

Klasterisasi Produksi Tanaman Perkebunan di Provinsi Jawa Timur Menggunakan Algoritma *Fuzzy C-Means*

Shofinatul Wahdah Nur Aulia¹, Putroue Keumala Intan²

^{1,2} Program Studi Matematika, UIN Sunan Ampel Surabaya
Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur, 60294
Email: shofinatul12@gmail.com¹, puput.in@gmail.com²
Korespondensi penulis : shofinatul12@gmail.com

Abstrak

Produksi tanaman perkebunan di Provinsi Jawa Timur menjadi mata pencaharian umum dikarenakan kebutuhan pokok utama pangan berasal dari pertanian atau perkebunan. Hasil produksi tanaman perkebunan memiliki beberapa perbedaan pada setiap daerah kabupaten atau kota. Karena terdapat perbedaan dari hasil produksi tanaman perkebunan, maka diperlukan klasterisasi pada data perkebunan. Tujuan dilakukannya penelitian ini adalah untuk mengetahui daerah yang cocok dan berpotensi dalam memproduksi tanaman perkebunan. Algoritma yang akan digunakan dalam penelitian ini adalah *Fuzzy C-Means*, dimana derajat keanggotaan menjadi penentu posisi data dalam klaster. Penerapan algoritma *Fuzzy C-Means* untuk mengetahui nilai yang optimal akan digunakan uji *silhouette coefficient* dengan hasil yang didapatkan sebesar 0,6413.

Kata Kunci: *Fuzzy C-Means*, Klasterisasi, Produksi Tanaman Perkebunan

Abstract

Production of plantation crops in East Java Province is a public livelihood because the main staple needs of food come from agriculture or plantations. The results of the production of plantation crops have some differences in each district or city. Because there are differences in the production of plantation crops, it is necessary to cluster the plantation data. The purpose of this research is to find out which areas are suitable and have the potential to produce plantation crops. The algorithm that will be used in this study is *Fuzzy C-Means*, where the degree of membership determines the position of the data in the cluster. The application of the *Fuzzy C-Means* algorithm to find out the optimal value will use the *silhouette coefficient* test with the results obtained at 0.6413.

Keywords: Clustering, *Fuzzy C-Means*, Production of Plantation Crops

1. Pendahuluan

Provinsi Jawa Timur yang terletak di bagian timur Pulau Jawa merupakan salah satu daerah yang memiliki potensi besar dalam sektor pertanian, khususnya produksi tanaman perkebunan [1]. Sektor pertanian menjadi peranan penting karena terletak pada posisinya sebagai produsen utama dalam rantai sebelum mencapai tahap pengolahan [2]. Jawa Timur menjadi salah satu produsen utama tanaman perkebunan di Indonesia dengan keadaan geografis yang subur, iklim tropis yang mendukung dan lahan yang luas [3].

Produksi tanaman perkebunan di Jawa Timur memiliki peran penting dalam perekonomian [4]. Sebagian besar individu memilih mata pencaharian sebagai petani dikarenakan kebutuhan pokok utama pangan bersumber dari pertanian ataupun perkebunan. Berdasarkan data perkebunan yang ada di beberapa Kabupaten atau Kota di Jawa Timur, terdapat perbedaan antara jenis dan hasil produksi perkebunan [5]. Jenis tanaman perkebunan yang dihasilkan, seperti kelapa, kopi, coklat (kakao), tebu, tembakau dan karet menjadi tanaman perkebunan utama [6].

Dalam rangka meningkatkan produksi tanaman perkebunan, diperlukan adanya pengklasteran atau pengelompokan guna mengidentifikasi daerah-daerah yang menghasilkan produksi secara optimal dan maksimal [7]. Algoritma yang dapat digunakan untuk pengklasteran adalah *Fuzzy C-Means* (FCM). Algoritma ini menggunakan derajat keanggotaan sebagai faktor penentu dalam menentukan posisi setiap data dalam suatu kluster [8]. Beberapa kelebihan dari FCM yaitu memberikan hasil klusterisasi yang realistis, fleksibilitas dalam menghadapi data yang ambigu dan memberikan gambaran yang akurat tentang keterkaitan data pada setiap kluster [9].

