

IMPLEMENTASI METODE FUZZY C-MEANS DALAM PENGELOMPOKAN HASIL PANEN PADI DI PROVINSI BALI

Implementation of the Fuzzy C-Means Method in Grouping Rice Yields in the Province of Bali

Indah Manfaati Nur¹, Afia Nur Lailatus Syifa², Miftakhul Kharis³, Shella Heidy Permatasari^{4*}

^{1,2,3,4}Statistika, Universitas Muhammadiyah Semarang

Jl. Kedungmundu No.18, Kedungmundu, Kota Semarang, 50273, Provinsi Jawa Tengah, Indonesia

E-mail Corespponding Author: heidypshella@gmail.com⁴

Abstrak: Peranan padi di Indonesia sebagai salah satu komoditas pertanian tidak dapat diabaikan dan menjadi sangat penting. Salah satu cara yang efektif untuk mencapai peningkatan hasil pertanian yang optimal adalah menggunakan metode pengelompokan. Melalui pengklasteran (*clustering*), pemerintah dan petani dapat mengidentifikasi tanaman yang memerlukan perhatian ekstra, sehingga memungkinkan mereka untuk mengelola pertanian dengan lebih efisien, terutama pada kluster-kelompok yang memiliki potensi tinggi. Terdapat beberapa metode pengelompokan dalam *cluster* yang dapat digunakan, seperti K-Means, Hirarki, Fuzzy C-Means, DBSCAN, dan metode lainnya. Metode Fuzzy C-Means (FCM) menjadi solusi yang menarik untuk mengelompokkan hasil panen padi dengan lebih efektif. Karena FCM tidak memiliki asumsi tertentu tentang bentuk *cluster* sehingga dapat menangani data dengan berbagai bentuk *cluster* yang beragam, dan juga lebih baik dalam mengatasi data *noise* atau *outlier* dibandingkan beberapa metode pengelompokan lainnya. Penelitian ini bertujuan untuk mengimplementasikan metode FCM dalam pengelompokan hasil panen padi di Provinsi Bali. Data yang digunakan adalah data sekunder yang diperoleh dari kabupaten/kota di Bali, termasuk variabel yang relevan seperti produktivitas, luas lahan, produksi padi, dan tinggi dataran. Hasilnya menunjukkan tiga *cluster* utama: kluster pertama terdiri dari Kabupaten Badung, Gianyar, dan Buleleng; *cluster* kedua terdiri dari Kabupaten Tabanan; dan *cluster* ketiga terdiri dari Kabupaten Jembrana, Klungkung, Bangli, Karangasem, dan Kota Denpasar. Untuk pengukuran kebaikan *cluster* digunakan nilai validasi. Hasil validasi untuk nilai FSI sebesar 0,758, nilai PEI menghasilkan nilai 0,221, PCI dan MPCI, diperoleh nilai 0,888 dan 0,832.

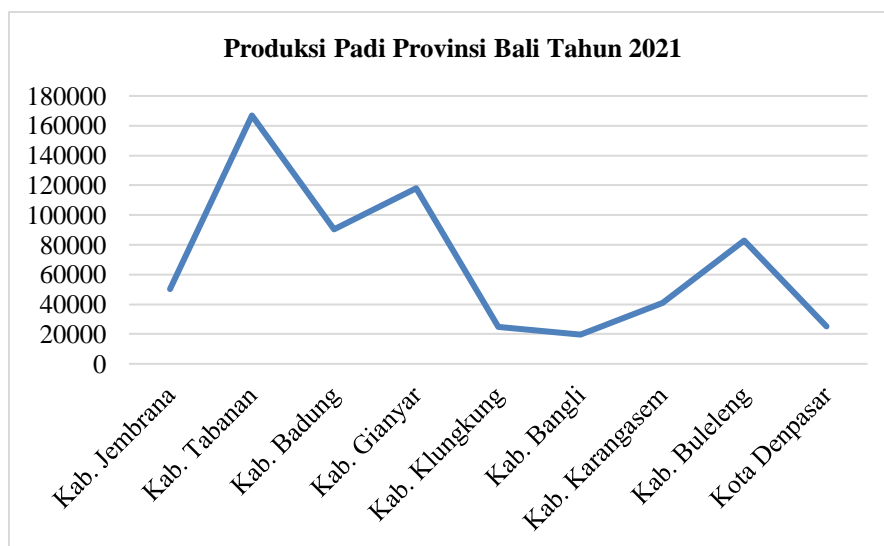
Kata Kunci: *Cluster, Fuzzy C-Means, Panen Padi.*

