

Perbandingan Algoritma K-Medoids Dan K-Means Dalam Pengelompokan Kecamatan Berdasarkan Produksi Padi Dan Palawija Di Jember

Akhmad Safrin Sadad Khan¹, Mohamat Fatekurohman², Yuliani Setia Dewi³

^{1,2,3}Matematika, Universitas Jember
E-mail : akhmadsafrincuy@gmail.com¹

Diajukan 24 Oktober 2023 *Diperbaiki* 28 Nopember 2023 *Diterima* 6 Desember 2023

Abstrak

Latar Belakang: Pengelolaan tanaman pangan sangat penting untuk mendukung ketahanan pangan. Dataset menunjukkan variasi hasil panen padi dan tanaman pokok lainnya. Variasi hasil panen tersebut memerlukan pengelompokan wilayah berdasarkan hasil panen. Algoritma yang umum digunakan dalam analisis *clustering* adalah *K-means* dan *K-medoids*. Terdapat pada kedua algoritma tersebut yaitu *K-means* kompleksitas waktu lebih cepat dan *K-medoids* lebih tahan dengan data *outlier*. Sehingga perbandingan kedua algoritma dapat membantu pemilihan algoritma yang lebih baik dalam kasus tertentu

Tujuan: memperoleh hasil perbandingan *cluster* terbaik dengan menggunakan algoritma *K-means* dan *K-medoids* di Kabupaten Jember berdasarkan produksi padi dan palawija dan mengetahui hasil *clustering* dengan algoritma pengelompokan terbaik Kecamatan Jember berdasarkan produksi padi dan palawija.

Metode: Algoritma *clustering* yang digunakan yaitu *K-means* dan *K-medoids*. Metode evaluasi menggunakan Davies Bouldin Index. Sumber data berasal dari data sekunder dari BPS Kabupaten Jember tahun 2020.

Hasil: Diperoleh algoritma terbaik yaitu *K-means* dengan DBI 0,648 lebih kecil dibandingkan *K-medoids* 0,886 dibagi menjadi 6 kluster yaitu kluster satu sebanyak 1 kecamatan, kluster dua sebanyak 3 kecamatan, kluster tiga sebanyak 2 kecamatan, kluster empat sebanyak 3 kecamatan, kluster lima sebanyak 8 kecamatan dan kluster 6 sebanyak 14 kecamatan.

Kesimpulan: *K-means* dengan 6 *cluster* menjadi algoritma terbaik untuk pengelompokan produksi tanaman pangan di Kabupaten Jember.

Kata kunci: *K-Means, K-Medoids, Davies Bouldin Index, Tanaman Pangan*

Abstract

Background: Food crop management is essential for ensuring food security. The dataset indicates variations in crop yields, such as rice and other staple crops. The variability in harvest yields necessitates the grouping of regions based on their crop yields. Two commonly used algorithms in clustering analysis are *K-means* and *K-medoids*. In both algorithms, *K-means* has a faster time complexity, while *K-medoids* is more robust in handling outlier data. Therefore, comparing the two algorithms can assist in choosing the more suitable algorithm for specific cases.

Objective: The objective of this study is to obtain the best cluster comparison results using the *K-means* and *K-medoids* algorithms in Jember Regency based on rice and staple crop production.

Methods: The clustering algorithms used are *K-means* and *K-medoids*, and the Davies Bouldin Index is used for evaluation. The data source is secondary data from BPS Jember Regency in 2020.

Results: The best-performing algorithm obtained is *K-means* with a Davies-Bouldin Index (DBI) of 0.648, which is smaller compared to *K-medoids* with a DBI of 0.886. Divided into six clusters. Cluster one consists of one sub-district, cluster two includes three sub-districts, cluster three comprises two sub-districts, cluster four has three sub-districts, cluster five consists of eight sub-districts, and cluster six includes fourteen sub-districts.

