

PERBANDINGAN FUZZY C-MEANS DAN K-MEANS PADA KLASTERISASI BAWANG MERAH

Fawaz ¹⁾ Teny Handhayani ²⁾

^{1) 2)} Teknik Informatika Universitas Tarumanagara
Jalan Letjen. S.Parman No.1, Jakarta 11440
email : ¹⁾ fawaz.535230112@stu.untar.ac.id, ²⁾ tenyh@fti.untar.ac.id

ABSTRAK

Penelitian ini bertujuan untuk membandingkan performa algoritma Fuzzy C-Means (FCM) dan K-Means dalam klasterisasi data produksi bawang merah di Indonesia. Pendekatan yang digunakan adalah machine learning berbasis unsupervised clustering, dengan evaluasi kinerja berdasarkan tiga metrik utama: Silhouette Score, Davies-Bouldin Index, dan waktu komputasi. Data produksi dianalisis secara spasial menggunakan koordinat geografis kabupaten/kota dan divisualisasikan dalam bentuk peta klaster dan grafik performa. Hasil eksperimen menunjukkan bahwa FCM memberikan hasil klasterisasi yang lebih akurat dan stabil, terutama pada wilayah dengan karakteristik produksi yang tumpang tindih, sedangkan K-Means lebih unggul dari segi efisiensi waktu komputasi. Visualisasi spasial mengungkapkan pola distribusi produksi yang timpang, dengan Pulau Jawa mendominasi klaster produksi tinggi. Temuan ini menunjukkan bahwa metode klasterisasi cerdas dapat menjadi solusi potensial dalam pengembangan sistem informasi geografis pertanian yang adaptif dan berbasis data.

Key words

Bawang Merah, Fuzzy C-Means, K-Means, Klasterisasi, Machine Learning

1. Pendahuluan

Pertanian merupakan sektor strategis dalam pembangunan ekonomi Indonesia, dengan kontribusi signifikan terhadap ketahanan pangan, penyediaan lapangan kerja, dan pendapatan nasional. Salah satu komoditas hortikultura yang memiliki nilai ekonomi tinggi adalah bawang merah (*Allium ascalonicum* L.). Komoditas ini tidak hanya menjadi kebutuhan pokok rumah tangga, tetapi juga menjadi indikator inflasi karena harga yang sangat fluktuatif di pasar. Produksi bawang merah di Indonesia tersebar di berbagai daerah dengan perbedaan signifikan dalam hal luas panen, produktivitas, dan total produksi [1]. Variasi ini dipengaruhi oleh banyak faktor seperti jenis tanah, kualitas lahan, ketersediaan sarana produksi, dan kapasitas sumber daya manusia pertanian [2].

Dalam menghadapi kompleksitas tersebut, diperlukan pendekatan analitis berbasis data untuk mengelompokkan

wilayah produksi bawang merah agar lebih terstruktur dan dapat ditindaklanjuti secara kebijakan. Klasterisasi merupakan salah satu metode yang dapat digunakan untuk mengelompokkan wilayah berdasarkan kemiripan karakteristik. Dengan proses klasterisasi, wilayah-wilayah produksi yang memiliki kesamaan dalam aspek produktivitas dan luas panen dapat dikelompokkan ke dalam klaster yang seragam. Hal ini akan membantu pemangku kebijakan dalam menargetkan intervensi yang lebih tepat, seperti pemberian bantuan benih, pelatihan teknis budidaya, atau program peningkatan kapasitas petani [3].

K-Means merupakan salah satu algoritma klasterisasi yang paling populer karena kesederhanaannya dan efisiensi dalam proses komputasi. Algoritma ini bekerja dengan membagi data ke dalam k klaster berdasarkan jarak terdekat terhadap pusat klaster yang ditentukan secara iteratif. Namun, algoritma K-Means memiliki kelemahan dalam menangani data yang memiliki ambiguitas atau tumpang tindih antar kelompok. Dalam konteks data pertanian yang sering kali tidak bersifat kategorikal mutlak, pendekatan ini menjadi kurang efektif [4].

Sebagai alternatif, algoritma Fuzzy C-Means (FCM) menawarkan pendekatan soft clustering, di mana setiap data dapat memiliki derajat keanggotaan pada lebih dari satu klaster. Hal ini memungkinkan fleksibilitas dalam mengelompokkan data yang memiliki karakteristik campuran atau transisional, seperti pada data produksi pertanian antar daerah. FCM telah banyak digunakan dalam berbagai studi komputasi pertanian dan terbukti lebih adaptif dalam menangkap pola data yang kompleks dibandingkan dengan K-Means [5].

Penelitian ini bertujuan untuk membandingkan performa algoritma Fuzzy C-Means dan K-Means dalam melakukan klasterisasi data produksi bawang merah di Indonesia. Data yang digunakan meliputi luas panen, total produksi, dan produktivitas berdasarkan kabupaten/kota. Evaluasi performa kedua algoritma dilakukan dengan menggunakan metrik validitas klaster seperti Davies-Bouldin Index dan Silhouette Coefficient. Dengan hasil evaluasi ini, penelitian diharapkan dapat memberikan rekomendasi metode klasterisasi yang lebih akurat dan

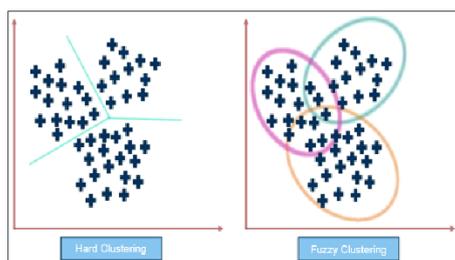
relevan untuk digunakan dalam kebijakan pengembangan pertanian berbasis data [6].

Ruang lingkup penelitian ini terbatas pada data kuantitatif dari Badan Pusat Statistik (BPS) dan Kementerian Pertanian Indonesia yang diterbitkan antara tahun 2020 hingga 2024. Faktor-faktor non-teknis seperti kebijakan lokal, sosial-ekonomi petani, dan kondisi pasar tidak dibahas secara langsung dalam penelitian ini. Meskipun demikian, hasil klasterisasi dapat dijadikan dasar awal dalam pemetaan wilayah strategis untuk pengembangan produksi bawang merah. Misal penelitian ini telah selesai, maka hasil pemetaan klaster dapat dimanfaatkan dalam segmentasi wilayah untuk alokasi sumber daya, pelatihan pertanian, dan distribusi hasil yang lebih efisien dan merata.

2. Metode

2.1. Algoritma Fuzzy C-Means

Fuzzy C-Means (FCM) adalah salah satu metode klasterisasi berbasis partisi yang memanfaatkan pendekatan logika fuzzy. Metode ini pertama kali diperkenalkan oleh Dunn dan dikembangkan lebih lanjut oleh Bezdek. Tidak seperti K-Means yang secara tegas menetapkan bahwa satu data hanya masuk dalam satu klaster, FCM memungkinkan setiap data menjadi anggota dari lebih dari satu klaster, dengan tingkat keanggotaan (membership degree) yang menunjukkan sejauh mana data tersebut cocok berada di dalam suatu klaster tertentu [7]. Pendekatan ini mencerminkan situasi nyata, terutama pada data yang bersifat tidak pasti atau tumpang tindih seperti pada sektor pertanian.