Beberapa penelitian sebelumnya telah membahas tentang pengklasteran. Sebagai contoh, dalam penelitian yang dilakukan oleh Surya [10], setelah melakukan uji validitas, diperoleh hasil *silhouette* sebesar 0,7695, dimana hasil tersebut menunjukkan bahwa kualitas kluster yang dihasilkan sudah optimal. Selanjutnya, penelitian yang dilakukan oleh Nurhidayati [11] berhasil mencapai tingkat akurasi di atas 90% yang menunjukkan keberhasilan cukup dalam pengklasteran. Dalam penelitian lain yang dilakukan oleh Vera [12] didapatkan tingkat akurasi dan nilai keanggotaan kluster yang cukup baik. Penelitian ini dilakukan dengan merujuk pada penelitian terdahulu. Akan tetapi, yang membedakan penelitian ini dengan penelitian terdahulu adalah kontribusi dalam bidang klusterisasi.

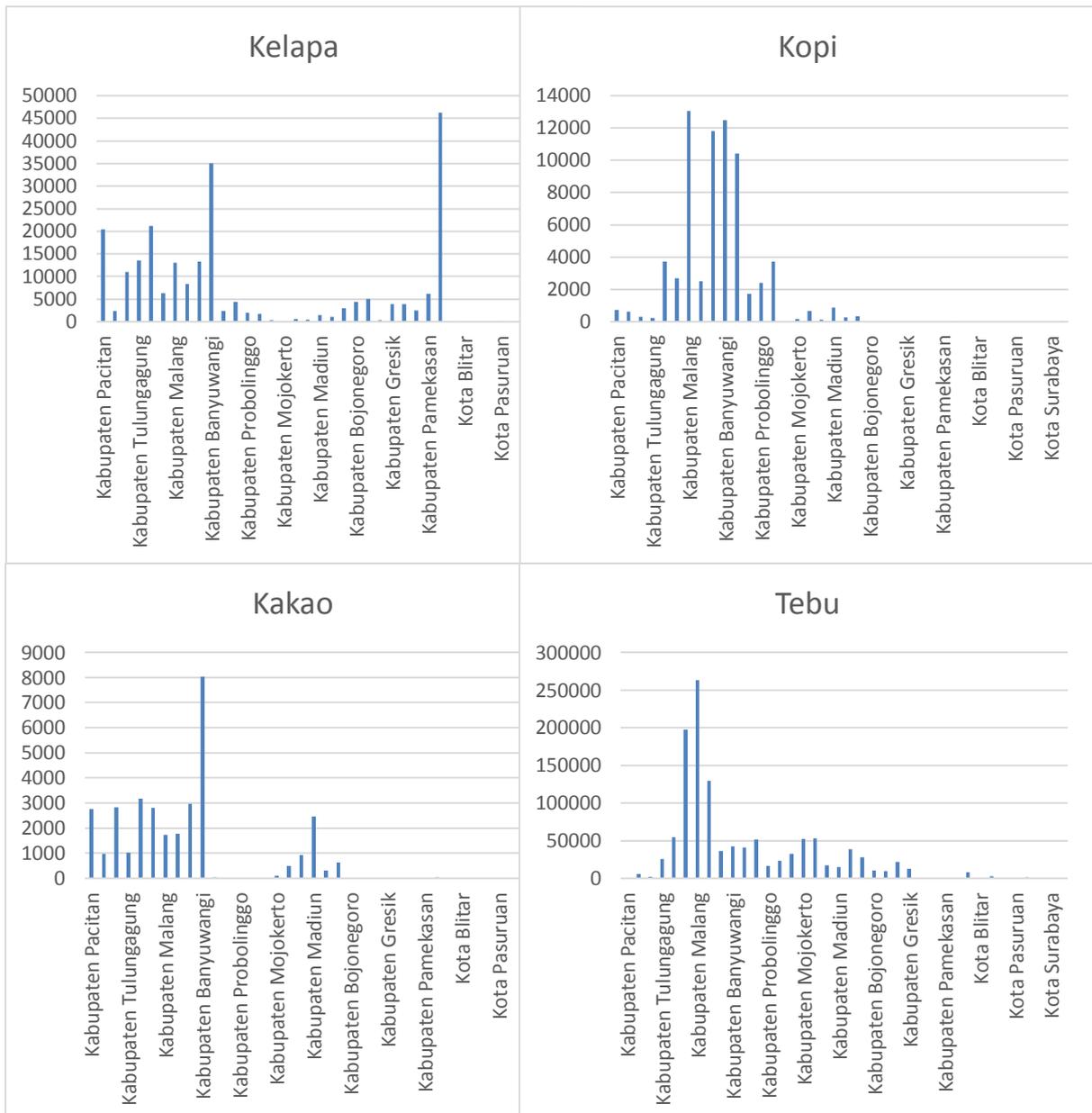
Dengan demikian, penelitian ini akan menerapkan algoritma *Fuzzy C-Means* untuk melakukan pengklasteran pada produksi tanaman perkebunan di Provinsi Jawa Timur. Dengan menggunakan algoritma *Fuzzy C-Means*, diharapkan dapat memperoleh hasil mengenai karakteristik produksi tanaman perkebunan di Jawa Timur. Tujuan dari penelitian ini adalah untuk mengidentifikasi daerah-daerah yang memiliki potensi tinggi dalam menghasilkan tanaman perkebunan, serta untuk menentukan daerah-daerah yang sesuai untuk kegiatan produksi tanaman perkebunan.