Conclusion: Therefore, based on the results, we can conclude that the *K-means* algorithm is the best algorithm for grouping food crop production in Jember Regency.

Keywords : *K-Means, K-Medoids, Davies Bouldin Index, Agricultural Production*

PENDAHULUAN

Indonesia memiliki iklim tropis yang mendukung petani cocok untuk membudidayakan padi dan palawija. Indonesia berhasil menjadi negara salah satu penghasil beras terbanyak di dunia pada tahun 2018 (Wijayanto & Fathoni, 2021). Negara Indonesia termasuk daerah agraris, mengingat pertanian di Indonesia merupakan salah satu sektor penting untuk pembangunan perekonomian Indonesia. Kabupaten Jember kaya akan potensi tanaman pangan dan dikenal sebagai penghasil utama padi di Provinsi Jawa Timur, bahkan di tingkat nasional. Kabupaten Jember menyumbangkan sekitar 23% dari total produksi padi Jawa Timur, sekitar 19% dari total produksi jagung Jawa Timur dan sekitar 8 % dari total produksi kedelai Jawa Timur (Badan Perencanaan Pembangunan Kabupaten Jember, 2015).

Pengelolaan tanaman pangan juga merupakan salah satu hal yang penting untuk mendukung ketahanan pangan (Efendi, 2016). Berdasarkan data hasil panen dari BPS Kabupaten Jember tahun 2020 menampilkan hasil produksi padi dan palawija yang bervariasi jumlahnya, sehingga saat ini belum ada sistem cocok digunakan untuk pengelompokan. Dalam data hasil panen yang bervariasi tersebut, perlu dilakukan pengelompokan daerah berdasarkan hasil panennya. Metode pengelompokan tersebut merupakan salah satu fungsi dari *data mining* (Han et al., 2022).

Teknik *data mining* berfungsi untuk menemukan wawasan atau pengetahuan baru, *data mining* juga merupakan bidang yang relatif baru untuk dijelajahi (Talakua et al., 2017). *Data mining* banyak digunakan salah satunya yaitu digunakan untuk pengelompokan (*Clustering*) (Holmes & Jain, 2011). Algoritma yang umum

digunakan dalam analisis *clustering* adalah *K-means* dan *K-medoids*. Kedua algoritma tersebut memiliki tujuan yang sama, yaitu mempartisi data menjadi beberapa *cluster* berdasarkan kesamaan karakteristik (Surya & Laurence Aroquiaraj, 2019). Namun ada perbedaan pada keduanya yaitu kelebihan yang dimiliki *K-means* dalam kompleksitas waktu lebih cepat dalam menangani data serupa dari *K-medoid*, sedangkan kelebihan *K-medoids* lebih tahan terhadap data outlier. Kedua algoritma tersebut juga memiliki kelemahan berbeda, seperti *K-means* tidak dapat menangani data outlier dan *K-medoids* dengan waktu komputasi membutuhkan waktu yang lebih lama, sehingga perbandingan kedua algoritma dapat membantu pemilihan algoritma yang lebih baik dalam kasus tertentu (Han et al., 2022).

Penerapan algoritma *K-means* dan *K-medoids* telah banyak ditemukan dalam berbagai studi dan kasus di berbagai bidang seperti penelitian Surya & Laurence Aroquiaraj (2019) menggunakan *clustering* dimana membandingkan performa *K-means* dan *K Medoids* dengan data pertanian menunjukkan bahwa algoritma *K-medoids* mencapai tingkat kesalahan yang lebih rendah dan tingkat akurasi tinggi yang memuaskan dibandingkan dengan algoritma *K-means*. Penelitian yang lain Fimawahib et al. (2022) dengan menggunakan *clustering* pada pengelompokan produksi beras dan padi pada ketahanan pangan di Provinsi Riau dengan menggunakan algoritma *K-medoids* dengan evaluasi *Davies Bouldin Index (DBI)* didapatkan nilai sebesar 0,626. Sehingga dikatakan cukup baik untuk evaluasi *cluster*, karena mendekati nol. Penelitian sebelumnya yang lain menurut (Sholikhah, 2022) memberikan kesimpulan *K-means* dengan hasil tingkat akurasi yang lebih tinggi saat digunakan dengan data kecil dan tidak ada *outlier*.