Abstract: The role of rice in Indonesia as an agricultural commodity cannot be ignored and is very important. One effective way to achieve an optimal increase in agricultural yields is to use the clustering method. Through clustering, the government and farmers can identify which crops require extra attention, enabling them to manage agriculture more efficiently, especially within clusters that have high potential value. There are several clustering methods that can be used, such as K-Means, Hierarchy, Fuzzy C-Means, DBSCAN, and other methods. The Fuzzy C-Means (FCM) method is an attractive solution for classifying rice yields more effectively. This study aims to implement the FCM method in classifying rice yields in the Province of Bali. Because FCM does not make specific assumptions about cluster shapes, it can handle data with various diverse cluster shapes, and it is also better at handling data noise or outliers compared to some other clustering methods. The data used is secondary data obtained from districts/cities in Bali, including relevant variables such as productivity, land area, rice production, and altitude. The results show three main clusters: the first cluster consists of Badung, Gianyar and Buleleng Regencies; the second cluster consists of Tabanan Regency; and the third cluster consists of Jembrana, Klungkung, Bangli, Karangasem, and Denpasar City Regencies. To measure the goodness of clusters, the validation value was used, the validation results for the FSI value were 0.758, the PEI value was 0.221, PCI and MPCI were obtained values of 0.888 and 0.832.

Keywords: *Cluster, Fuzzy C-Means, Rice Yield*

1. PENDAHULUAN

Pertumbuhan ekonomi di Indonesia pada Tahun 2021 secara keseluruhan mengalami peningkatan positif sebesar 3,69%, hal ini menunjukkan perbaikan dibandingkan dengan Tahun 2020 yang terjadi kontraksi ekonomi sebesar 2,07% sebagai dampak pandemi covid-19 [1]. Dampaknya, pada Tahun 2020 ekonomi Indonesia mengalami deflasi atau penurunan yang signifikan karena kondisi perekonomian di Indonesia memiliki fluktuasi yang tidak stabil [2]. Di era globalisasi saat ini, sektor pertanian memiliki peranan penting dalam mendorong pertumbuhan ekonomi Indonesia. Sektor pertanian menjadi pilar utama dalam kegiatan ekonomi di Indonesia, karena tidak hanya berfungsi sebagai penyedia pangan bagi masyarakat sehari-hari, tetapi juga sebagai sumber pendapatan devisa negara [3]. Peran strategis dan penting sektor pertanian dalam ekonomi nasional tidak dapat diragukan. Terutama, sektor pertanian masih menjadi sumber lapangan pekerjaan bagi sebagian besar penduduk Indonesia yang tinggal di daerah pedesaan dan berperan dalam penyediaan bahan pangan bagi masyarakat. Hampir seluruh penduduk Indonesia mengandalkan tanaman padi sebagai sumber utama bahan pangan mereka. Oleh karena itu, tanaman padi memiliki nilai yang signifikan dalam aspek spiritual, budaya, ekonomi, dan politik bagi bangsa Indonesia karena mempengaruhi kehidupan banyak orang.



Gambar 1. Grafik (Ton)

Setiap daerah di Indonesia memiliki potensi dan sumber daya yang berbeda-beda. Salah satunya adalah Provinsi Bali, yang memiliki kekayaan sumber daya yang penting dan signifikan terutama sumberdaya pariwisata dan pertaniannya. Pulau Bali memiliki daya tarik unik dengan keindahan alamnya serta berbagai atraksi wisata budaya yang menjadi ciri khasnya. Selain dikenal sebagai tujuan wisata, Pulau Bali juga dikenal baik dalam sektor pertaniannya, seperti sistem subak yang diakui sebagai warisan budaya dunia [4]. Pada saat ini, tidak terdapat peningkatan yang signifikan pada pertumbuhan produksi padi di Provinsi Bali. Kebanyakan petani di wilayah tersebut menghadapi keterbatasan lahan untuk bercocok tanam, sehingga laba yang mereka peroleh dari usaha pertanian padi cenderung minim [5]. Meskipun luas area usaha tani padi sawah di Provinsi Bali relatif kecil, hanya mencakup 14,40% dari total luas wilayahnya, namun tingkat produktivitasnya terlihat lebih tinggi apabila dibandingkan dengan rata-rata nasional [6]. Pada rentang Januari hingga Desember 2021, produksi padi di Bali mencapai 618.910,81 ton Gabah Kering Giling (GKG). Bulan April menjadi bulan dengan produksi tertinggi, mencapai 78.874,36 ton. Dua kabupaten dengan produksi padi (GKG) tertinggi adalah Tabanan dengan produksi sebesar 166.880,26 ton dan Gianyar dengan produksi sebesar 117.895,26 ton. Melihat banyaknya data yang dihasilkan dari produksi tersebut, dibutuhkan metode seperti *data mining* untuk mencari informasi tertutup dari data tersebut [7].