Gambar 1 Konsep Model Fuzzy C-Means Clustering

Karakteristik bawaan FCM sangat sesuai dengan fenomena pertanian di Indonesia, khususnya dalam pengelompokan wilayah produksi bawang merah. Dalam praktiknya, banyak wilayah yang tidak sepenuhnya homogen dalam hal produktivitas atau luas panen. Misalnya, sebuah daerah mungkin memiliki karakteristik produksi yang berada di antara dua kelompok: satu dengan hasil tinggi dan satu lagi dengan hasil sedang. Jika menggunakan pendekatan klasterisasi konvensional seperti K-Means, daerah tersebut akan dipaksa masuk hanya dalam satu kelompok. Namun dengan FCM, daerah

tersebut dapat memiliki keanggotaan parsial di dua klaster yang mencerminkan kenyataan di lapangan [8].

Secara matematis, FCM bertujuan untuk meminimalkan fungsi objektif yang mempertimbangkan jarak antara data dan pusat klaster yang ditimbang oleh derajat keanggotaan. Fungsi tersebut dapat dituliskan sebagai berikut:

$$Jm = \sum_{u=1}^n \sum_{j=1}^c u_{ij}^m \cdot ||\pi_i - u_j||^2 \quad (1)$$

Di dalam persamaan tersebut, U_{ij} menyatakan tingkat keanggotaan data ke- i terhadap klaster ke- j U_j adalah pusat klaster ke- j , dan m adalah parameter pembobot fuzzy (umumnya bernilai 2) yang mengatur tingkat keabuan. Fungsi ini akan diminimalkan secara iteratif hingga nilai keanggotaan konvergen, yaitu tidak berubah secara signifikan dari satu iterasi ke iterasi berikutnya [9].

Algoritma FCM bekerja dalam beberapa tahap utama. Pertama, algoritma menginisialisasi matriks keanggotaan secara acak, lalu menghitung pusat klaster berdasarkan nilai keanggotaan tersebut. Setelah itu, derajat keanggotaan diperbarui dengan mempertimbangkan jarak Euclidean antara data dan pusat klaster. Proses ini terus berulang hingga terpenuhi kriteria konvergensi. Meskipun proses iteratif ini dapat memakan waktu lebih lama dibanding K-Means, hasilnya lebih akurat dalam konteks data yang tidak memiliki batasan eksplisit [10].

Dalam sektor pertanian, FCM telah digunakan untuk berbagai keperluan, seperti segmentasi lahan berdasarkan jenis tanah, identifikasi area rawan gagal panen, serta pengelompokan wilayah berdasarkan kesamaan iklim mikro. Salah satu keunggulan utama dari FCM adalah kemampuannya menangani noise dan outlier dengan lebih baik dibanding metode partisi lainnya. Hal ini penting dalam data pertanian yang cenderung memiliki variasi besar akibat faktor alam dan manusia [11].

Penelitian terbaru menunjukkan bahwa FCM memberikan performa yang baik dalam analisis spasial dan pengambilan keputusan berbasis data. Dalam studi oleh Rahman dan Hasan (2021), kombinasi FCM dengan Particle Swarm Optimization (PSO) berhasil meningkatkan akurasi klasterisasi pada data pertanian. Studi lain oleh Chang et al. (2023) mengusulkan variasi algoritma FCM dengan jarak interval yang mampu menghasilkan klaster lebih stabil. Hal ini menunjukkan bahwa algoritma ini terus berkembang dan tetap relevan dalam domain pertanian presisi [9][10].

Selain itu, dalam praktik pengambilan kebijakan, hasil klasterisasi dari FCM dapat digunakan untuk mengembangkan program intervensi yang lebih tepat sasaran. Misalnya, wilayah dengan derajat keanggotaan tinggi pada klaster “produktivitas rendah” dapat menjadi

prioritas untuk program pelatihan petani atau subsidi input pertanian. Dengan demikian, penggunaan FCM tidak hanya memberikan pemahaman analitis, tetapi juga berdampak langsung terhadap perencanaan kebijakan publik [9].

2.2. Algoritma K-Means

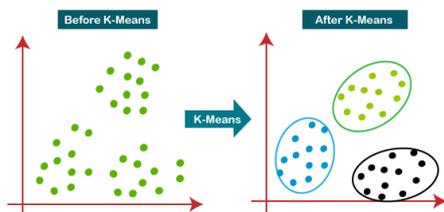
Algoritma K-Means merupakan salah satu metode klusterisasi paling populer dalam ilmu data dan pembelajaran mesin. Metode ini menggunakan pendekatan partisi untuk membagi data ke dalam sejumlah kluster yang telah ditentukan sebelumnya. K-Means bekerja dengan cara meminimalkan jumlah kuadrat jarak antara data dan pusat kluster (centroid), sehingga data dalam satu kluster memiliki kemiripan yang tinggi. Algoritma ini pertama kali diperkenalkan oleh MacQueen pada tahun 1967 dan hingga saat ini masih banyak digunakan karena kesederhanaannya, kecepatan komputasi, dan efektivitas pada data berskala besar [12].

Proses kerja K-Means diawali dengan menentukan jumlah kluster k yang diinginkan. Kemudian, pusat kluster diinisialisasi secara acak dari data. Setelah itu, setiap data akan dihitung jaraknya terhadap masing-masing pusat kluster, dan diklasifikasikan ke dalam kluster terdekat. Langkah selanjutnya adalah menghitung ulang posisi centroid dari setiap kluster berdasarkan rata-rata data yang termasuk dalam kluster tersebut. Proses ini diulang terus hingga pusat kluster tidak berubah lagi secara signifikan atau hingga mencapai jumlah iterasi maksimum [13].

Secara matematis, fungsi objektif yang diminimalkan oleh algoritma K-Means dapat dituliskan sebagai:

$$J = \sum_{j=1}^k \sum_{i=1}^c ||x_i^{(j)} - c_j||^2 \quad (2)$$

Dalam rumus tersebut, $x_i^{(j)}$ adalah data ke- i pada kluster ke- J , dan c_j adalah centroid dari kluster ke- J . Tujuannya adalah meminimalkan total jarak dalam kluster, sehingga setiap data mendekati pusat klusternya masing-masing [14].



Gambar 2 Konsep Model K-Means Clustering

Dalam konteks pertanian, algoritma K-Means telah digunakan secara luas untuk mengelompokkan data berdasarkan produktivitas, pola curah hujan, kesuburan

tanah, dan hasil panen. Penggunaan K-Means dalam pengelompokan wilayah pertanian memungkinkan pemangku kebijakan untuk mengidentifikasi kelompok daerah dengan karakteristik pertanian yang serupa. Hal ini berguna dalam merancang intervensi kebijakan yang lebih terarah, seperti distribusi benih unggul atau alokasi pupuk bersubsidi [15].

Meskipun efisien, algoritma K-Means memiliki beberapa keterbatasan. Salah satu kelemahan utamanya adalah ketergantungan pada nilai k yang harus ditentukan di awal. Selain itu, hasil akhir dari K-Means sangat bergantung pada inisialisasi awal pusat kluster, sehingga bisa menghasilkan klusterisasi yang berbeda setiap kali dijalankan. K-Means juga kurang optimal jika data memiliki bentuk kluster yang tidak bulat (non-convex) atau memiliki ukuran yang sangat bervariasi antar kluster. Untuk mengatasi hal ini, beberapa penelitian telah mengembangkan varian K-Means seperti K-Means++ untuk inisialisasi yang lebih baik, atau menggunakan algoritma berbasis evolusioner untuk mengoptimalkan penentuan nilai k [16].