Produksi tanaman mengacu pada jumlah atau hasil dari tanaman yang dihasilkan dari

suatu wilayah tertentu dalam suatu periode waktu tertentu. Produksi pertanian menggunakan tanaman atau hewan untuk menghasilkan produk untuk mempertahankan atau meningkatkan kehidupan manusia (Arifien et al., 2022). Produksi pertanian mengacu juga pada proses menanam dan memanen tanaman, serta beternak, untuk menyediakan makanan, bahan mentah, dan mendukung ketahanan pangan suatu negara (Pretty, 2008).

Data mining merupakan penggalian informasi dalam jumlah yang sangat besar. *Data mining* mencakup dari statistik data, pengumpulan data, analisis data dan ekstraksi data. *Knowledge discovery* dan *information harvesting* merupakan nama lain *data mining* (Han et al., 2022). *Data mining* merupakan informasi yang berguna dengan proses logis. Setelah ditemukan informasinya dan polanya, dapat digunakan sebagai alat pendukung pengambilan keputusan dalam mengembangkan bisnis dan lain – lain (Arhami & Nasir, 2020).

Clustering juga bisa disebut *segmentation*. Metode ini merupakan sebuah teknik yang digunakan untuk menemukan sub-kelompok atau *cluster*, dimana *cluster* objek-objek yang berada di dalam sebuah *cluster* akan memiliki kemiripan tinggi. Tujuan dari *clustering* adalah membagi seluruh data menjadi beberapa *cluster* (Han et al., 2022). Dalam setiap *cluster* memiliki tingkat kemiripan yang tinggi. Ukuran kemiripan dihitung dengan jarak. Jarak antar *cluster* sejauh mungkin dan jarak dalam *cluster* sedekat mungkin (Tendean & Purba, 2020). *Clustering* mempunyai dua metode untuk pengelompokan data yaitu *hierarchical clustering* dan *non-hierarchical clustering*. Metode *K-means* dan *K-medoids* termasuk *Partitional clustering* (Tan et al., 2013).

K-means adalah algoritma pengelompokan yang populer yang

mempartisi data ke dalam k *cluster* dengan memperbarui secara iteratif centroid *cluster* dan menugaskan setiap titik data ke centroid terdekat (Umagapi et al., 2023). *K-means* merupakan algoritma *clustering* yang digunakan dalam suatu pengelompokan secara membagi data ke dalam *cluster* lalu memisahkan data ke dalam kelompok yang berbeda. Algoritma ini mampu memperpendek jarak antara data ke *clusternya* (Han et al., 2022).

Algoritma *K-medoids* memakai objek data yang mewakili suatu *cluster* pada kumpulan objek (*medoid*). *Medoid* adalah objek data yang merupakan pusat dari *cluster* dan memiliki jarak rata-rata terkecil dari semua objek dalam *cluster* tersebut (Setyawati, 2017). *K-medoids* lebih tahan terhadap *outlier* dibandingkan dengan *K-means*. *K-medoids* mencari k sebagai objek perwakilan untuk meminimalkan jumlah ketidakmiripan antar objek data, sedangkan *K-means* menggunakan jumlah jarak *Euclidean* untuk objek data. Jadi metrik jarak *K-medoids* mengurangi gangguan dan nilai-nilai data yang jauh dari kelompok. (Arora et al., 2016).