Data mining dikenal sebagai teknik yang digunakan untuk menggali pola-pola tersembunyi pada sekumpulan data dengan maksud untuk menghasilkan informasi baru. Metode yang digunakan dalam *data mining* bervariasi tergantung pada tujuan penggunaan dataset. Ini termasuk estimasi, prediksi, klasifikasi, *clustering*, dan asosiasi. Dalam pemetaan data, salah satu teknik yang digunakan adalah *clustering* atau *clustering*. *Clustering* adalah sebuah teknik dalam *data mining* yang digunakan untuk mengelompokkan data berdasarkan kesamaannya ke dalam kelompok yang disebut *cluster*. [8] Terdapat beberapa algoritma yang digunakan dalam *clustering*, salah satunya adalah algoritma *Fuzzy C-Means* (FCM) *Clustering* [9]. Proses FCM *clustering* melibatkan pembobotan tingkat keanggotaan dalam himpunan fuzzy. Metode ini memungkinkan suatu bagian data memiliki dua kelompok atau lebih. Dalam FCM, objek cenderung ditempatkan dalam kelompok berdasarkan tingkat keanggotaannya [10]. Dengan menerapkan metode FCM, kelompok tanaman padi yang berbeda dengan karakteristik produktivitas yang serupa dapat diidentifikasi. Pengelompokan ini dapat dianalisis lebih lanjut untuk mendapatkan wawasan tentang faktor-faktor yang berkontribusi terhadap produktivitas tinggi atau rendah, seperti kondisi tanah yang optimal, praktik irigasi yang tepat, pengelolaan hara yang sesuai, atau strategi pengendalian hama yang efektif [11].

Metode FCM menggabungkan konsep logika *fuzzy* yang memungkinkan keanggotaan parsial titik data dalam kelompok yang berbeda. Hal ini sangat berguna ketika berhadapan dengan data yang kompleks dan ambigu karena memberikan pendekatan yang lebih fleksibel untuk pengelompokan dibandingkan dengan algoritma pengelompokan tradisional. Metode FCM menetapkan derajat keanggotaan ke setiap titik data yang menunjukkan sejauh mana observasi dimiliki oleh setiap *cluster* [12]. Hal ini memungkinkan penilaian produktivitas tanaman padi yang lebih bernuansa, dengan mempertimbangkan berbagai faktor yang mempengaruhi kinerja tanaman. Dengan menggunakan metode FCM, dapat diperoleh pemahaman yang komprehensif tentang variasi produktivitas tanaman padi dalam suatu area atau wilayah tertentu. Informasi ini dapat digunakan untuk menyesuaikan praktik pertanian dan strategi manajemen untuk kelompok tertentu, yang pada akhirnya mengarah pada peningkatan produktivitas dan alokasi sumber daya. Selain itu, metode FCM dapat membantu dalam mengidentifikasi pola dan tren produktivitas tanaman padi, memfasilitasi intervensi yang ditargetkan, dan pengambilan keputusan yang terinformasi dalam budidaya padi.

Salah satu penelitian yang dilakukan oleh S. Mulyani, B. N. Sari, dan A. A. Ridha (2020) [15] mengenai *clustering* menggunakan metode FCM, menunjukkan bahwa hasil pengelompokan dengan menggunakan metode FCM memiliki struktur yang baik yaitu diatas 0,5 (mendekati 1) dengan teknik evaluasi *Silhouette Coefficient*. Adapun juga penelitian yang dilakukan oleh M. Orisa and A. Faisol [16], yang mengelompokkan data menerapkan metode *Fuzzy C-Means* (FCM) dengan meminimalkan fungsi keanggotaan dari kumpulan data, sehingga setiap potongan data memiliki kemungkinan menjadi anggota dari lebih dari satu kelompok. Dan juga penelitian yang dijalankan oleh N. Uinnuha (2020) [11], dihasilkan 3 *cluster*, dengan hasil dari tiga *cluster* diperoleh 1 provinsi pada *cluster* pertama dengan hasil perkebunan tinggi, 27 provinsi pada *cluster* kedua dengan hasil perkebunan rendah dan 6 provinsi pada *cluster* ketiga dengan hasil perkebunan sedang. Pada pengujian indeks siluet, nilai indeks siluet seluruh *cluster* adalah 0,84321191 yang berarti bahwa hasil pengelompokan sudah sesuai. Maka dari itu, kami memilih metode FCM guna dilakukannya *clustering* pada data produktivitas panen padi di Provinsi Bali. Tujuan penelitian ini adalah mengelompokkan dan mengidentifikasi *cluster* tanaman padi yang memiliki karakteristik produktivitas serupa dengan menggunakan metode *Fuzzy C-Means* (FCM) sehingga dapat memahami faktor-faktor yang berkontribusi terhadap kinerja tanaman padi.