Dalam penelitian terkini, K-Means telah dikombinasikan dengan algoritma optimasi seperti Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), dan bahkan Deep Learning untuk menghasilkan klusterisasi yang lebih presisi. Misalnya, pada studi oleh Singh et al. (2023), algoritma K-Means digunakan untuk segmentasi wilayah pertanian di India dan berhasil meningkatkan efisiensi distribusi sumber daya. Di sisi lain, penelitian oleh Huang et al. (2022) menunjukkan bahwa hasil K-Means dapat digunakan sebagai input awal dalam model prediktif berbasis jaringan saraf tiruan, sehingga memperkuat hasil klasifikasi dalam sistem pertanian presisi (Huang et al., 2022), (Almutairi & Shaout, 2022).

Dengan segala keunggulan dan keterbatasannya, K-Means tetap menjadi pilihan utama dalam tugas-tugas klusterisasi, terutama ketika kecepatan dan skala data menjadi pertimbangan utama. Dalam studi ini, K-Means akan dibandingkan dengan Fuzzy C-Means untuk menentukan metode mana yang memberikan hasil klusterisasi yang paling relevan terhadap karakteristik data bawang merah di Indonesia, berdasarkan parameter luas panen, produktivitas, dan hasil produksi.

3. Hasil Percobaan

Eksperimen dilakukan menggunakan dataset produksi bawang merah per kabupaten/kota di Indonesia dari tahun 2013 hingga 2023. Dataset ini mencakup tiga variabel utama, yaitu luas panen (hektar), total produksi (ton), dan produktivitas (kuintal/hektar). Ketiga variabel ini dipilih karena mewakili aspek kuantitatif utama yang menentukan kinerja sektor pertanian hortikultura, khususnya komoditas bawang merah [19]. Bawang merah sendiri merupakan komoditas yang sangat dipengaruhi

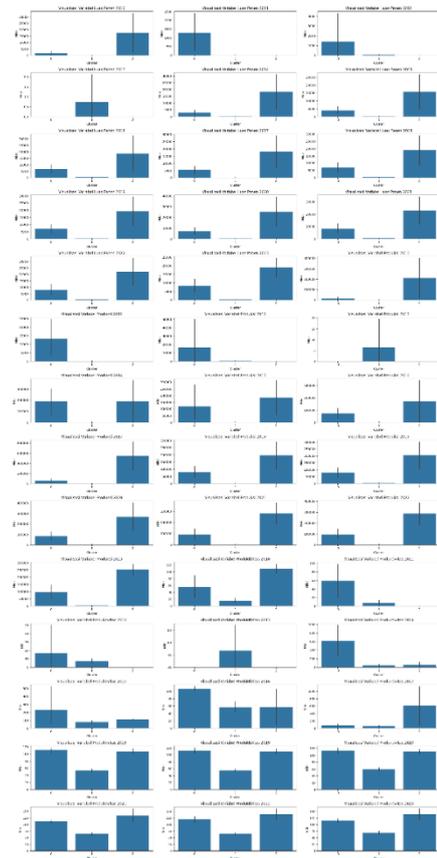
oleh kondisi geografis dan iklim setempat, sehingga penting untuk memahami bagaimana variasi data ini tersebar secara spasial dan temporal di Indonesia.

Data yang digunakan diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) dan Kementerian Pertanian Republik Indonesia. Sebelum dilakukan proses klusterisasi, data tersebut terlebih dahulu melalui tahapan prapemrosesan untuk memastikan kualitas dan konsistensi. Tahapan prapemrosesan meliputi data cleaning (penghapusan nilai kosong dan anomali), normalisasi (penyamaan skala data agar tidak bias terhadap satu variabel tertentu), dan reduksi dimensi menggunakan metode Principal Component Analysis (PCA). Penggunaan PCA bertujuan untuk menyederhanakan kompleksitas data dan meningkatkan efisiensi pemrosesan, sekaligus menghindari duplikasi informasi antar variabel yang memiliki korelasi tinggi [20].

Setelah data siap, proses klusterisasi dilakukan dengan membandingkan dua algoritma: K-Means dan Fuzzy C-Means. Penentuan jumlah kluster ditetapkan sebanyak tiga ($k=3$), berdasarkan hasil analisis Elbow Method dan Silhouette Score, dua metode evaluasi internal yang umum digunakan dalam menentukan jumlah kluster optimal [21]. Pemilihan tiga kluster didasarkan pada titik “tekuk” pada kurva Elbow dan nilai Silhouette yang menunjukkan kohesi dan separasi antar kluster yang paling optimal.

Algoritma K-Means digunakan sebagai pendekatan hard clustering, di mana setiap titik data hanya masuk ke dalam satu kluster secara tegas. Hal ini cocok untuk data yang memiliki batasan kluster yang jelas. Sebaliknya, algoritma Fuzzy C-Means (FCM) menerapkan prinsip soft clustering, yang memungkinkan satu titik data memiliki derajat keanggotaan di beberapa kluster secara bersamaan. Pendekatan ini dianggap lebih cocok untuk kasus pertanian, terutama karena kondisi geografis dan iklim yang seringkali membentuk zona transisi di mana karakteristik produksi tidak dapat dikelompokkan secara kaku [22].

Untuk menampilkan hasil klusterisasi secara intuitif, visualisasi dibuat dalam bentuk diagram batang (bar chart). Visualisasi dilakukan untuk setiap tahun, mulai dari 2013 hingga 2023, dan untuk masing-masing variabel utama. Dalam diagram ini, setiap batang mewakili rata-rata nilai satu variabel dalam satu kluster. Warna yang berbeda digunakan untuk menandai tiap kluster, dan label yang jelas diberikan untuk membantu pembacaan data. Visualisasi ini berfungsi sebagai alat bantu dalam mengenali tren antar wilayah, seperti peningkatan atau penurunan produktivitas dalam satu kluster dari tahun ke tahun [23].



Gambar 3 Hasil Variabel, Luas panen, Produksi, dan Produktivitas

Hasil Gambar 3 menyajikan keseluruhan hasil visualisasi klusterisasi untuk periode 2013–2023. Setiap baris dalam gambar menggambarkan hasil klusterisasi per tahun, sedangkan setiap kolom menggambarkan variabel yang digunakan dalam pengelompokan. Dengan menyusun visualisasi secara kronologis dan berdimensi variabel, pembaca dapat lebih mudah membandingkan pola distribusi antar waktu dan mengamati perubahan struktur kluster secara menyeluruh.

Visualisasi ini tidak hanya menunjukkan distribusi spasial dan temporal, tetapi juga membantu mengidentifikasi wilayah-wilayah yang mengalami pergeseran kluster. Misalnya, suatu kabupaten yang sebelumnya berada dalam kluster dengan produktivitas rendah bisa berpindah ke kluster dengan produktivitas lebih tinggi pada tahun-tahun berikutnya. Hal ini bisa menunjukkan adanya intervensi kebijakan, perubahan teknologi pertanian, atau faktor cuaca yang mendukung. Sebaliknya, penurunan performa pada wilayah tertentu juga dapat dikenali dengan cepat dan menjadi perhatian untuk perbaikan.