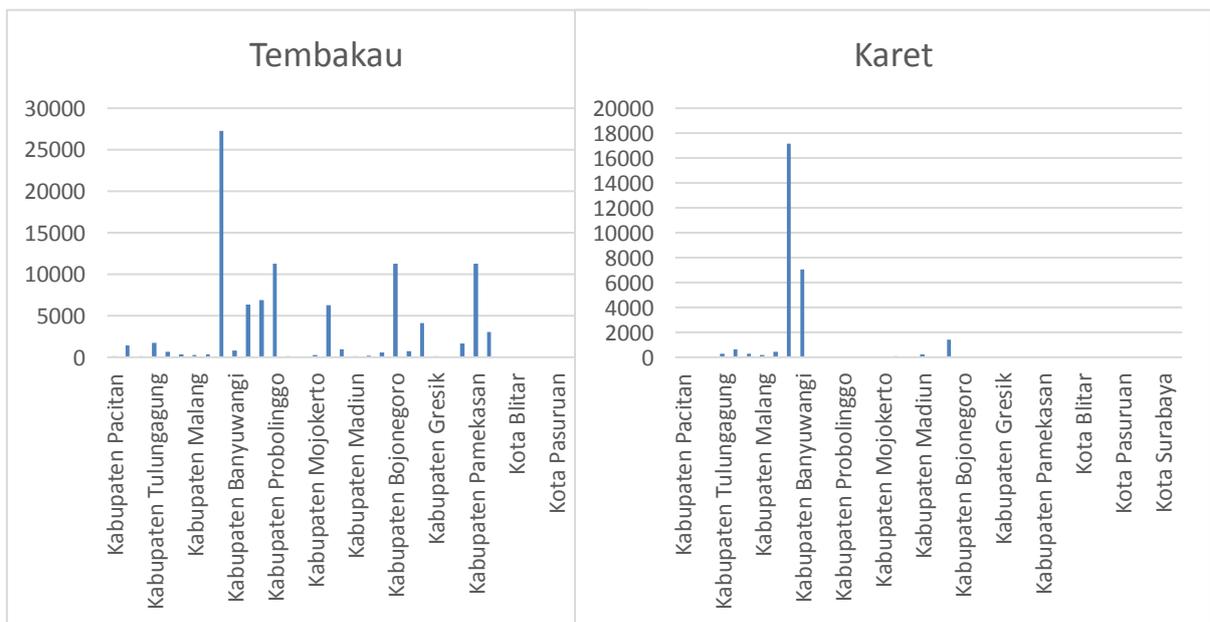
2. Metode Penelitian

Penelitian ini menggunakan metode *Fuzzy C-Means* (FCM) sebagai pendekatan utama. *Fuzzy C-Means* (FCM) atau biasa dikenal dengan *Fuzzy Isodata* memakai model pengelompokan *fuzzy* yang menjadikan data dari semua variabel terbentuk dan ditentukan dengan derajat atau tingkat keanggotaannya yang berbeda antara 0 sampai

dengan 1 [13]. Kelebihan dari FCM ini sendiri ialah terletak pada penempatan pusat kluster yang lebih tepat dibanding metode kluster yang lain. Cara untuk menangani pusat kluster agar dapat dilihat bahwa pusat kluster bergerak menuju lokasi yang tepat [14].