2. METODOLOGI

2.1. Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data yang telah ada sebelumnya atau dikenal sebagai data sekunder, dengan mengumpulkan data hasil panen padi dari berbagai wilayah di Provinsi Bali, termasuk variabel-variabel yang relevan seperti produktivitas panen, luas lahan, dan faktor-faktor lain yang berpotensi mempengaruhi hasil panen padi. Data penelitian disajikan pada Tabel 1.

Tabel 1. Data Penelitian yang Digunakan

Kota/Kabupaten	Produktivitas Padi	Luas Panen Padi (Ha ²)	Produksi Padi (Ton)	Tinggi Dataran (Mdpl)
Kab. Jembrana	5,313755198	9.458,17	50.258,4	12
Kab. Tabanan	5,712780546	29.211,74	166.880,26	124
Kab. Badung	6,086198093	14.825,85	90.233,06	25
Kab. Gianyar	6,028018374	19.557,88	117.895,26	126
Kab. Klungkung	6,248217893	3.970,02	24.805,55	93
Kab. Bangli	5,060471152	3.883,67	19.653,2	425
Kab. Karangasem	5,986095426	6.827,25	40.868,57	102
Kab. Buleleng	5,879700005	14.112,22	82.975,62	60
Kota Denpasar	7,554274693	3.354,51	25.340,89	25

Sumber: BPS Provinsi Bali [13]

2.2. Tahapan Analisa Data

Pada tahap ini, data akan dianalisis menggunakan KDD (*Knowledge Discovery in Databases*). Keseluruhan proses KDD dimulai dengan tiga tahap yang melibatkan pengolahan data sebelum menghasilkan informasi [7], yaitu:

a) Data Mining

Data mining adalah konsep yang merujuk pada penemuan pengetahuan di dalam basis data. Hal ini melibatkan penerapan statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengenali informasi berharga serta pengetahuan yang terkait dari beragam basis data yang besar [15]. Dalam penelitian ini, digunakan proses *data mining* yang melibatkan penerapan teknik *clustering* menggunakan algoritma FCM dengan tujuan untuk mengidentifikasi dan mendapatkan karakteristik dari data yang terkumpul.

b) Fuzzy C-Means (FCM)

Penelitian ini menggunakan metode FCM, yang merupakan sebuah algoritma yang tergolong dalam kategori pembelajaran tanpa pengawasan. Algoritma ini digunakan untuk memproses himpunan data yang mengandung *outlier*, di mana terdapat objek-objek yang berjarak jauh dari *cluster* yang sudah ada sehingga tidak dapat dengan jelas diklasifikasikan ke dalam *cluster* tertentu [16]. *Clustering fuzzy* adalah sebuah metode pengelompokan yang menggunakan fungsi keanggotaan *fuzzy* untuk menentukan kelompok berdasarkan jarak. Teknik ini sebenarnya merupakan hasil dari pengembangan metode *partitional* dengan melakukan penambahan elemen *fuzzy*, yang memungkinkan pengelompokan data bahkan ketika kelompok-kelompok data yang memiliki distribusi ambigu atau tidak terlihat dengan jelas [14].

FCM merupakan sebuah metode pengelompokan data dimana setiap titik data dalam suatu *cluster* didefinisikan oleh derajat keanggotaannya. Berikut adalah deskripsi algoritma FCM menurut Kusumadewi dan Purnomo [7]:

1. Tentukan data yang akan berada di *cluster* X, berupa matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut masing – masing data). X_{ij} = data sampel ke – i ($i = 1, 2, \dots, n$), j – atribut ($j = 1, 2, \dots, m$).

dengan batasan:

- a. Jumlah *cluster* (c)
- b. Pangkat (w)
- c. Maksimum iterasi (Maxit)
- d. Error terkecil yang diharapkan (ξ)
- e. Fungsi obyektif awal ($P_0 = 0$)
- f. Iterasi awal ($t = 1$)

2. Membangkitkan bilangan random secara acak untuk $\mu_{ik}, i = 1, 2, \dots, n; k = 1, 2, \dots, c$; sebagai elemen-elemen matrik partisi awal U, setiap baris memiliki jumlah nilai elemen kolom yang sama yaitu 1 (satu).

$$\sum_{i=1}^c \mu_{ci} = 1 \quad (1)$$

3. Menghitung pusat *cluster* ke-k: V_{kj} dengan $k = 1, 2, \dots, c$; dan $j = 1, 2, \dots, m$.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w \times X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (2)$$

4. Menghitung fungsi obyektif pada iterasi ke- t, P_t :

Fungsi tujuan digunakan sebagai syarat perulangan untuk mendapatkan pusat *cluster* yang tepat. Untuk mendapatkan kecenderungan data masuk ke *cluster* mana pada langkah terakhir. Untuk iterasi awal, nilai $t = 1$.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (3)$$