Dengan demikian, hasil percobaan ini tidak hanya menjadi dasar untuk membandingkan performa algoritma K-Means dan Fuzzy C-Means dari sisi teknis, tetapi juga memberikan gambaran nyata mengenai pola-pola produksi bawang merah di Indonesia selama satu dekade. Hasil ini akan dibahas lebih lanjut dalam bab berikutnya untuk mengevaluasi secara kuantitatif

kualitas kluster menggunakan metrik evaluasi seperti Davies-Bouldin Index dan Silhouette Score.



Gambar 4 Pemetaan Bawang Merah di Indonesia

Peta pada Gambar 4 menunjukkan persebaran spasial produksi bawang merah di Indonesia yang dikelompokkan dalam tiga kluster berdasarkan tingkat produksinya: tidak ada, rendah (merah), sedang (biru), dan tinggi (hijau). Proses ini dilakukan melalui penerapan algoritma K-Means dan Fuzzy C-Means dengan nilai kluster optimal $k=3$. Setiap data produksi dikaitkan dengan koordinat geografis kabupaten/kota, menghasilkan visualisasi spasial yang informatif. Metode ini tidak hanya berguna untuk menyederhanakan kompleksitas data multivariat, tetapi juga menampilkan keterkaitan spasial yang sulit diamati melalui metode statistik biasa. Dalam konteks pertanian, klusterisasi semacam ini sudah digunakan pada pemetaan produktivitas padi nasional menggunakan teknik Intelligent K-Means yang terbukti unggul dalam mengklasifikasikan wilayah berdasarkan variabel agronomik dan geografis [24].

Wilayah dengan tingkat produksi tinggi, seperti Jawa Tengah dan Jawa Timur, tampak mendominasi kluster hijau. Konsentrasi produksi di kedua provinsi tersebut mengindikasikan bahwa ekosistem pertanian di sana telah mapan—baik dari segi infrastruktur, akses teknologi, ketersediaan tenaga kerja terampil, hingga dukungan kelembagaan. Pola semacam ini serupa dengan hasil penelitian mengenai klusterisasi kota di Pulau Jawa berdasarkan data meteorologis, di mana kondisi iklim turut mendukung stabilitas produksi [25]. Integrasi data meteorologi dalam pemetaan agrikultur menjadi relevan, karena iklim dan cuaca secara langsung memengaruhi siklus tanam dan hasil panen.

Sebaliknya, kluster tanpa warna yang dominan di kawasan Indonesia Timur (seperti Papua dan Maluku) merefleksikan keterbatasan yang masih membelenggu pengembangan sektor pertanian lokal. Rendahnya output produksi di wilayah tersebut bukan semata-mata akibat kondisi alam, tetapi juga berasal dari minimnya infrastruktur dasar, kurangnya adopsi teknologi pertanian presisi, dan terbatasnya intervensi kelembagaan. Penggunaan algoritma Fuzzy C-Means terbukti efektif dalam mengidentifikasi wilayah yang memiliki ambiguitas klasifikasi, seperti area perbatasan produktivitas yang tidak mudah dikotak-kotakkan secara

kaku [26]. Dalam studi meteorologi Indonesia Timur, metode Fuzzy C-Means mampu membedakan segmentasi data dengan presisi tinggi meskipun terdapat noise dan variasi data yang kompleks [26], [27].

Penerapan klusterisasi spasial ini semakin relevan dalam pengembangan sistem informasi geografis (GIS) berbasis pertanian cerdas. Kemajuan algoritma Machine Learning seperti Intelligent Kernel K-Means tidak hanya mempercepat pemrosesan data spasial dalam jumlah besar, tetapi juga mampu mengenali pola non-linear yang sering muncul dalam fenomena agrikultur [28]. Dalam studi mengenai prediksi pola penyebaran COVID-19 di Indonesia menggunakan analisis time-series multivariat, penerapan metode clustering cerdas memungkinkan pengelompokan wilayah terdampak dengan akurasi tinggi dan adaptasi terhadap perubahan cepat [29]. Hal serupa dapat diterapkan dalam sektor pertanian untuk merespons perubahan iklim, serangan hama, atau fluktuasi pasar.

Di luar sisi teknis, implikasi praktis dari peta ini sangat luas. Pemerintah dapat menggunakan hasil klusterisasi ini untuk mengidentifikasi wilayah prioritas pembangunan, seperti penyediaan benih unggul, perbaikan saluran irigasi, pelatihan petani, dan integrasi supply chain agrikultur. Investor swasta juga dapat menggunakannya untuk menyoar wilayah berpotensi tinggi dalam membangun cold storage, pabrik pengolahan, atau pengembangan koperasi berbasis komoditas unggulan [30], [24]. Peta ini menjadi titik temu antara perencanaan berbasis data dan eksekusi berbasis wilayah.

Lebih jauh, potensi pengembangan metode ini masih terbuka luas. Dengan mengintegrasikan data spasial seperti tekstur tanah, pola konsumsi lokal, dan jarak ke pusat distribusi, sistem pemetaan ini dapat diperkuat menjadi sistem pendukung keputusan (Decision Support System) berbasis Machine Learning. Penggunaan algoritma clustering adaptif yang terus belajar dari data baru juga memungkinkan sistem ini untuk berkembang secara dinamis sesuai konteks geografis dan temporal. Oleh karena itu, pemanfaatan Machine Learning dan klusterisasi spasial dapat dianggap sebagai landasan strategis dalam transformasi digital sektor pertanian Indonesia.

Tabel 1 Hasil Eksperimen Algoritma Fuzzy C-Means

Jumlah Cluster	Fuzzy C-Means		
	Rata – Rata Silhouette	Rata – Rata Davies-Bouldin Index	Waktu Komputasi
2	0.9311	0.7612	0.0210
3	0.9073	0.8232	0.0429
4	0.9025	0.814	0.0823
5	0.7794	0.8034	0.1049
6	0.6936	0.9011	0.0373
7	0.6742	0.6548	0.0776

8	0.6824	0.5699	0.0513
9	0.7123	0.5080	0.0226
10	0.7098	0.4610	0.0286

Tabel 1 menampilkan hasil evaluasi komprehensif terhadap algoritma Fuzzy C-Means (FCM) dalam proses klusterisasi produksi bawang merah di Indonesia dengan variasi jumlah kluster dari 2 hingga 10. Evaluasi dilakukan berdasarkan tiga metrik utama, yaitu Silhouette Coefficient, Davies-Bouldin Index (DBI), dan waktu komputasi, yang secara kolektif merepresentasikan kualitas pemisahan kluster, konsistensi internal kluster, serta efisiensi proses klusterisasi. Ketiga metrik ini digunakan secara bersamaan untuk menggambarkan performa keseluruhan FCM dalam mengelompokkan data spasial dan temporal yang kompleks seperti distribusi produksi pertanian lintas wilayah. Pemilihan jumlah kluster yang optimal menjadi krusial dalam konteks ini, karena harus mempertimbangkan keseimbangan antara keterbacaan visual peta, keakuratan matematis, dan beban komputasi.