Kedua Algoritma *clustering* tersebut di evaluasi dengan *Davies-Bouldin Index* (DBI). DBI digunakan untuk mengevaluasi kualitas *cluster* dengan mengukur metrik kohesi dan metrik sparasi. Metrik kohesi menunjukkan keterikatan atau kesamaan anggota dalam *cluster*. Metrik separasi menunjukkan sejauh mana *cluster* berbeda satu sama lain. *Cluster* dikatakan optimal, jika nilai DBI rendah dalam kohesi dan tinggi dalam separasi (Meiriza & Ali, 2023). Langkah Langkah proses DBI adalah *Sum of square Within Cluster*, *Sum of square Between-cluster*, Rasio dan DBI (Jollyta et al., 2021).

Berdasarkan permasalahan yang terjadi, maka tujuan dalam penelitian ini adalah memperoleh hasil perbandingan *cluster* terbaik dengan menggunakan algoritma di Kabupaten Jember berdasarkan

produksi tanaman pangan. Penelitian ini bermanfaat bagi pemerintah dan petani agar meningkatkan pemahaman tentang produksi padi dan palawija di Kabupaten Jember dengan mengelompokkan kecamatan dapat mengetahui informasi kondisi di Kabupaten Jember. Penelitian ini dapat mengambil keputusan tentang pengembangan pertanian di wilayah tersebut.

METODE

Desain Penelitian

Desain penelitian yang digunakan adalah pendekatan kuantitatif. Metode yang digunakan yaitu K-means dan K-medoids sebagai studi perbandingan. Metode clustering menggunakan software R Studio dengan evaluasi DBI berdasarkan kohesi dan separasi.

Populasi Dan Sampel

Populasi yang menjadi fokus penelitian adalah seluruh kecamatan di Jember, sementara sampelnya adalah sebagian dari kecamatan-kecamatan dengan produksi tanaman pangan ini yang akan diamati dan menjadi representatif dari seluruh populasi tersebut.

Teknik Sampling

Pengambilan data pada website resmi BPS (2020) Kabupaten Jember.

Subyek Penelitian

Subjek penelitian ini adalah produksi tanaman pangan di Kabupaten Jember berdasarkan kecamatan pada tahun 2020. Variabel penelitian dapat dilihat pada Tabel 1 sebagai berikut :

Tabel 1. Variabel Penelitian

Variabel	Definisi	Satuan
Produksi Padi	Jumlah produksi padi bentuk gabah kering	Ton
Produksi Jagung	Jumlah produksi jagung bentuk pipilan kering	Ton
Produksi Kedelai	Jumlah produksi kedelai bentuk biji kering	Ton
Produksi Kacang tanah	Jumlah produksi kacang tanah bentuk biji kering	Ton
Produksi Ubi kayu	Jumlah produksi ubi kayu bentuk umbi basah	Ton
Produksi Ubi jalar	Jumlah produksi ubi jalar bentuk umbi basah	Ton

Teknik Analisis Data.

Teknik analisis data melakukan perbandingan algoritma *clustering K-means* dan *K-medoids* menggunakan software Rstudio. Langkah-langkah penyelesaian data sebagai berikut :

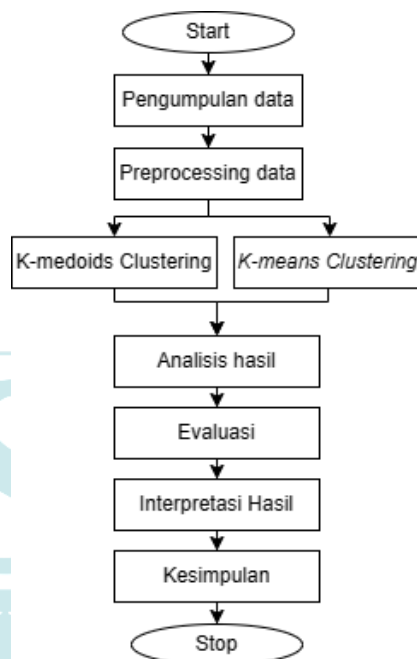
1. Pengumpulan data 6 jenis produksi tanaman pangan diambil dari website BPS Kabupaten Jember
2. preprocessing data dilakukan serangkaian proses pada data mentah sebelum digunakan untuk analisis .data cleaning membersihkan data yang tidak perlu dan data selection memilih variabel yang digunakan
3. Melakukan pengolahan data dengan menggunakan algoritma *K-means* dan *K-medoid*. Kedua algoritma tersebut mengelompokkan data secara partisi, meningkatkan kesamaan karakteristik antara data dalam satu *cluster*, serta mengurangi kesamaan karakteristik data antar *cluster*
4. Melakukan evaluasi terhadap model yang telah dibuat dengan menggunakan DBI yang bertujuan untuk mengetahui kualitas dari suatu objek pada suatu *cluster*. Sehingga pada hasil akhirnya