Data yang diambil dalam penelitian ini fokus pada pembahasan dan analisis data produksi tanaman perkebunan di Provinsi Jawa Timur pada tahun 2022 yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur [15]. Penelitian ini menggunakan data produksi tanaman perkebunan yang meliputi kelapa, kopi, coklat (kakao), tebu, tembakau dan karet sebagai objek penelitian. Grafik data untuk penelitian bisa dilihat pada Gambar 1:





Gambar 1. Grafik Data Tanaman Perkebunan

Klasterisasi (*Clustering*) adalah sebuah metode yang digunakan untuk mengelompokkan data ke dalam beberapa klaster berdasarkan kesamaan parameter tertentu [16]. Untuk mengetahui kualitas dan kekuatan suatu klaster, digunakanlah uji *silhouette* yang melibatkan perbandingan jarak antara data dalam klaster yang sama dengan data dalam klaster lainnya [17].

Uji *silhouette coefficient* merupakan uji yang dilakukan guna melihat kualitas dan kekuatan dari suatu klaster dengan cara jarak data yang dibandingkan pada klaster yang sama dengan jarak data yang lain [18] seperti pada persamaan 1, dimana $S_{(i)}$ adalah *silhouette coefficient* dengan $a_{(i)}$ adalah rata-rata jarak objek yang berada pada klaster yang sama dan $b_{(i)}$ adalah nilai terkecil dari rata-rata objek pada klaster yang berbeda [19].

$$S_{(i)} = \frac{b_{(i)} - a_{(i)}}{\max\{a_{(i)}, b_{(i)}\}} \quad (1)$$

Langkah-langkah algoritma pengklasteran *Fuzzy C-Means* yang dapat dilihat pada Gambar 2, dimulai dengan langkah pertama yaitu menginputkan data dan diakhiri dengan pengecekan kondisi berhenti saat melakukan pembaruan matriks partisi U . Berikut adalah rincian langkah-langkah tersebut [20]:

- 1) Input data yang akan diklaster X berupa matriks berukuran $n \times m$ (n = jumlah data, m = variabel input tiap data). X_{ij} = data sampel ke i ($i = 1, 2, 3, \dots, n$), variabel input ke j ($j = 1, 2, \dots, m$).
- 2) Tentukan jumlah klaster (c), pangkat (w), maksimum iterasi (\maxIter), *error* terkecil yang diharapkan (ϵ), fungsi objektif awal ($P_0 = 0$), iterasi awal ($t = 1$).
- 3) Bangkitkan bilangan random μ_{ik} , sebagai elemen-elemen matriks partisi awal U dengan $i = 1, 2, \dots, n$ dan $k = 1, 2, 3, \dots, m$. Partisi awal μ_{ik} adalah derajat keanggotaan

yang mengacu pada seberapa besar data yang termasuk dalam sebuah kluster dengan Q_i adalah jumlah nilai derajat keanggotaan.

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (2)$$

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (3)$$

- 4) Hitung pusat kluster ke- k dengan $k = 1, 2, \dots, c$ dan $j = 1, 2, \dots, m$ dimana V_{kj} adalah pusat kluster, X_{ij} adalah data sampel ke- i dan atribut ke- j , i adalah iterasi, μ_{ik} adalah perubahan matriks partisi awal dan w adalah bobot.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w) * X_{ij}}{\sum_{i=1}^n (\mu_{ik})^w} \quad (4)$$

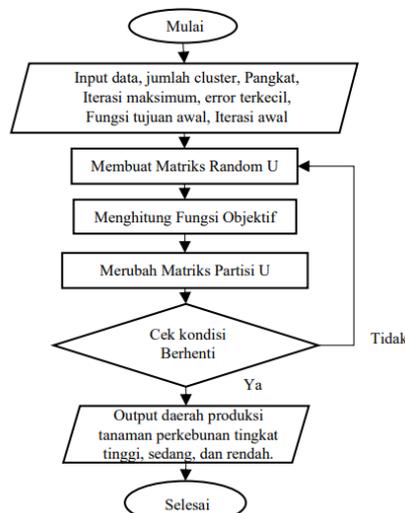
- 5) Hitung fungsi objektif pada iterasi ke- t dengan nilai fungsi objektif yang dinotasikan sebagai P_t , $\sum_{i=1}^n$ adalah jumlah data yang di kluster dan $\sum_{k=1}^c$ adalah jumlah perhitungan kluster awal.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (5)$$