5. Menghitung perubahan matrik partisi:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{w-1}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{w-1}} \quad (4)$$

dengan $i = 1, 2, \dots, n$; dan $k = 1, 2, \dots, c$

6. Memeriksa kondisi berhenti:

- a. Jika $(|P_t - P_{t-1}| < \xi)$ atau $(t > MakIter)$ maka berhenti;
- b. Jika tidak, atau $t = t + 1$, ulangi langkah ke- d (menghitung V_{kj}).

c) *Postprocessing*

Postprocessing bertujuan untuk memberikan bantuan kepada pengguna dalam memperoleh pemahaman yang lebih baik terhadap informasi yang dihasilkan. Karena interpretasi langsung terhadap pusat *cluster* masih sulit, diperlukan suatu metode bernama *Simple Additive Weighting* (SAW) yang dapat melakukan perbandingan terhadap nilai pusat *cluster* terakhir tersebut. Metode SAW juga dikenal dengan sebutan metode penjumlahan terbobot. Konsep dasar metode SAW adalah untuk menghitung penjumlahan terbobot dari rating kinerja pada setiap alternatif untuk semua atribut. Berikut adalah langkah-langkah dalam metode SAW [7]:

1. Membuat matriks keputusan Z berukuran $m \times n$, dimana m = alternatif yang akan dipilih dan n = kriteria.
2. Memberikan nilai x setiap alternatif (i) pada setiap kriteria (j) yang sudah ditentukan, dimana, $i = 1, 2, \dots, m$ dan $j = 1, 2, \dots, n$ pada matriks keputusan Z,

$$\begin{bmatrix} x_{11} & x_{12} & x_{1j} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ x_{i1} & x_{i2} & x_{ij} \end{bmatrix} \quad (5)$$

3. Memberikan nilai bobot preferensi (W) oleh pengambil keputusan untuk masing-masing kriteria yang sudah ditentukan.

$$W = W_1 W_2 W_3 \dots W_j \quad (6)$$

4. Melakukan normalisasi matriks keputusan Z dengan cara menghitung nilai rating kinerja ternormalisasi (r_{ij}) dari alternatif A_i pada atribut C_j ;

$$r_{ij} = \begin{cases} \frac{x_{ij}}{MAX_i(x_{ij})} , & \text{jika } j \text{ adalah atribut keuntungan} \\ \frac{MIN_i(x_{ij})}{x_{ij}} , & \text{jika } j \text{ adalah atribut biaya} \end{cases} \quad (7)$$

dengan ketentuan:

- a. Atribut keuntungan dapat diidentifikasi ketika atribut tersebut memberikan manfaat yang signifikan bagi pengambil keputusan. Di sisi lain, atribut biaya adalah atribut yang berkontribusi pada pengeluaran atau biaya yang meningkat seiring dengan nilai yang lebih besar bagi pengambil keputusan.
 - b. Jika atribut tersebut merupakan atribut keuntungan, maka nilai (x_{ij}) dari setiap kolom atribut akan dibagi dengan nilai maksimum ($MAX x_{ij}$) dari setiap kolom. Namun, jika atribut tersebut adalah atribut biaya, maka nilai minimum ($MIN x_{ij}$) dari setiap kolom atribut akan dibagi dengan nilai (x_{ij}) dari setiap kolom.
5. Hasil dari nilai rating kinerja ternormalisasi (r_{ij}) membentuk matriks ternormalisasi (N);

$$N = \begin{bmatrix} r_{11} & r_{12} & r_{1j} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ r_{i1} & r_{i2} & r_{ij} \end{bmatrix} \quad (8)$$

6. Proses perankingan dilakukan dengan mengalikan matriks yang telah dinormalisasi (N) dengan nilai bobot preferensi (W).
7. Untuk menentukan nilai preferensi dari setiap alternatif (V_i), dilakukan dengan menjumlahkan hasil perkalian antara matriks yang telah dinormalisasi (N) dengan nilai bobot preferensi (W).

$$V_i = \sum_{j=1}^n w_j r_{ij} \quad (9)$$

Apabila nilai V_i lebih tinggi, hal tersebut menunjukkan bahwa alternatif A_i adalah alternatif terbaik.

d) Evaluasi hasil pengelompokan

Setelah menerapkan FCM pada data, akan diamati beberapa hasil seperti kelompok dan jumlah iterasi maksimum yang tercapai. *Silhouette Coefficient* adalah sebuah metode yang digunakan untuk mengevaluasi kualitas dan kekuatan dari *cluster*. Metode ini menggabungkan dua pendekatan, yaitu metode kohesi yang mengukur kedekatan antar objek dalam suatu *cluster*, dan metode separasi yang mengukur sejauh mana *cluster* tersebut terpisah dari *cluster* lainnya. Rentang nilai Indeks Silhouette adalah dari -1 hingga 1, dimana nilai yang lebih tinggi menandakan kualitas *cluster* yang lebih baik. Nilai positif menunjukkan bahwa objek tercluster dengan baik, sedangkan nilai negatif menunjukkan bahwa objek mungkin lebih cocok dalam kluster lain [17][18].