Perbandingan Algoritma K-Medoids Dan K-means Dalam ...

mendapatkan nilai efisiensi terbaik dari dua metode algoritma yang digunakan.

- Melakukan interpretasi untuk memahami makna dari *cluster-cluster* yang terbentuk setelah menjalankan algoritma *clustering*.

Berikut merupakan langkah-langkah penelitian dalam diagram alir pada Gambar 1



Gambar 1. Diagram alir langkah-langkah penelitian

HASIL DAN PEMBAHASAN

Data yang dikumpulkan adalah data produksi beberapa tanaman pangan padi dan palawija diambil dari Badan Pusat Statistik tahun 2020. Pengumpulan data dilakukan menggunakan data sekunder yang langsung diambil dari website resmi BPS. Data terdiri dari beberapa tabel produksi untuk masing-masing tanaman pangan. Nama-nama kecamatan di Kabupaten Jember, yang akan dikelompokkan berdasarkan variabel, yaitu produksi dari 6 jenis tanaman pangan.

Pengelompokan *K-means clustering* dan *K-medoids clustering*. Pertama perlu dilakukan menentukan jumlah *cluster*, jika

jumlah *cluster* terlalu banyak maka interpretasi hasil menjadi sulit dan masing-masing *cluster* mungkin hanya berisi sedikit atau bahkan satu data saja. Sehingga sulit untuk menarik kesimpulan tentang karakteristik atau pola dalam setiap *cluster*, jadi jumlah klaster yang lebih sedikit dan terkonsentrasi dapat memberikan informasi yang lebih berguna dan signifikan. Jumlah *cluster* yang dipilih yaitu yaitu $k=2$, $k=3$, $k=4$, $k=5$, $k=6$.

Tabel 1. Hasil percobaan *K-means*

Jumlah k	Cluster					
	1	2	3	4	5	6
k = 2	5	26	-	-	-	-
k = 3	14	12	5	-	-	-
k = 4	2	12	14	14	-	-
k = 5	14	2	9	3	3	-
k = 6	3	14	2	3	8	1

Tabel 2. Hasil percobaan *K-medoids*

Jumlah k	Cluster					
	1	2	3	4	5	6
k = 2	5	26	-	-	-	-
k = 3	12	5	14	-	-	-
k = 4	4	4	14	9	-	-
k = 5	3	3	2	14	9	-
k = 6	3	3	6	6	9	1

Tabel 1 merupakan hasil uji coba k dengan masing masing *cluster* algoritma *K-means*, Tabel 2 merupakan hasil uji coba k dengan masing masing *cluster* *K-medoids*. Beberapa hasil jumlah k Tabel 1 dan Tabel 2 terdapat kesamaan karena kedua metode merupakan metode partisi, namun perbedaan kedua algoritma tersebut pada perubahan centroid baru. Setelah menggunakan metode *K-means* dan *K-medoids* untuk melakukan pengelompokan data, langkah berikutnya dalam penelitian ini adalah menentukan jumlah *cluster* yang paling optimal menggunakan Rstudio. Melakukan dengan melihat Davies-Bouldin Index (DBI) untuk setiap nilai *cluster* $k=2$, $k=3$, $k=4$, $k=5$, $k=6$. Dipilih sampai $k=6$, karena jika terlalu banyak *cluster* dapat membuat sulit untuk diinterpretasi hasilnya. Hasil perbandingan DBI dari kedua metode

Perbandingan Algoritma K-Medoids Dan K-means Dalam ...