- 6) Hitung perubahan matriks partisi dengan rumusan Persamaan (6) untuk $i = 1, 2, \dots, n$ dan $k = 1, 2, 3, \dots, m$

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \quad (6)$$

- 7) Cek kondisi berhenti pada pembaruan matriks.
- Jika: $(P_t - P_{t-1}) < \epsilon$ atau $(t > \text{maxIter})$ maka berhenti.
 - Jika tidak: $t = t+1$, maka ulangi Langkah ke-4 (persamaan 4)



Gambar 2. Algoritma Clustering

Data yang diinput berupa matriks 38×6 sesuai dengan parameter yang telah ditentukan. Selanjutnya, matriks partisi U dibuat dengan derajat keanggotaan yang memiliki nilai awal 1, sesuai dengan Persamaan (3). Kemudian, fungsi objektif dihitung menggunakan Persamaan (6) dan pembaruan dilakukan sesuai dengan Persamaan (7) untuk meminimalkan kesalahan yang telah ditetapkan seperti yang dijelaskan dalam Persamaan (8). Langkah berikutnya adalah untuk melihat derajat keanggotaan dari matriks U yang telah diperbarui.

3. Hasil dan Pembahasan

Penelitian dilakukan untuk mengklasterkan daerah di Provinsi Jawa Timur yang memiliki potensi produksi tanaman perkebunan pada tahun 2022 menggunakan metode algoritma *Fuzzy C-Means*. Nilai klaster dapat berbeda jika pengklasteran dilakukan ulang karena matriks partisi U awal dibuat secara random. Namun, perubahan tersebut tidak akan signifikan dan tidak akan mempengaruhi keanggotaan klaster.

Langkah awal yang harus dilakukan dalam pengklasteran data menggunakan *Fuzzy C-Means* ialah menentukan komponen penting yang diperlukan. Komponen tersebut meliputi:

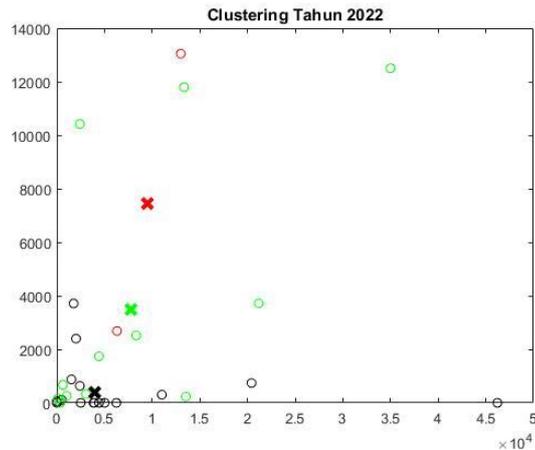
- a) Jumlah dari cluster : 3
- b) Bobot pangkat : 2
- c) Iterasi maksimum : 100
- d) Kemungkinan eror terkecil : 1×10^{-5}
- e) Fungsi tujuan awal : 0
- f) Iterasi awal : 1

Langkah selanjutnya yaitu melakukan pengklasteran daerah produksi tanaman perkebunan pada tahun 2021. Dilakukan proses uji validasi untuk menentukan suatu *cluster* yang optimal. Pada penelitian ini dilakukan uji *silhouette coefficient* dimana nilainya berada pada rentan (-1) sampai dengan (1). Jika nilai indeks yang dihasilkan mendekati 1, maka kualitas dari *cluster* semakin baik.

Tabel 1. Nilai *Silhouette Coefficient*

Cluster	<i>Silhouette Index Value</i>
3	0,6413
4	0,4263
5	0,4864

Berdasarkan Tabel 1, dapat dilihat bahwa klaster 3 memiliki nilai *silhouette* yang paling optimal yaitu 0,6413. Nilai tersebut melebihi 0,5 yang menunjukkan bahwa klaster tersebut telah menghasilkan hasil klasterisasi yang baik. Visualisasi hasil klasterisasi dapat dilihat dalam Gambar 3.



Gambar 3. Hasil Clustering

Berdasarkan Gambar 2 dapat dilihat bahwa simbol (x) melambangkan pusat dari masing-masing kluster. Sementara simbol (o) berwarna merah mewakili data pada kluster 1 dengan tingkatan tertinggi, simbol (o) berwarna hijau mewakili data pada kluster 2 dengan tingkatan sedang dan simbol (o) berwarna hitam mewakili data pada kluster 3 dengan tingkatan terendah. Informasi terkait klusterisasi data produksi tanaman perkebunan di Provinsi Jawa Timur dapat dilihat pada Tabel 2.

