2	0.678672	0.389911	0.465765	0.50141	0.428832	0.462129
Tukdana	C2	0.601291	0.336534	0.324251	0.353846	0.3104	0.38697
Krangkeng	C2	0.635273	0.370672	0.403982	0.389295	0.289084	0.354825
Juntinyuat	C2	0.659967	0.418663	0.395203	0.395188	0.324697	0.373167
Sliyeg	C2	0.670211	0.427611	0.381828	0.407654	0.355733	0.420817
Losarang	C2	0.665794	0.395106	0.456329	0.409056	0.415084	0.338369
Kandanghaur	C2	0.652907	0.43922	0.539831	0.396328	0.372458	0.29304
Bongas	C2	0.579692	0.350703	0.320075	0.363374	0.285584	0.349592
Anjatan	C2	0.603464	0.625906	0.555	0.590921	0.499828	0.623237
Bangodua	C1	0.61119	0.290424	0.244437	0.271724	0.23263	0.301051
Widasari	C1	0.672981	0.256593	0.240864	0.203692	0.198769	0.226835
Kertasemaya	C1	0.624204	0.270281	0.260896	0.243046	0.253424	0.282694
Sukagumiwa...	C1	0.694301	0.194922	0.199289	0.201807	0.176355	0.198618
Karangampel	C1	0.703005	0.159509	0.132024	0.15839	0.133907	0.136877
Kedokan Bu...	C1	0.70336	0.156642	0.147315	0.164687	0.144307	0.146413
Jatibarang	C1	0.659428	0.26391	0.234932	0.249351	0.206159	0.252266
Balongan	C1	0.699067	0.109621	0.109259	0.126798	0.108111	0.11236
Indramayu	C1	0.694506	0.106953	0.118901	0.0965913	0.0809858	0.0845407
Sindang	C1	0.697781	0.166088	0.0827385	0.139882	0.128733	0.110178
Cantigi	C1	0.693135	0.0671644	0.125795	0.119765	0.10301	0.0772156
Pasekan	C1	0.663839	0	0	0	0	0
Lohbener	C1	0.697917	0.164812	0.190761	0.199273	0.176302	0.179855
Arahan	C1	0.699054	0.125496	0.179096	0.19578	0.158048	0.163753
Sukra	C1	0.565056	0.306399	0.290722	0.30692	0.258769	0.266555
Patrol	C1	0.644367	0.270116	0.270225	0.255679	0.228779	0.2357

Gambar 5. Hasil clustering K-Means menggunakan orange3

### 3.4. Analisis Perbandingan Clustering

Perbandingan antara hasil clustering menggunakan perhitungan manual dan menggunakan aplikasi orange3, data pada cluster C3, C2, dan C1 memiliki kesamaan baik berdasarkan kecamatan dan cluster yang membagninya.

Tabel 8. Perbandingan clustering manual dan orange3

No	Clustering Manual			Clustering Orange3		
	C3	C2	C1	C3	C2	C1
1	Gantar	Haurgeulis	Bangodua	Gantar	Haurgeulis	Bangodua
2	Kroya	Gabuswetan	Widasari	Kroya	Gabuswetan	Widasari
3	Terisi	Cikedung	Kertasemaya	Terisi	Cikedung	Kertasemaya
4		Lelea	Sukagumiwang		Lelea	Sukagumiwang
5		Tukdana	Karangampel		Tukdana	Karangampel
6		Krangkeng	Kedokan Bunder		Krangkeng	Kedokan Bunder
7		Juntinyuat	Jatibarang		Juntinyuat	Jatibarang
8		Sliyeg	Balongan		Sliyeg	Balongan
9		Losarang	Indramayu		Losarang	Indramayu
10		Kandanghaur	Sindang		Kandanghaur	Sindang
11		Bongas	Cantigi		Bongas	Cantigi
12		Anjatan	Pasekan		Anjatan	Pasekan