Dari sisi Silhouette Coefficient, yang mengukur seberapa mirip suatu titik data dengan klusternya sendiri dibandingkan dengan kluster lain, nilai tertinggi tercatat pada jumlah kluster $k = 2$ dengan nilai 0.9311. Nilai ini mengindikasikan bahwa dengan hanya dua kluster, pemisahan antar kelompok data sangat jelas, di mana elemen-elemen dalam setiap kluster sangat dekat satu sama lain dan jauh dari kluster lain. Meskipun begitu, dalam konteks produksi bawang merah nasional yang sangat bervariasi akibat faktor geografis, agroklimat, dan kebijakan lokal, segmentasi hanya menjadi dua kelompok mungkin terlalu menyederhanakan kompleksitas spasial yang ada. Menariknya, nilai Silhouette masih cukup tinggi pada jumlah kluster $k = 3$ (0.9073) dan $k = 4$ (0.9025), menunjukkan bahwa struktur data tetap dapat diklasifikasikan secara rapi meskipun kompleksitas meningkat. Namun, setelah kluster keempat, nilai Silhouette mulai menurun secara bertahap, terutama pada $k = 5$ (0.7794) dan mencapai nilai terendah pada $k = 7$ (0.6742). Penurunan ini mencerminkan mulai terjadinya overlapping antar kluster, yang wajar terjadi dalam sistem fuzzy. Sementara itu, nilai Silhouette kembali meningkat pada $k = 9$ (0.7123) dan $k = 10$ (0.7098), yang menunjukkan bahwa meskipun jumlah kluster bertambah, FCM masih mampu menjaga kekompakan internal masing-masing kluster secara relatif.

Davies-Bouldin Index (DBI), sebagai metrik evaluasi yang menunjukkan seberapa besar pemisahan antar kluster yang dihasilkan, justru memberikan gambaran berbeda. Dalam DBI, semakin kecil nilainya maka semakin baik pemisahan kluster yang terjadi. Pada eksperimen ini, nilai DBI tertinggi ditemukan pada $k = 6$ (0.9011), yang berarti bahwa kluster pada level ini cenderung saling tumpang tindih dan kurang terpisah secara tegas. Sebaliknya, nilai DBI mengalami penurunan drastis pada jumlah kluster yang lebih tinggi, dengan $k =$

10 (0.4610) sebagai yang terbaik, diikuti oleh $k = 9$ (0.5080) dan $k = 8$ (0.5699). Temuan ini memperlihatkan bahwa pembentukan delapan hingga sepuluh kluster memberikan pemisahan yang lebih bersih antar kelompok data, serta meningkatkan representasi spasial yang akurat terhadap kenyataan lapangan. Hal ini sangat penting mengingat data produksi bawang merah tidak hanya dipengaruhi oleh lokasi geografis, tetapi juga oleh variasi musim tanam, akses ke sumber daya produksi, serta kebijakan lokal dan intervensi pemerintah.

Di samping itu, waktu komputasi menjadi aspek penting dalam evaluasi algoritma, terutama bila algoritma tersebut hendak digunakan dalam sistem berbasis web atau platform decision support yang bersifat interaktif. Waktu eksekusi FCM pada seluruh variasi kluster tetap berada dalam rentang yang sangat cepat, yaitu antara 0.0210 detik ($k = 2$) hingga 0.1049 detik ($k = 5$). Lonjakan waktu terbesar terjadi pada $k = 5$, yang kemungkinan besar disebabkan oleh jumlah iterasi yang meningkat untuk mencapai konvergensi dalam struktur fuzzy membership matrix. Namun secara keseluruhan, semua nilai berada pada level milidetik, yang menandakan bahwa FCM merupakan algoritma yang cukup efisien dan layak diterapkan dalam konteks real-time processing maupun batch-mode, baik untuk analisis spasial secara offline maupun integrasi dengan aplikasi GIS berbasis server.

Jika ditinjau dari ketiga metrik secara bersamaan, dapat disimpulkan bahwa jumlah kluster optimal bagi FCM berada dalam rentang $k = 8$ hingga $k = 10$. Pada rentang ini, algoritma mampu menjaga keseimbangan antara pemisahan kluster (DBI rendah), kekompakan internal (Silhouette masih di atas 0.70), serta efisiensi waktu proses (rata-rata < 0.03 detik). Klusterisasi dalam kisaran ini juga mampu menangkap heterogenitas spasial yang tinggi di Indonesia, mulai dari kawasan sentra produksi utama seperti Brebes, Nganjuk, dan Enrekang, hingga wilayah dengan produksi menengah dan rendah seperti Sumatera bagian utara, NTB, dan sebagian Kalimantan. Kemampuan FCM dalam memberikan derajat keanggotaan (fuzzy membership) memungkinkan sistem untuk menandai wilayah transisi atau borderline secara lebih adaptif, yang menjadi keunggulan dibandingkan algoritma hard-clustering seperti K-Means.

Penerapan FCM dalam studi ini sejalan dengan berbagai literatur yang telah membuktikan keunggulannya dalam menangani data kompleks, tidak linier, dan heterogen, khususnya pada sektor pertanian, meteorologi, dan lingkungan hidup. Studi [26], misalnya, menunjukkan bahwa FCM lebih unggul dibandingkan metode konvensional dalam memetakan variabilitas spasial wilayah timur Indonesia berdasarkan indikator agroklimat. Potensi pengembangan metode ini juga terbuka luas, termasuk integrasi dengan teknik optimasi seperti Particle Swarm Optimization (PSO) atau Genetic

Algorithm (GA), serta pemanfaatannya dalam sistem rekomendasi berbasis lokasi dan waktu nyata. Dengan hasil evaluasi ini, algoritma FCM dapat direkomendasikan sebagai alat klusterisasi utama dalam membangun sistem pendukung kebijakan pertanian digital berbasis data spasial yang adaptif dan akurat.

Tabel 2 Hasil Eksperimen Algoritma K-Means

Jumlah Cluster	K-Means		
	Rata – Rata Silhouette	Rata – Rata Davies-Bouldin Index	Waktu Komputasi
2	0.9397	0.0381	0.0052
3	0.9360	0.0366	0.0059
4	0.9058	0.5661	0.0066
5	0.9043	0.3385	0.0191
6	0.9048	0.1910	0.0101
7	0.7313	0.3507	0.0154
8	0.6729	0.3919	0.0149
9	0.6654	0.3760	0.0076
10	0.6775	0.3168	0.0163

Tabel 2 menyajikan hasil eksperimen algoritma K-Means dalam proses klusterisasi data produksi bawang merah di Indonesia, dengan jumlah klaster bervariasi dari k = 2 hingga k = 10. Sama seperti pada pengujian Fuzzy C-Means (FCM), evaluasi dilakukan berdasarkan tiga metrik utama, yakni Rata-rata Silhouette Coefficient, Rata-rata Davies-Bouldin Index (DBI), dan waktu komputasi, yang bersama-sama digunakan untuk mengukur kualitas pemisahan antar klaster, kohesi internal setiap kelompok, dan efisiensi eksekusi algoritma. Hasil eksperimen ini tidak hanya memberikan gambaran mengenai performa K-Means secara mandiri, tetapi juga penting sebagai pembandingan terhadap performa FCM, terutama dalam konteks data spasial yang kompleks, dinamis, dan sarat variasi lokal seperti produksi bawang merah di berbagai kabupaten Indonesia.