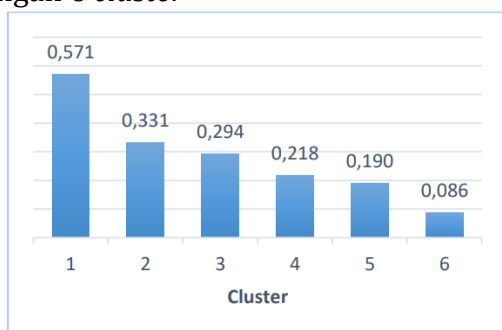
tersebut dapat dilihat dalam Tabel 3.

Tabel 3. Hasil Perbandingan DBI

cluster	K-means	K-medoids
k = 2	0,716	1,282
k = 3	0,746	0,929
k = 4	0,780	0,944
k = 5	0,804	0,886
k = 6	0,648	1,161

Berdasarkan tabel 3 Menunjukkan hasil terbaik nilai DBI terdapat pada algoritma *K-means* dengan nilai DBI yaitu 0,648. Nilai tersebut dikatakan baik sesuai prinsip DBI dengan semakin kecil nilai DBI atau mendekati nol maka semakin optimal *cluster* yang dihasilkan. DBI menggunakan pengukuran kohesi yang menunjukkan *cluster* berbeda satu sama lain dan pengukuran separasi anggota *cluster* berbeda satu sama lain.

Interpretasi hasil merupakan tahapan terakhir tahapan *Data mining*. Setelah analisis kedua algoritma *clustering* dan evaluasi algoritma *clustering* terbaik selesai dengan mendapatkan metode *K-means* dan *cluster* 6 yang paling optimal, Dimana menginterpretasikan hasilnya dengan memberikan nama hasil *clustering* diperlukan memberikan kategori agar dapat membedakan karakteristik masing-masing dari kelompok. Berikut merupakan gambar yang digunakan untuk memberikan kategori hasil *clustering* algoritma *K-means* dengan 6 *cluster*



Gambar 2. Rata-rata per klaster

Berdasarkan Gambar 2, merupakan hasil rata-rata per *cluster* dari urutan tinggi ke rendah. Gambar 2 menjelaskan urutan

tertinggi ke rendah dengan hasil nilai rata-rata per *cluster*. Berikut pada Tabel 3 merupakan penjelasan lebih rinci untuk rata-rata setiap variabel dari Gambar 2.

Tabel 3. Rata Rata per *cluster* algoritma *K-means*

cluster	Rata Rata Produksi					
	Padi	Jagung	Ubi Kayu	Ubi Jalar	Kacang Tanah	Kedelai
1	40406	50567	517	2414	4	2
2	51693	6951	19003	505	281	5181
3	23715	40179	12733	63	35	30
4	43173	20637	2178	0	2	232
5	42058	5955	903	339	145	283
6	21826	4472	1145	516	20	52

Tabel 3 digunakan untuk melihat interpretasi hasil agar mudah memahami data yang banyak. interpretasi hasil dilakukan dengan memberikan nama masing-masing sesuai dengan *cluster*. Karena terdapat 6 *cluster* dan penamaan kategori yang sulit, maka dilakukan dengan bantuan nilai kuartil untuk menentukan tingkat kategori produksi tanaman pangan tinggi, sedang dan rendah sebagai batas pembagian interval. Nilai kuartil pertama (Q1) sebagai batas rendah, nilai kuartil kedua (Q2 atau median) sebagai batas sedang dan nilai kuartil ketiga (Q3) sebagai batas tinggi.