Tabel 2. Hasil Clustering Produksi Tanaman Perkebunan

<i>Cluster</i>	<i>Kategori</i>	<i>Kabupaten / Kota</i>
1	Tinggi	Kab. Kediri Kab. Malang
2	Sedang	Kab. Tulungagung Kab. Blitar Kab. Lumajang Kab. Jember Kab. Banyuwangi Kab. Bondowoso Kab. Situbondo Kab. Sidoarjo Kab. Mojokerto Kab. Jombang Kab. Magetan Kab. Ngawi
3	Rendah	Kab. Pacitan Kab. Ponorogo Kab. Trenggalek Kab. Probolinggo Kab. Pasuruan Kab. Nganjuk Kab. Madiun Kab. Bojonegoro Kab. Tuban Kab. Lamongan

Kab. Gresik
Kab. Bangkalan
Kab. Sampang
Kab. Pamekasan
Kab. Sumenep
Kota Kediri
Kota Blitar
Kota Malang
Kota Probolinggo
Kota Pasuruan
Kota Mojokerto
Kota Madiun
Kota Surabaya
Kota Batu

Dari hasil proses pengklasteran terlihat bahwa klaster 1 berada pada bagian paling atas, sementara klaster 3 berada pada bagian paling bawah. Dengan demikian, dapat disimpulkan bahwa klaster 1 merupakan daerah dengan tingkat produksi perkebunan tingkat tinggi, klaster 2 merupakan daerah dengan tingkat produksi sedang dan klaster 3 merupakan daerah dengan tingkat produksi rendah.

4. Kesimpulan

Berdasarkan hasil dan analisis menggunakan algoritma *Fuzzy C-Means* untuk mengklaster daerah produksi tanaman perkebunan dengan jumlah klaster sebanyak 3, ditemukan bahwa Kabupaten Kediri dan Kabupaten Malang merupakan daerah produksi dengan tingkat tinggi dengan nilai *silhouette coefficient* masing-masing sebesar 0,8075 dan 0,9054. Penelitian ini dapat dikatakan berjalan dengan baik karena algoritma *Fuzzy C-Means* berhasil digunakan untuk mengklaster daerah dengan tingkat produksi tanaman perkebunan. Hasil klastering menunjukkan adanya perbedaan tingkat produksi antara Kabupaten atau Kota dengan masing-masing klaster yang mewakili tingkat produksi yang berbeda. Hal ini dapat dijadikan acuan bagi pihak berwenang dalam mengambil kebijakan dan Langkah dalam memproduksi tanaman perkebunan. Selain itu, nilai *silhouette coefficient* yang didapatkan juga cukup tinggi, menunjukkan bahwa klaster yang terbentuk memiliki batas yang jelas dan data dalam setiap klaster memiliki tingkat kesamaan yang tinggi.

Daftar Pustaka

- [1] A. F. Sunartomo, "Kapasitas Penyuluh Pertanian Dalam Upaya Meningkatkan Produktivitas Pertanian di Jawa Timur," *Agriekonomika*, vol. 5, no. 2, 2016, doi: 10.21107/agriekonomika.v5i2.1343.
- [2] M. Chusnah *et al.*, "Peningkatan Daya Saing PKK Desa Gongseng Jombang melalui Pelatihan Pengolahan Hasil Pertanian," vol. 4, no. 1, 2023.
- [3] A. Septya, P. Pradana, and F. T. Industri, "Produksi Tanaman Pangan Kabupaten Kediri Jawa Timur," vol. 3, no. 2, pp. 9–15, 2019.