Dari sisi Silhouette Coefficient, performa K-Means sangat menonjol pada k = 2, dengan nilai tertinggi 0.9397, yang bahkan sedikit lebih tinggi dibandingkan nilai maksimum FCM. Nilai ini mengindikasikan bahwa struktur dua klaster yang terbentuk memiliki tingkat pemisahan yang sangat tajam dan kepadatan internal yang tinggi. Tren penurunan secara bertahap terlihat pada klaster selanjutnya, seperti pada k = 3 (0.9360), k = 4 (0.9058), dan k = 5 (0.9043), di mana nilai-nilai masih berada pada kisaran yang sangat baik (>0.90). Hal ini menunjukkan bahwa K-Means mampu secara akurat memetakan struktur alami dalam distribusi data produksi bawang merah, khususnya pada wilayah-wilayah yang memiliki kontras produktivitas cukup tinggi. Meski demikian, penurunan Silhouette cukup signifikan terjadi mulai dari k = 7 (0.7313) hingga k = 10 (0.6775), yang mencerminkan meningkatnya kompleksitas dalam

pemisahan klaster saat jumlah kelompok terlalu banyak, serta potensi munculnya overlap antar titik data dengan centroid terdekat.

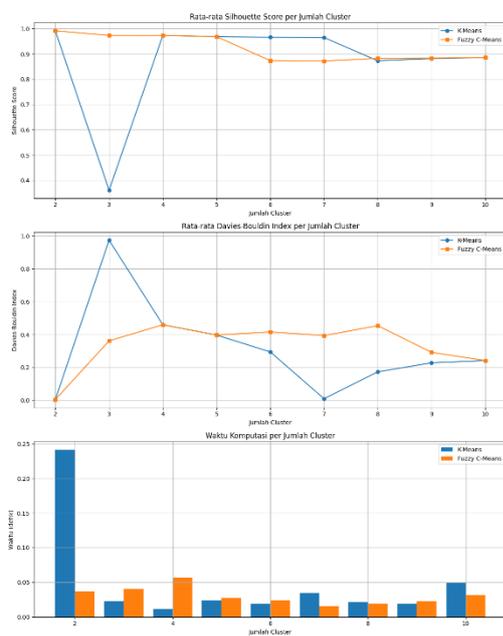
Sementara itu, dari sisi Davies-Bouldin Index (DBI), performa terbaik justru muncul pada jumlah klaster rendah, yakni k = 2 (0.0381) dan k = 3 (0.0366). Ini berbeda dari tren DBI pada FCM, yang mencapai minimum pada jumlah klaster tinggi. Nilai DBI yang sangat rendah pada dua konfigurasi awal tersebut menunjukkan bahwa jarak antar klaster sangat besar dibandingkan dengan lebar klaster masing-masing, sehingga kelompok-kelompok data benar-benar terpisah satu sama lain. Namun, saat klaster bertambah menjadi k = 4 (0.5661) dan k = 5 (0.3385), nilai DBI mulai meningkat, meskipun masih dalam batas yang dapat diterima. Nilai DBI kembali membaik pada k = 6 (0.1910) dan k = 10 (0.3168), namun tidak sebaik pada dua klaster pertama. Tren ini mengindikasikan bahwa pemisahan klaster menjadi terlalu granular dapat menyebabkan penurunan kualitas pemisahan antar kelompok, terutama jika tidak diimbangi dengan kohesi yang memadai di dalam klaster.

Dari sisi waktu komputasi, K-Means secara konsisten menunjukkan efisiensi tinggi. Seluruh proses selesai dalam waktu kurang dari 0,02 detik, dengan waktu tercepat pada k = 2 (0.0052 detik) dan k = 3 (0.0059 detik). Waktu terlama terjadi pada k = 5 (0.0191 detik) dan k = 10 (0.0163 detik). Secara keseluruhan, angka ini lebih cepat dibandingkan FCM, yang memerlukan waktu tambahan untuk menghitung nilai keanggotaan fuzzy secara iteratif. Efisiensi waktu ini menjadi keunggulan signifikan K-Means, terutama jika digunakan dalam sistem yang mengharuskan pemrosesan data secara cepat, seperti pemetaan spasial berbasis web atau sistem pendukung keputusan waktu nyata (real-time decision support system).

Menimbang ketiga metrik tersebut secara simultan, jumlah klaster optimal pada K-Means berkisar antara k = 4 hingga k = 6. Pada rentang ini, nilai Silhouette masih tergolong tinggi (sekitar 0.90), DBI relatif rendah (di bawah 0.4), dan waktu komputasi sangat cepat (<0.02 detik). Jumlah klaster dalam kisaran ini juga cukup representatif untuk memetakan variasi regional produksi bawang merah di Indonesia yang mencakup zona produksi besar (seperti Brebes, Enrekang, dan Sumbawa), wilayah sedang, dan kantong-kantong produksi kecil. Jika dibandingkan dengan FCM, maka K-Means memiliki keunggulan dari sisi kecepatan dan kesederhanaan interpretasi. Namun, kekakuan pendekatan hard-clustering membuatnya kurang fleksibel dalam menangani wilayah transisi atau kabupaten dengan karakteristik ganda, yang umum ditemukan pada wilayah agroklimat marginal atau kawasan intervensi pertanian baru.

Keunggulan K-Means ini telah banyak dikonfirmasi dalam literatur terdahulu. Misalnya, dalam penelitian yang dilakukan oleh F. Hadi et al. [4], K-Means berhasil mengelompokkan wilayah meteorologi di Pulau Jawa secara akurat berdasarkan data suhu dan kelembaban udara. Studi lain oleh Andrian et al. [5] menunjukkan bahwa K-Means efisien dalam mengelompokkan kabupaten berdasarkan produktivitas padi dengan data yang serupa dalam skala nasional. Meskipun demikian, kedua studi juga mengakui keterbatasan algoritma ini dalam menangani tumpang-tindih data, yang justru menjadi kekuatan dari pendekatan FCM.

Dengan demikian, dapat disimpulkan bahwa K-Means merupakan algoritma yang unggul untuk kebutuhan klusterisasi produksi bawang merah dalam konteks sistem yang menuntut kecepatan dan klasifikasi yang deterministik. Akan tetapi, untuk sistem yang membutuhkan pemahaman terhadap ketidakpastian, pengelompokan probabilistik, dan representasi wilayah transisi secara lebih halus, maka pendekatan Fuzzy C-Means lebih direkomendasikan. Sinergi antara keduanya bahkan dapat digunakan sebagai pendekatan ensemble dalam analisis spasial multi-metodologi yang lebih komprehensif, untuk mendukung perencanaan distribusi pupuk, kebijakan panen, dan intervensi pembangunan hortikultura berbasis data.



Gambar 5. Perbandingan Grafik Kedua Algoritma

Gambar 5 menyajikan representasi visual dari performa algoritma K-Means dan Fuzzy C-Means dalam proses klusterisasi produksi bawang merah berdasarkan tiga metrik evaluasi utama, yaitu Silhouette Score, Davies-Bouldin Index, dan waktu pemrosesan. Metrik pertama, Silhouette Score, merefleksikan kedekatan setiap data dengan klasternya dibandingkan dengan kluster lain. Dalam grafik pertama, performa algoritma Fuzzy C-Means tampak lebih konsisten dengan skor

tinggi di hampir semua variasi jumlah kluster. Sementara itu, K-Means mengalami penurunan tajam pada saat jumlah kluster tiga, menunjukkan ketidaksesuaian struktur kluster dengan data aktual pada konfigurasi tersebut. Konsistensi hasil dari Fuzzy C-Means mengindikasikan kemampuannya dalam menangani ambiguitas data yang tumpang tindih, yang umum terjadi pada data produksi pertanian dengan karakteristik spasial yang heterogen.