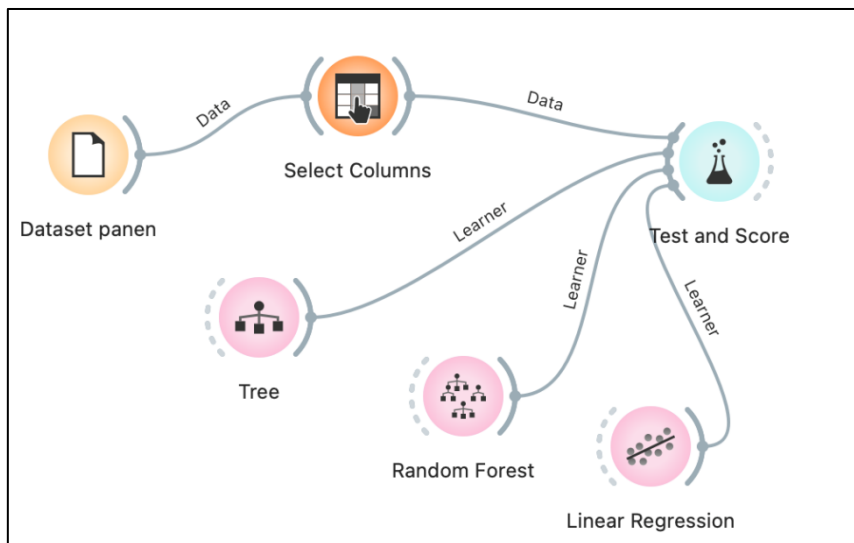
DOI: 10.52362/jisamar.v8i3.1572

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).

13	Lohbener	Lohbener
14	Arahan	Arahan
15	Sukra	Sukra
16	Patrol	Patrol

### 3.5. Test and Score Metode Prediksi

Pada penelitian ini untuk melihat nilai error terkecil dari algoritma-algoritma prediksi yang disediakan aplikasi orange3, dimana algoritma yang di *test and score* adalah *random forest*, *tree*, dan *linear regression*. Model algoritma yang memiliki nilai *error* terkecil yang akan digunakan untuk memprediksi hasil panen padi.



Gambar 6. Desain model *test and score* algoritma untuk prediksi panen padi

*Test and score* dilakukan menggunakan dataset panen padi, dimana *features data* dari tahun 1999 sampai tahun 2022, dan tahun 2023 adalah *target* prediksi. Hasil *test and score* menunjukkan algoritma *random forest* memiliki *error* terendah dibandingkan *tree* dan *linear regression*.

Test and Score				
Model	MSE	RMSE	MAE	MAPE
Tree	239.376	15.472	8.448	0.151
Random Forest	129.186	11.366	6.298	0.126
Linear Regression	158.403	12.586	7.547	0.133

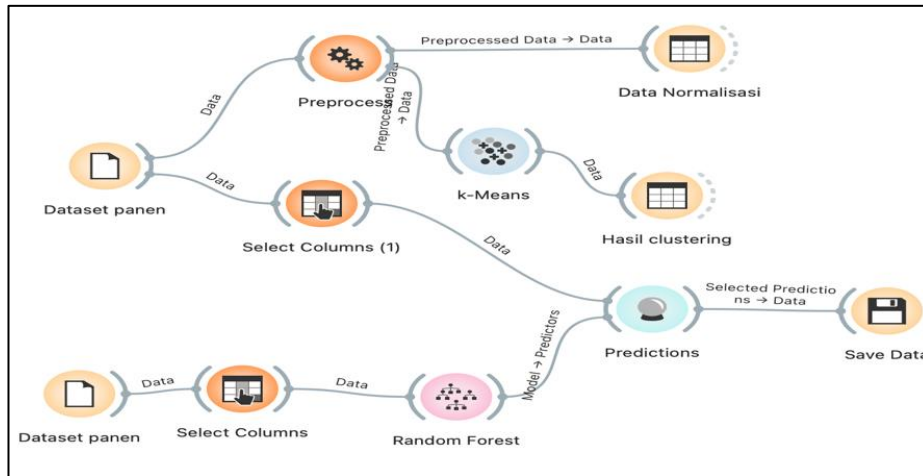
Gambar 7. Hasil *test and score*

Berdasarkan pengujian *test and score*, ditentukan bahwa model *random forest* yang akan digunakan sebagai model untuk prediksi panen padi berdasarkan kecamatan di Indramayu, dimana tahun 1999 sampai 2022 sebagai *features* dan 2023 sebagai *target*. desain model pada aplikasi *orange3* seperti berikut.



DOI: 10.52362/jisamar.v8i3.1572

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



Gambar 8. Desain model prediksi hasil panen padi

Hasil prediksi yang dilakukan menggunakan algoritma *random forest* pada *dataset* hasil panen padi. Data hasil prediksi dapat dilihat pada *widget predictions*.