$$S_i = \left(\frac{b_i - a_i}{\max(a_i, b_i)} \right) \quad (10)$$

dimana:

S_i : *Silhouette coefficient*

a_i : rata-rata jarak dari objek $ke - i$ dengan seluruh objek di *cluster*

b_i : rata-rata jarak dari objek ke- i dengan cluster apa pun yang tidak mengandung objek

Terdapat beberapa langkah untuk menunjukkan *cluster* yang telah dilakukan. Tahapan pertama melibatkan penghitungan beberapa indeks validasi. Untuk indeks validasi yang digunakan, yaitu *Fuzzy Silhouette Index* (FSI), *Partition Entropy Index* (PEI), *Partition Coefficient Index* (PCI), dan *Modified Partition Coefficient Index* (MPCI) [19][20]. Kriteria indeks ini untuk *cluster* disajikan dalam Tabel 2.

Tabel 2. Nilai Kriteria Subjektif Pengukuran Cluster

Nilai	Interpretasi
0,71 - 1,00	Struktur <i>cluster</i> sangat baik
0,51 - 0,70	Struktur <i>cluster</i> baik
0,26 - 0,50	Struktur <i>cluster</i> lemah
<0,25	Struktur <i>cluster</i> buruk

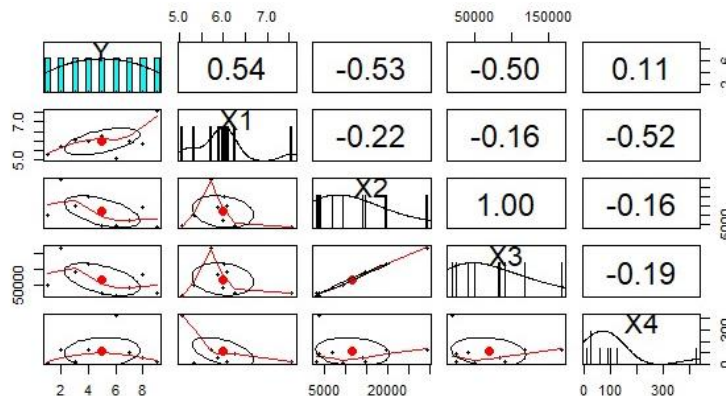
Nilai kriteria subjektif pengukuran *cluster* digunakan untuk mengukur nilai validasi dari FSI, PCI, dan MPCI. Dapat dilihat untuk nilai validasi pada FSI, PCI, dan MPCI yang sangat baik adalah yang mendekati nilai 1. Sedangkan untuk nilai validasi PEI, nilai kriteria pengukuran yang memiliki struktur *cluster* sangat baik adalah yang mendekati nilai 0 [21].

3. HASIL DAN PEMBAHASAN

Analisis proses mining dengan data produktivitas panen, luas lahan, dan produksi padi di Provinsi Bali Tahun 2021, terdapat 36 data yang telah diolah untuk menghitung jarak ke setiap titik pusat *cluster*. Kemudian dilakukan *clustering* untuk setiap kabupaten/kota dan juga validasi untuk metode *Fuzzy C-Means* FCM. Penghitungan yang dilakukan menggunakan *syntax* Rstudio.

3.1. Korelasi Antar Atribut

Korelasi antar atribut mengacu pada hubungan statistik antara dua atau lebih variabel atau atribut dalam suatu dataset. Korelasi ini mengukur sejauh mana perubahan dalam hubungan antara perubahan dalam satu variabel dengan perubahan dalam variabel lainnya. Korelasi dapat membantu kita memahami pola dan hubungan antara atribut dalam dataset. Korelasi tersebut disajikan dalam Gambar 2 berikut ini.



Gambar 2. Korelasi Antar Atribut

Gambar 2 menunjukkan hasil dari korelasi antar atribut disertai dengan plot dan juga grafik. Apabila hasil dari korelasi mendekati 1, artinya korelasi antar variabelnya semakin kuat. Kemudian, untuk hasil negatif menunjukkan bahwa korelasi berbanding terbalik yang artinya apabila salah satu nilainya bertambah, maka korelasinya akan menurun. Sedangkan untuk hasil positif menunjukkan bahwa korelasi berbanding searah, apabila salah satu nilainya bertambah maka korelasinya akan bertambah.