- [4] B. M. Pratama, M. I. Ali, M. H. Wati, R. Rhonsina, and S. Wulandari, "Kontribusi Subsektor Perkebunan Dalam Mendukung Pertumbuhan Ekonomi Masyarakat di Sumatera Utara," *El-Mal: Jurnal Kajian Ekonomi & Bisnis Islam*, vol. 4, no. 6, pp. 1637–1645, 2023, doi: 10.47467/elmal.v4i6.3079.
- [5] A. Rohmatullah, D. Rahmalia, and M. S. Pradana, "Klasterisasi Data Pertanian di Kabupaten Lamongan Menggunakan Algoritma K-Means dan Fuzzy C-Means," *Jurnal Ilmiah Teknosains*, vol. V, no. 2, pp. 686–93, 2019.
- [6] R. A. Widyowanti, Sunardi, T. Setyorini, and R. A. Renjani, "Pendampingan Pembuatan dan Aplikasi Pelet Pupuk Limbah Biogas untuk Tanaman Perkebunan," 2021.
- [7] A. R. Said, D. Arifianto, and H. A. Al Faruq, "Pengelompokan Kecamatan di Kabupaten Jember Berdasarkan Tanaman Pangan dengan Algoritma Fuzzy C-Means Dan Metode Elbow," *Jurnal Smart Teknologi*, vol. 2, no. 1, pp. 1–12, 2020.
- [8] Faza Adhzima, Yandra Arkeman, and Irman Hermadi, "The Clustering Rice Plant Diseases Using Fuzzy C-Means and Genetic Algorithm," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 2, pp. 240–245, 2022, doi: 10.29207/resti.v6i2.3912.
- [9] N. K. Munadi, A. Fahmi, and U. K. Usman, "Simulasi Fuzzy C-Means Clustering Terhadap Efisiensi Energi Pada Komunikasi D2D Fuzzy C-Means Clustering Simulation Of D2D Communication For Energy Efficiency," vol. 8, no. 6, pp. 3892–3899, 2022.
- [10] S. Agung Priambodo and A. Zakki Falani, "Pemanfaatan Data Mining Untuk Klasterisasi Potensi Produksi Beras di Kabupaten Blitar dengan Menggunakan Metode Fuzzy C-Means," *Nopember*, vol. 12, no. 2, pp. 30–36, 2020.
- [11] Nurhidayati and I. Marzuki, "Deteksi Otomatis Penyakit Daun Jagung Menggunakan Teknik Klasterisasi Data dan Operasi Morfologi," *Energy-Jurnal Ilmiah Ilmu-Ilmu Teknik*, vol. 10, no. 1, pp. 1–6, 2020.
- [12] V. Herlinda and D. Darwis, "Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means," *Darwis, Dartono*, vol. 2, no. 2, pp. 94–99, 2021.
- [13] J. Inayah, D. A. S. N. Maghfiroh, and D. C. R. Novitasari, "Clustering Daerah Rawan Kriminalitas Menggunakan Algoritma Fuzzy C-Means," *Jurnal Ilmiah Informatika Komputer*, vol. 27, no. 2, pp. 95–106, 2022, doi: 10.35760/ik.2022.v27i2.6019.
- [14] 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," ... *Student Journal for ...*, vol. III, pp. 1–8, 2022.
- [15] B. P. Statistik, "Produksi Tanaman Perkebunan Menurut Komoditas dan Kabupaten / Kota," 2023.
- [16] D. Abdullah, S. S. Ansari, S. Ahmar, and R. R. Rahmat, "The Application of K - means Clustering for Province Clustering in Indonesia of The Risk of The COVID - 19 Pandemic Based," *Quality & Quantity*, vol. 56, no. 3, pp. 1283–1291, 2022, doi: 10.1007/s11135-021-01176-w.
- [17] A. Shoolihah, M. T. Furqon, and A. W. Widodo, "Implementasi Metode Improved K-Means untuk Mengelompokkan Titik Panas Bumi," *Universitas Brawijaya*, 2017.

- [18] T. L. Afandi, B. Warsito, and R. Santoso, "Implementasi K-Medoids dan Model Weighted-Length Recency Frequency Monetary (W-LRFM) Untuk Segmentasi Pelanggan Dilengkapi Gui R," *Jurnal Gaussian*, vol. 11, no. 3, pp. 429–438, Jan. 2023, doi: 10.14710/j.gauss.11.3.429-438.
- [19] N. Nur Adifia Nanda, N. Ulinuha, and H. Khaulasari, "Kesejahteraan Sosial Menggunakan Fuzzy C-Means Clustering Grouping of Districts in the Bojonegoro Regency Based on the Type of PMKS Using Fuzzy C-Means Clustering," *Techno.com*, vol. 22, no. 2, pp. 499–507, 2023.
- [20] N. Ulinuha, "Provincial Clustering in Indonesia Based on Plantation Production Using Fuzzy C-Means," *Jurnal Ilmiah Teknologi dan Informasi*, vol. 9, no. 1, pp. 8–12, 2020.