Tabel 4. Interval pembagian produksi

Quartil	Padi	Jagung	Ubi Kayu	Ubi jalar	Kacang tanah	Kedelai
Q3 (Tinggi)	> 31355	> 18458	> 2233	> 415	> 16	> 41
Q2 (Sedang)	31355 - 21801	18458 - 2981	2233 - 15	415 - 0	16 - 0	41 - 0
Q1 (Rendah)	< 21801	< 2981	< 15	< 0	< 0	< 0

Tabel 4 menyajikan data untuk menemukan nilai produksi tinggi, sedang dan rendah. Tahap selanjutnya yaitu

memberikan nama kategori pada masing-masing *cluster* dapat dilakukan berdasarkan karakteristik produksi tanaman pada *cluster*.

- a. *Cluster* satu dikategorikan *cluster produksi tanaman tinggi* karena memiliki tingkat produksi yang tinggi untuk sebagian besar jenis tanaman.
- b. *Cluster* dua dikategorikan *cluster produksi jagung dan ubi jalar tinggi* karena memiliki tingkat produksi yang tinggi untuk jagung dan ubi jalar, namun rendah untuk jenis tanaman lainnya.
- c. *Cluster* tiga dikategorikan *cluster produksi campuran* karena memiliki kombinasi tingkat produksi yang beragam untuk setiap jenis tanaman, dengan nilai yang bervariasi dari rendah hingga sedang.
- d. *Cluster* empat dikategorikan *cluster produksi padi tinggi* karena memiliki tingkat produksi yang tinggi untuk padi, namun rendah untuk jenis tanaman lainnya.
- e. *Cluster* lima dikategorikan *cluster produksi tanaman sedang* karena memiliki tingkat produksi yang cenderung sedang untuk semua jenis tanaman.
- f. *Cluster* enam dikategorikan *cluster produksi tanaman rendah* karena memiliki tingkat produksi yang rendah untuk semua jenis tanaman.

PENUTUP

esimpulan

Berdasarkan hasil yang diuraikan peneliti, maka dapat disimpulkan sebagai berikut.

1. *K-means* menjadi algoritma terbaik untuk pengelompokan produksi tanaman pangan di Kabupaten Jember dengan nilai DBI lebih kecil yaitu 0,648 dibandingkan DBI algoritma *K-medoids* sebesar 0,886.
2. Hasil pengelompokan *K-means* menjadi

6 *cluster* yaitu diperoleh *cluster* satu dengan nama *cluster* produksi tanaman tinggi sebanyak 1 kecamatan. *Cluster* dua dengan nama *cluster* produksi jagung tinggi, ubi jalar dan kacang tanah rendah sebanyak 3 kecamatan. *Cluster* tiga dengan nama ubi kayu tinggi, jagung rendah sebanyak 2 kecamatan. *Cluster* empat dengan nama *cluster* produksi padi dan jagung sedang, ubi kayu dan kedelai rendah sebanyak 3 kecamatan. *Cluster* lima dengan nama *cluster* produksi tanaman sedang sebanyak 8 kecamatan. *Cluster* enam dengan nama *cluster* produksi tanaman sangat rendah sebanyak 14 kecamatan.

Saran

Berdasarkan hasil penelitian yang diperoleh, disarankan penelitian selanjutnya menggunakan perbandingan partisi *clustering* seperti *X-means* atau algoritma *hierarchical clustering*. Penelitian selanjutnya juga dapat melakukan evaluasi perbandingan yang lain seperti *dunn index*.

DAFTAR PUSTAKA

- Arhami, M., Kom, M., & Muhammad Nasir, S. T. (2020). *Data Mining-Algoritma dan Implementasi*. Penerbit Andi.
- Arifien, Y., Putra, R. P., Wibaningwati, D. B., Anasi, P. T., Masnang, A., Rizki, F. H., ... & Indrawati, E. (2022). *Pengantar Ilmu Pertanian*. Get Press.
- Arora, P., & Varshney, S. (2016). Analysis of k-means and k-medoids algorithm for big data. *Procedia Computer Science*, 78, 507-512. <https://doi.org/10.1016/j.procs.2016.02.095>
- Badan Perencanaan Pembangunan Kabupaten Jember. (2015). *PENYUSUNAN MASTERPLAN PERTANIAN*.