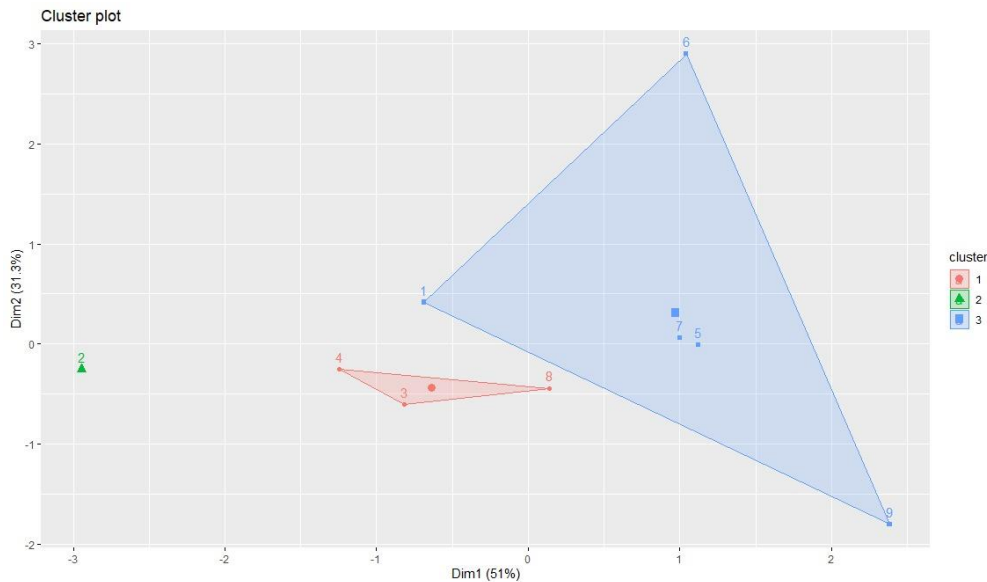
3.2. Melakukan *Clustering Fuzzy C-Means (FCM)*

Sebelum melakukan *clustering*, kita harus menentukan *cluster* atau jumlah pengelompokannya terlebih dahulu. Dalam penelitian ini, digunakan *cluster* sebanyak 3 ($k=3$). Berikut adalah Tabel 3 yang merupakan hasil dari matriks derajat keanggotaan yang menunjukkan keanggotaan dari masing-masing *cluster*.

Tabel 3. Matriks Derajat Keanggotaan

Kota/Kabupaten	Membership Degrees Matrix		
	Cluster 1	Cluster 2	Cluster 3
Kab. Jembrana	0,168206276	0,023418567	0,808375157
Kab. Tabanan	0,000743111	0,999037926	0,000218963
Kab. Badung	0,995814342	0,001611032	0,002574626
Kab. Gianyar	0,738799665	0,201673223	0,059527112
Kab. Klungkung	0,008047711	0,001913622	0,990038667
Kab. Bangli	0,022621599	0,005778047	0,97160035
Kab. Karangasem	0,034522790	0,006126483	0,95935073
Kab. Buleleng	0,949080492	0,014615719	0,03630379
Kota Denpasar	0,007248110	0,001712648	0,99103924

Hasil *output* Tabel 3 menunjukkan derajat keanggotaan dari setiap variabel. Derajat keanggotaannya paling tinggi atau mendekati nilai satu, masuk ke dalam *cluster* tersebut. Misal saja pada data pertama, nilai membership degrees terbesarnya adalah 0,808 terletak pada *cluster* 3, artinya adalah data pertama termasuk ke dalam *cluster* ketiga dan seterusnya. Hasil pengelompokan dengan FCM dapat dilihat pada Gambar 3.



Gambar 3. Hasil Plot Clustering

Berdasarkan output hasil pengelompokkan dengan FCM dalam bentuk plot *cluster* tersebut menunjukkan terdapat 3 *cluster* dalam penelitian ini. *Cluster* pertama berwarna merah dengan keanggotaan *cluster*nya yaitu Kabupaten Badung, Kabupaten Gianyar, dan Kabupaten Buleleng. *Cluster* kedua berwarna hijau dengan keanggotaan *cluster*nya yaitu Kabupaten Tabanan. *Cluster* ketiga berwarna biru dengan keanggotaan *cluster*nya yaitu Kabupaten Jembrana, Kabupaten Klungkung, Kabupaten Bangli, Kabupaten Karangasem, dan Kota Denpasar.

Selanjutnya akan dilakukan validasi dengan menghitung *Fuzzy Silhouette Index* (FSI), kemudian *Partition Entropy Index* (PEI), *Partition Coefficient Index* (PCI), dan *Modified Partition Coefficient Index* (MPCI). Nilai validasi *cluster* disajikan pada Tabel 4.