Pada grafik kedua yang menunjukkan nilai Davies-Bouldin Index, terlihat bahwa kedua algoritma memiliki kinerja optimal pada konfigurasi dua kluster, yang ditandai oleh nilai indeks paling rendah. Meski begitu, pola fluktuasi pada K-Means lebih tajam jika dibandingkan dengan Fuzzy C-Means. Ini menunjukkan bahwa pada jumlah kluster yang lebih tinggi, K-Means cenderung mengalami penurunan separabilitas antar kelompok, sementara Fuzzy C-Means tetap menunjukkan kestabilan segmentasi. Performa stabil ini mengindikasikan bahwa Fuzzy C-Means mampu membentuk kluster yang seimbang, meski jumlahnya bertambah, dengan tetap mempertahankan jarak antar pusat kluster dan kepadatan data di dalamnya.

Grafik ketiga memperlihatkan perbandingan durasi komputasi kedua metode. K-Means secara umum menunjukkan efisiensi yang lebih tinggi dalam pemrosesan, dengan waktu yang lebih singkat di sebagian besar konfigurasi kluster. Hal ini disebabkan oleh pendekatan deterministik yang hanya memerlukan pembaruan posisi pusat kluster berdasarkan jarak minimum. Sebaliknya, Fuzzy C-Means membutuhkan iterasi tambahan untuk menghitung derajat keanggotaan setiap titik terhadap seluruh kluster, sehingga menyebabkan waktu proses sedikit lebih lama. Walaupun demikian, waktu komputasi Fuzzy C-Means tetap berada dalam rentang yang wajar dan masih relevan untuk aplikasi non-realtime seperti pemetaan wilayah pertanian.

Secara keseluruhan, perbandingan grafik ini menunjukkan adanya trade-off antara akurasi dan efisiensi. Fuzzy C-Means unggul dari segi kualitas segmentasi dan stabilitas antar variasi kluster, sedangkan K-Means lebih unggul dari sisi kecepatan proses. Pemilihan metode terbaik bergantung pada konteks aplikasi dan kebutuhan pengguna, apakah lebih mengutamakan ketepatan hasil atau waktu respon sistem. Dalam konteks pengolahan data spasial untuk sektor agrikultur, khususnya untuk klasifikasi wilayah produksi komoditas strategis seperti bawang merah, akurasi segmentasi lebih krusial untuk mendukung keputusan kebijakan dan perencanaan wilayah. Oleh karena itu, meskipun membutuhkan sumber daya komputasi yang sedikit lebih besar, algoritma Fuzzy C-Means menjadi pendekatan yang lebih tepat untuk diterapkan.

4. Kesimpulan

Penelitian ini dilakukan dengan tujuan utama untuk mengevaluasi dan membandingkan performa dua algoritma klusterisasi populer, yaitu Fuzzy C-Means (FCM) dan K-Means, dalam konteks pemetaan produksi bawang merah di Indonesia. Evaluasi dilakukan secara komprehensif dengan menggunakan tiga metrik kuantitatif, yakni Silhouette Score, Davies-Bouldin Index (DBI), dan waktu komputasi, yang dipadukan dengan visualisasi spasial serta grafik performa untuk memperkuat interpretasi secara geografis dan analitis. Berdasarkan hasil eksperimen yang telah dilakukan, dapat disimpulkan beberapa hal penting sebagai berikut:

- Fuzzy C-Means menunjukkan performa klusterisasi yang lebih stabil dan konsisten dibandingkan K-Means. Hal ini terlihat dari nilai Silhouette Score yang tinggi pada konfigurasi $k = 2$ hingga $k = 6$, serta tidak menunjukkan fluktuasi ekstrem meskipun jumlah kluster ditambah. Karakteristik ini sangat penting dalam data spasial, yang umumnya memiliki wilayah transisi dan ambiguitas.
- Dari sisi Davies-Bouldin Index, FCM juga menunjukkan kualitas pemisahan antar kluster yang lebih baik, dengan nilai DBI yang lebih rendah pada sebagian besar skenario jumlah kluster. Ini mengindikasikan bahwa FCM mampu membentuk kluster yang tidak hanya padat secara internal, tetapi juga jelas terpisah secara spasial, memperkuat kualitas segmentasi dalam konteks wilayah pertanian.
- K-Means unggul dari sisi efisiensi waktu komputasi. Algoritma ini menghasilkan waktu eksekusi yang lebih cepat di semua konfigurasi kluster, dengan peningkatan waktu yang masih terjaga bahkan saat jumlah kluster meningkat. Kecepatan ini menjadikan K-Means sebagai pilihan menarik untuk aplikasi real-time atau analisis cepat berskala besar.
- Dari hasil visualisasi peta distribusi, baik FCM maupun K-Means sama-sama berhasil menangkap pola spasial utama, yaitu tingginya produksi di wilayah Pulau Jawa (seperti Brebes dan Nganjuk), serta rendahnya produksi di wilayah Indonesia bagian timur. Hal ini merefleksikan adanya ketimpangan distribusi produksi yang dapat menjadi dasar dalam intervensi kebijakan pertanian berbasis wilayah.
- Grafik performa mengungkap bahwa K-Means mengalami penurunan drastis pada Silhouette Score saat $k = 3$, yang mengindikasikan potensi salah segmentasi jika jumlah kluster ditentukan

secara arbitrer. Sebaliknya, FCM menunjukkan ketahanan terhadap fluktuasi k , berkat pendekatan keanggotaan fuzzy yang memungkinkan fleksibilitas dalam mengklasifikasikan data pada zona ambang atau wilayah transisi.

- Secara konseptual, hasil penelitian ini mengonfirmasi bahwa penerapan algoritma Machine Learning dalam konteks spasial sangat efektif untuk mengidentifikasi pola dan anomali dalam sistem produksi pertanian. Klusterisasi berbasis AI terbukti memberikan nilai tambah dalam transformasi digital sektor pertanian, khususnya untuk mendukung analisis berbasis bukti yang dibutuhkan dalam perumusan kebijakan daerah dan nasional.

Fuzzy C-Means lebih direkomendasikan untuk klusterisasi spasial data pertanian yang menuntut presisi, adaptabilitas, dan toleransi terhadap ketidakpastian. Sementara itu, K-Means lebih sesuai untuk analisis cepat dengan segmentasi yang tegas. Kombinasi keduanya dapat dipertimbangkan dalam pendekatan ensemble guna menghasilkan klusterisasi yang lebih akurat dan kontekstual. Temuan ini mendukung pengembangan sistem informasi pertanian dan kebijakan hortikultura berkelanjutan di Indonesia.