Predictions								
Shown regression error: Difference								
	Random Forest	error	2023	Kecamatan	2022	2021	2000	1999
1	61.630	4.158	57.472	Haurgeulis	60.138	54.772	69.076	83.889
2	119.534	7.547	111.987	Gantar	143.691	126.485	163.397	148.500
3	77.541	0.856	76.685	Kroya	102.356	138.447	119.239	113.705
4	77.541	0.786	76.755	Gabuswetan	93.844	83.667	85.632	74.946
5	77.541	1.965	75.576	Cikedung	94.104	94.318	95.942	93.477
6	112.807	-21.117	133.924	Terisi	111.162	114.900	90.230	90.650
7	65.671	5.378	60.293	Lelea	73.711	75.493	77.065	75.512
8	45.342	-2.944	48.286	Bangodua	44.719	46.492	47.409	53.654
9	55.256	1.405	53.851	Tukdana	55.174	56.861	59.164	65.313
10	43.290	-0.913	44.203	Widasari	44.251	37.902	42.291	43.583
11	46.925	1.070	45.855	Kertasemaya	46.875	42.871	50.552	51.163
12	34.054	-2.706	36.760	Sukagumiwa...	38.805	37.664	38.903	39.754
13	57.911	-0.060	57.971	Krangkeng	65.618	61.337	55.942	60.951
14	32.133	-0.353	32.486	Karangampel	29.994	32.182	32.487	31.376
15	32.133	-0.007	32.140	Kedokan Bu...	31.997	32.977	34.059	32.670
16	61.433	-2.330	63.763	Juntinyuat	64.468	62.081	61.325	63.440
17	64.389	-0.454	64.843	Sliyeg	62.716	63.655	66.016	69.906
18	45.060	-0.026	45.086	Jatibarang	43.474	43.667	43.408	47.034
19	29.580	3.115	26.465	Balongan	27.012	28.193	28.588	28.049
20	23.375	-2.768	26.143	Indramayu	28.275	24.379	24.488	24.274
21	31.594	-1.686	33.280	Sindang	23.538	29.845	31.705	27.753
22	21.061	-0.280	21.341	Cantigi	29.178	27.305	27.817	23.280
23	18.431	5.196	13.235	Pasekan	12.700	12.183	12.247	12.802
24	31.536	-1.590	33.126	Lohbener	37.688	37.344	38.895	37.208
25	31.536	3.155	28.381	Arahan	36.160	36.903	36.136	35.023
26	62.330	1.410	60.920	Losarang	72.475	63.832	74.987	58.718
27	67.300	1.056	66.244	Kandanghaur	83.413	62.225	68.544	52.567
28	53.947	-1.614	55.561	Bongas	54.627	58.064	55.413	60.241
29	77.541	-11.234	88.775	Anjatan	85.400	86.795	87.796	97.374
30	49.074	-1.140	50.214	Sukra	50.782	50.936	51.360	48.973
31	46.925	1.090	45.835	Patrol	48.097	44.466	46.827	44.786

Show performance scores						
Model	MSE	RMSE	MAE	MAPE	R2	
Random Forest	24.973	4.997	2.884	0.057	0.962	

Gambar 9. Hasil prediksi menggunakan *random forest*



### 3.6. Analisa Cluster dan Hasil Prediksi

Tahapan Analisa cluster dan hasil prediksi merupakan proses akhir dalam penelitian ini, dimana pada tahapan ini dilakukan Analisa berapa banyak kecamatan yang diprediksi mengalami penurunan panen padi yang dibagi dalam cluster masing-masing, kemudian dilihat apakah jenis cluster berpengaruh terhadap persentase penurunan panen padi, sehingga dapat menjadi acuan untuk Dinas Pertanian dan Ketahanan Pangan atau instansi terkait untuk merumuskan kebijakan.