- Efendi, E. (2016). Implementasi sistem pertanian berkelanjutan dalam mendukung produksi pertanian. *Warta Dharmawangsa*, (47). <https://doi.org/10.46576/wdw.v0i4.7.231>
- Fimawahib, L., Bakti, I. R., & Supriyanto, A. (2022). Algoritma K-Medoids untuk Pengelompokan Produksi Padi dan Beras sebagai Upaya Optimalisasi Ketahanan Pangan di Provinsi Riau. *SATIN-Sains dan Teknologi Informasi*, 8(2), 13-24. <https://doi.org/10.33372/stn.v8i2.877>
- Han, J., Pei, J., & Tong, H. (2022). Data mining: concepts and techniques. Morgan kaufmann.
- Holmes, D. E., & Jain, L. C. (Eds.). (2011). *Data Mining: Foundations and Intelligent Paradigms: Volume 1: Clustering, Association and Classification (Vol. 23)*. Springer Science & Business Media.
- Jollyta, D., Siddik, M., Mawengkang, H., & Efendi, S. (2021). *Teknik Evaluasi Cluster Solusi Menggunakan Python Dan Rapidminer*. Deepublish.
- Meiriza, A., & Ali, E. (2023). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Program BPJS Ketenagakerjaan. *Indonesian Journal of Computer Science*, 12(2). <https://doi.org/10.33022/ijcs.v12i2.3184>
- P Pretty, J. (2008). Agricultural sustainability: concepts, principles and evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1491), 447-465.
- Setyawati, A. W. (2017). Implementasi Algoritma Partitioning Around Medoid (PAM) untuk Pengelompokan Sekolah Menengah Atas di DIY Berdasarkan Nilai Daya Serap Ujian Nasional. *Fakultas Sains dan Teknologi, Universitas Sanata Darma, Yogyakarta*.
- Sholikhah, N. A. (2022). Studi Perbandingan Clustering Kecamatan di Kabupaten Bojonegoro Berdasarkan Keaktifan Penduduk Dalam Kepemilikan Dokumen Kependudukan. *Jurnal Statistika dan Komputasi*, 1(1), 42-53. <https://doi.org/10.32665/statkom.v1i1.443>
- Surya, P., & Laurence Aroquiaraj, I. (2019). Performance analysis of K-means and K-medoid clustering algorithms using agriculture dataset. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 6(1). <https://ssrn.com/abstract=3345800>
- Talakua, M. W., Leleury, Z. A., & Taluta, A. W. (2017). Analisis cluster dengan menggunakan metode k-means untuk pengelompokan Kabupaten/Kota di provinsi maluku berdasarkan indikator indeks pembangunan manusia tahun 2014. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 11(2), 119-128. <https://doi.org/10.30598/barekengv011iss2pp119-128>
- Tan, P. N., Steinbach, M., & Kumar, V. (2013). *Data mining cluster analysis: basic concepts and algorithms*. Introduction to data mining, 487, 533.
- Tendean, T., & Purba, W. (2020). Analisis Cluster Provinsi Indonesia Berdasarkan Produksi Bahan Pangan Menggunakan Algoritma K-Means. *Jurnal Sains dan Teknologi*, 1(2), 5-11. <https://doi.org/10.34013/saintek.v1i2.31>

Umagapi, I. T., Umaternate, B., Hazriani, H., & Yuyun, Y. (2023). Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa. *Prosiding SISFOTEK*, 7(1), 303-308.

Wijayanto, S., & Fathoni, M. Y. (2021). Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means. *JUPITER (Jurnal Penelitian Ilmu dan Teknik Komputer)*, 13(2), 212-219.