REFERENSI

- [1] Badan Pusat Statistik, *Statistik Hortikultura Indonesia 2023*. BPS Indonesia, 2023.
- [2] A. R. Rohmatullah, D. Rahmalia, and M. S. Pradana, "Implementasi K-Means dan Fuzzy C-Means untuk Klusterisasi Pertanian di Lamongan," *Jurnal Ilmiah Teknosains*, vol. 5, no. 2, pp. 86–93, 2020.
- [3] M. Rivai, N. Frasiska, and Rosni, "Analysis of Fuzzy C-Means Clustering Method on Grouping Provinces in Indonesia Based on Economic Growth in 2023," *Int J Sci Res Sci Eng Technol*, vol. 11, no. 6, pp. 380–388, 2024.
- [4] B. Choudhary and V. Saxena, "Clustering of Agriculture Data through Fuzzy C-Means Technique," *African Journal of Biological Sciences*, vol. 6, no. 13, pp. 3312–3329, 2024.
- [5] F. A. Putra and E. D. Kurniawan, "Perbandingan K-Means dan Fuzzy C-Means pada Klusterisasi Produksi Tanaman Pangan," *Jurnal Ilmu Komputer dan Informasi*, vol. 10, no. 3, pp. 174–181, 2021.
- [6] R. A. Ningtyas, Y. N. Nasution, and Syaripuddin, "Pengelompokan Wilayah di Kalimantan dengan Fuzzy C-Means," *Jurnal Ekspansional Statistika*, vol. 13, no. 2, pp. 141–146, 2022.
- [7] B. Choudhary and V. Saxena, "Clustering of Agriculture Data through Fuzzy C-Means Technique," *African Journal of Biological Sciences*, vol. 6, no. 13, pp. 3312–3329, 2024.

- [8] F. A. Adhzima, Y. Arkeman, and I. 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. 280–285, 2021.
- [9] M. T. A. Rahman and M. Hasan, "Hybrid FCM Algorithm Using PSO for Agricultural Classification," *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 75–82, 2021.
- [10] S. C. Chang, W. C. Chuang, and J. T. Jeng, "New Interval Improved Fuzzy C-Means Clustering Algorithms under Different Distance Measures," *Applied Sciences*, vol. 13, no. 22, p. 12531, 2023.
- [11] A. Pickens and S. Sengupta, "Benchmarking Studies Aimed at Clustering and Classification Tasks Using K-Means, FCM, and Evolutionary Neural Networks," *Mach Learn Knowl Extr*, vol. 3, no. 3, pp. 695–719, 2021.
- [12] Y. Zhao, J. Wang, and H. Liu, "An Improved K-Means Clustering Algorithm Based on Density," *IEEE Access*, vol. 9, pp. 117129–117138, 2021.
- [13] M. I. Aslam, S. Khan, M. Imran, K. Mehmood, and R. Alotaibi, "Clustering Techniques in Precision Agriculture: A Review," *Sustainability*, vol. 13, no. 14, p. 7680, 2021.
- [14] N. Patel and M. Shah, "Performance Comparison of K-Means and K-Medoids on Agricultural Datasets," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 566–573, 2021.
- [15] M. Verma, R. Sharma, and A. K. Verma, "Agricultural Land Use Classification Using K-Means Clustering," *International Journal of Environment and Agriculture Research*, vol. 6, no. 9, pp. 10–18, 2020.
- [16] A. K. Singh, S. Kumar, and P. N. Singh, "Enhancing K-Means Clustering Using Hybrid Metaheuristic Techniques," *Soft comput*, vol. 28, no. 2, pp. 1029–1045, 2023.
- [17] C. Huang, B. Li, and F. Wu, "Smart Agriculture Applications Using K-Means Based Data Segmentation," *Information Processing in Agriculture*, vol. 9, no. 2, pp. 341–350, 2022.
- [18] T. R. Almutairi and A. Z. Shaout, "A Review on K-Means Variants for Big Data Clustering," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 5, pp. 1971–1983, 2022.
- [19] S. Tuslaela, R. Rusdiansyah, H. Suspendar, and N. Suharyanti, "Implementation of K-Means Clustering in Food Security by Regency in East Java Province in 2022," *Sinkron: Jurnal Penelitian dan Pengabdian kepada Masyarakat*, vol. 9, no. 1, pp. 54–60, 2024.
- [20] W. Kusuma, A. Setiawan, and A. Prasetyo, "Comparison of k-means clustering with hierarchical agglomerative clustering for the analysis of food security of rice sector in Indonesia," *Journal of Intelligent Decision Support System*, vol. 8, no. 1, pp. 22–33, 2025.
- [21] Y. A. Ishak, D. Nugraha, and R. Astuti, "Advanced Clustering Approach for Mapping Regions of Paddy Productivity in Indonesia Using Intelligent K-Means," in *2025 4th International Conference on Data Science and Information Technology (DSIT)*, 2025.
- [22] G. Srinivasan, A. Kannan, and A. Rahman, "Energy efficient hierarchical clustering based dynamic data fusion algorithm for wireless sensor networks in smart agriculture," *Sci Rep*, vol. 15, no. 1, p. 7207, 2025.
- [23] T. Ahmed, Y. Zhang, and V. Kumar, "Fuzzy Methods in Smart Farming: A Systematic Review," *Comput Electron Agric*, vol. 211, p. 108041, 2024.
- [24] M. A. Rahmadani and others, "Advanced Clustering Approach for Mapping Regions of Paddy Productivity in Indonesia Using Intelligent K-Means," in *2024 8th International Conference on Informatics and Computational Sciences (ICICoS)*, 2024, pp. 64–69. doi: 10.1109/ICICoS60247.2024.10932942.
- [25] G. Rachmawati and others, "K-Means Using Dynamic Time Warping for Clustering Cities in Java Island According to Meteorological Conditions," in *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2023, pp. 96–101. doi: 10.1109/ISITIA58215.2023.10381899.
- [26] I. M. Kartini, Y. Surya, and B. Triyono, "Clustering Data Meteorologi Wilayah Indonesia Timur Dengan Metode K-Means dan Fuzzy C-Means," *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 17, no. 2, pp. 123–130, 2023, [Online]. Available: <https://ejournal.nusamandiri.ac.id/index.php/inti/article/view/5039>
- [27] S. Fadhilah, B. Widodo, and P. Lestari, "An Analysis of Meteorological Data in Sumatra and Nearby Using Agglomerative Clustering," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 3, pp. 412–419, 2023, [Online]. Available: <https://jurnal.iaii.or.id/index.php/RESTI/article/view/5663>
- [28] D. N. Iliescu, "Fully Unsupervised Clustering in Nonlinearly Separable Data Using Intelligent Kernel KMeans," in *2014 International Joint Conference on Neural Networks (IJCNN)*, 2014, pp. 2474–2481. doi: 10.1109/IJCNN.2014.6889674.
- [29] M. D. Pratama and others, "An Intelligent Clustering Approach For Analyzing A Multivariate Time Series Dataset, Case Study COVID-19 Outbreak in Indonesia," in *2023 7th International Conference on Informatics and*

Computational Sciences (ICICoS), 2023, pp. 77–82. doi: 10.1109/ICICoS60247.2023.10367007.

- [30] R. K. U. Haq and others, “Student Graduation Time Prediction Using Intelligent K-Medoids Algorithm,” in *2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2017, pp. 113–118. doi: 10.1109/ICACSIS.2017.8257122.

Fawaz, saat ini sebagai mahasiswa program studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara.

Teny Handhayani, memperoleh gelar Sarjana Komputer (S.Kom.) dari Institut Pertanian Bogor (IPB) dan Magister Ilmu Komputer (M.Kom.) dari Universitas Indonesia (UI). Kemudian beliau memperoleh gelar Ph.D. dalam bidang Ilmu Komputer dari University of York, Inggris. Saat ini beliau aktif sebagai Staf Pengajar pada program studi Teknik Informatika, Universitas Tarumanagara.