Tabel 9. Persentase jumlah panen berdasarkan *cluster*

<i>Cluster</i>	<i>Kecamatan</i>	<i>Status Prediksi</i>	<i>Jumlah Panen Turun</i>	<i>Persentase</i>
C3	Gantar	Naik	1	33,3%
	Kroya	Naik		
	Terisi	Turun		
C2	Haurgeulis	Naik	6	50 %
	Gabuswetan	Turun		
	Cikedung	Naik		
	Lelea	Naik		
	Tukdana	Naik		
	Krangkeng	Turun		
	Juntinyuat	Turun		
	Sliyeg	Turun		
	Losarang	Naik		
	Kandanghaur	Naik		
	Bongas	Turun		
	Anjatan	Turun		
C1	Bangodua	Turun	11	68,75 %
	Widasari	Turun		
	Kertasemaya	Naik		
	Sukagumiwang	Turun		
	Karangampel	Turun		
	Kedokan Bunder	Turun		
	Jatibarang	Turun		
	Balongan	Naik		
	Indramayu	Turun		
	Sindang	Turun		
	Cantigi	Turun		
	Pasekan	Naik		
	Lohbener	Turun		
	Arahan	Naik		
	Sukra	Turun		
Patrol	Naik			



Hasil dari analisis data yang dilakukan pada cluster C3 terdapat 1 kecamatan yang mengalami prediksi turun dengan persentase 33,3 %, cluster C2 6 kecamatan dengan persentase 50 %, dan C1 11 kecamatan dengan persentase 68 %.

#### IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan untuk penerapan *K-Means clustering* untuk pemetaan produktivitas padi dan prediksi panen di Kabupaten Indramayu, hasil perhitungan *clustering K-Means* yang dilakukan menggunakan *dataset* historis produksi padi 5 tahun dengan rentang 2019 – 2023 secara manual memperoleh dua iterasi, dimana hasil clustering yang didapatkan membagi kecamatan yang berada di Indramayu menjadi tiga cluster yaitu C1 dengan kategori rendah sebanyak 16 kecamatan , C2 sedang 12 kecamatan, dan C3 tinggi sebanyak 3 kecamatan. Untuk mendapatkan hasil yang akurat, hasil *clustering* yang dilakukan secara manual tersebut dibandingkan menggunakan hasil *clustering* aplikasi *orange3*. Hasil dari perbandingan kedua metode tersebut menunjukkan hasil yang sama baik secara pembentukan *cluster* dan kecamatan yang termasuk didalamnya.

*Dataset* historis produksi panen padi juga dapat dimanfaatkan untuk melakukan prediksi produksi panen menggunakan algoritma *random forest* yang terdapat pada aplikasi *orange3*, dimana hasil dari prediksi tersebut dianalisis bersama hasil *clustering* dengan hasil *cluster* C1 yang memiliki potensi produksi rendah memiliki persentase penurunan produksi padi yang tinggi yaitu 68,75 %, *cluster* C2 persentase penurunan sebanyak 50 %, dan C3 memiliki persentase 33,3%.

Dari informasi tersebut, diketahui bahwasanya daerah *cluster* yang memiliki potensi panen rendah juga memiliki potensi prediksi penurunan tingkat produksi padi dengan persentase yang tinggi, sehingga dengan hasil penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi Dinas Pertanian dan Ketahanan Pangan Kabupaten Indramayu khususnya dalam menentukan kebijakan-kebijakan untuk meningkatkan hasil produksi beras, sehingga ketahanan pangan dapat terjaga.

#### REFERENSI

- [1] D. M. Azzahra *et al.*, “Faktor-faktor yang mempengaruhi impor beras di Indonesia.”. E-Journal Perdagangan Industri dan Moneter Vol. 9. No. 3, September-Desember 2021. ISSN: 2303-1204 (*online*)
- [2] D. Isnawati, “Analisis Permintaan Bahan Pokok Beras Pada Saat Pandemi Covid-19 Di Indonesia”. Nomicpedia: Journal of Economics and Business Innovation Volume 2 Nomor 1, Maret, 2022. E-ISSN : 2775-4774. Available: <https://journal.inspirasi.or.id/index.php/nomicpedia>
- [3] M. Yusuf Nugroho and U. Duta Bangsa Surakarta, “Analisis Faktor-Faktor yang Mempengaruhi Produksi Padi di Sumatera Menggunakan Metode Regresi Linier.” Prosiding Seminar Nasional Teknologi Informasi dan Bisnis (SENATIB) 2023. E-ISSN 2962-1968
- [4] Naira Afshari Putri, Raden Roro Aura Putri Indira, and Vany Seftiani Kurnia, “Dampak Pemberhetian Ekspor Beras Dari Negara-Negara Pengekspor Utama Terhadap Ketahanan Pangan Indonesia,” *Jurnal Ilmu Manajemen, Ekonomi dan Kewirausahaan*, vol. 4, no. 1, pp. 107–114, Jan. 2024, doi: 10.55606/jimek.v4i1.2680.
- [5] D. Riyatno, M. A. S. Tyas, and A. Muksin, “Upaya Pemerintah Kabupaten Indramayu dalam Mengembangkan Pertanian Pangan sebagai Potensi Unggulan,” *Jurnal Syntax Admiration*, vol. 4, no. 9, pp. 2194–2201, Sep. 2023, doi: 10.46799/jsa.v4i9.1015.
- [6] Dinas Ketahanan Pangan Dan Pertanian, “Jumlah Produksi Padi Berdasarkan Kecamatan di Kabupaten Indramayu,” <https://opendata.indramayukab.go.id/dataset/jumlah-produksi-padi-berdasarkan-kecamatan-di-kabupaten-indramayu-2>.
- [7] D. T. Cahaya *et al.*, “PENERAPAN METODE K-MEANS CLUSTERING UNTUK PENGELOMPOKAN POTENSI PADI DI KOTA PAGAR ALAM,”. JATI (Jurnal Mahasiswa Teknik Informatika). Vol. 8 No. 2, April 2024.
- [8] M. Firman, A. Halik, and L. Septiana, “Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi 4.0 Internasional. Analisa Data Untuk Prediksi Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan



DOI: 10.52362/jisamar.v8i3.1572

Ciptaan disebarluaskan di bawah [Lisensi Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



- Algoritma K-Means Clustering,” *Journal of Information System, Applied, Management, Accounting and Research*, vol. 6, no. 4, pp. 856–870, 2022, doi: 10.52362/jisamar.v6i4.939.
- [9] A. Yudhistira and R. Andika, “Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering,” *Journal of Artificial Intelligence and Technology Information (JAITI)*, vol. 1, no. 1, pp. 20–28, Feb. 2023, doi: 10.58602/jaiti.v1i1.22.
- [10] A. Satria, R. M. Badri, and I. Safitri, “Prediksi Hasil Panen Tanaman Pangan Sumatera dengan Metode Machine Learning,” *Digital Transformation Technology*, vol. 3, no. 2, pp. 389–398, Sep. 2023, doi: 10.47709/digitech.v3i2.2852.
- [11] Jailani, M. Syahrani, et al., “Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah”. *HSAN: Jurnal Pendidikan Islam*, 2023, 1.2: 1-9. E-ISSN : 1987 – 1298.
- [12] Y. Yuliska and K. U. Syaliman, “Literatur Review Terhadap Metode, Aplikasi dan Dataset Peringkasan Dokumen Teks Otomatis untuk Teks Berbahasa Indonesia,” *IT Journal Research and Development*, vol. 5, no. 1, pp. 19–31, Jul. 2020, doi: 10.25299/itjrd.2020.vol5(1).4688.
- [13] Harmain, Ahmad, et al “Normalisasi Data Untuk Efisiensi K-Means Pada Pengelompokan Wilayah Berpotensi Kebakaran Hutan Dan Lahan Berdasarkan Sebaran Titik Panas”. *TEKNIMEDIA: Teknologi Informasi dan Multimedia*, 2021, 2.2: 83-89. E-ISSN : 2722 - 6271
- [14] R. A. Indraputra and R. Fitriana, “K-Means Clustering Data COVID-19”. *Jurnal Teknik Industri* Vol. 10 No. 3. ISSN 1411-6340.
- [15] F. Marisa *et al.*, “Digitasi Produktivitas Panen Padi Berbasis K-Means Clustering,” *SMARTICS Journal*, vol. 7, no. 1, 2021, doi: 10.21067/smartics.v7i1.5270.
- [16] E. Mardiani *et al.*, “Membandingkan Algoritma Data Mining Dengan Tools Orange untuk Social Economy,” *Digital Transformation Technology*, vol. 3, no. 2, pp. 686–693, Nov. 2023, doi: 10.47709/digitech.v3i2.3256.
- [17] Adiguno, Siswo; Syahira, Yohanni; Yetri, Milfa. “Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda”. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 2022, 1.4: 275-281.
- [18] N. I. Hommy D. E. Sinaga, “Perbandingan double moving average dengan double exponential smoothing pada peramalan bahan medis habis pakai”. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, 2018, 4.2: 197-204. ISSN : 2407 – 1811.
- [20] T. Tendean, and W. Purba, “Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means,” *SAINTEK (Jurnal Sains dan Teknologi)*, vol. 1, no. 2, 2020 pp. 5–11.