Tabel 4. Nilai Validasi Cluster

Nilai Validasi Cluster	
<i>Fuzzy Silhouette Index (FSI)</i>	0,758
<i>Partition Entropy Index (PEI)</i>	0,221
<i>Partition Coefficient Index (PCI)</i>	0,888
<i>Modified Partition Coefficient (MPCI)</i>	0,832

Berdasarkan data yang terdapat pada Tabel 4 bahwa hasil validasi *clustering* data menggunakan metode FCM memiliki kualitas yang sangat baik dengan nilai FSI = 0,758. Namun, hasil validasi menggunakan metode PEI menunjukkan nilai PEI = 0,221, yang menandakan bahwa kualitas *clustering* data sangat baik jika mendekati 0. Sementara itu, untuk PCI dan MPCI, nilai yang diperoleh adalah 0,888 dan 0,832. Nilai validasi yang mendekati 1 menunjukkan bahwa *clustering* data telah dilakukan dengan sangat baik untuk PCI dan MPCI.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan beberapa hal sebagai berikut.

1. *Fuzzy C-Means* (FCM) *clustering algorithm* dapat digunakan untuk mengelompokkan hasil panen padi di Bali berdasarkan karakteristik produktivitas yang serupa. Metode ini dapat membantu petani dan pengelola pertanian untuk menyesuaikan praktik pertanian dan strategi manajemen untuk kelompok tertentu, sehingga dapat meningkatkan produktivitas dan alokasi sumber daya. Selain itu, dibahas pula metode *post-processing Simple Additive Weighting* (SAW) untuk menilai kualitas dan kekuatan *cluster*, serta *Silhouette Coefficient Method* untuk mengevaluasi kualitas *cluster*.
2. Hasil pengelompokkan terdapat tiga *cluster* untuk mengelompokkan kabupaten/kota di Provinsi Bali. *Cluster* pertama dengan keanggotaan *cluster*-nya yaitu Kabupaten Badung, Kabupaten Gianyar, dan Kabupaten Buleleng. *Cluster* kedua dengan keanggotaan *cluster*-nya yaitu Kabupaten Tabanan. *Cluster* ketiga dengan keanggotaan *cluster*-nya yaitu Kabupaten Jembrana, Kabupaten Klungkung, Kabupaten Bangli, Kabupaten Karangasem, dan Kota Denpasar.
3. Hasil validasi *clustering* data menggunakan metode FCM menunjukkan kualitas yang sangat baik dengan nilai FSI = 0,758. Sedangkan untuk metode validasi menggunakan PEI menghasilkan nilai PEI = 0,221 yang menandakan kualitas *clustering* data yang sangat baik ketika mendekati 0. Selain itu, untuk PCI dan MPCI, diperoleh nilai 0,888 dan 0,832. Nilai validasi mendekati 1 menunjukkan bahwa *clustering* data telah dilakukan dengan sangat baik untuk PCI dan MPCI.

DAFTAR PUSTAKA

- [1] B. P. Statistik, "BPS Catat Ekonomi Indonesia Tumbuh 3,69 Persen Pada 2021," *Kominfo.go.id*,

2021. <https://www.kominfo.go.id/content/detail/39835/bps-catat-ekonomi-indonesia-tumbuh-369-persen-pada-2021/0/berita> (accessed May 13, 2023).
- [2] R. N. Hayati, “Pemulihan Perekonomian Indonesia Setelah Kontraksi Akibat Pandemi Covid-19,” *Kementerian Keuangan Republik Indonesia*, 2022. <https://www.djkn.kemenkeu.go.id/kpknl-banjarmasin/baca-artikel/14769/Pemulihan-Perekonomian-Indonesia-Setelah-Kontraksi-Akibat-Pandemi-Covid-19.html> (accessed May 13, 2023).
- [3] S. I. Kusumaningrum, “Pemanfaatan Sektor Pertanian Sebagai Penunjang Pertumbuhan Perekonomian Indonesia,” *J. Transaksi*, vol. 11, no. 1, pp. 80–89, 2019, [Online]. Available: <http://ejournal.atmajaya.ac.id/index.php/transaksi/article/view/477>
- [4] E. P. Yuendini, I. N. Rachmi, N. Nurul, A. Puspitasari, and R. Harini, “Analisis Potensi Ekonomi Sektor Pertanian dan Sektor Pariwisata di Provinsi Bali Menggunakan Teknik Analisis Regional,” *J. Geogr.*, vol. 16, no. 2, pp. 128–136, 2019, doi: 10.15294/jg.v16i2.20831.
- [5] S. Suharyanto, J. H. Mulyo, D. H. Darwanto, and S. Widodo, “Analisis Produksi dan Efisiensi Pengelolaan Tanaman Terpadu Padi Sawah di Provinsi Bali,” *J. Penelit. Pertan. Tanam. Pangan*, vol. 34, no. 2, p. 131, 2015, doi: 10.21082/jpptp.v34n2.2015.p131-143.
- [6] S. Suharyanto, J. Rinaldy, and N. Ngurah Arya, “Analisis Risiko Produksi Usahatani Padi Sawah,” *Agrar. J. Agribus. Rural Dev. Res.*, vol. 1, no. 2, pp. 70–77, 2015, doi: 10.18196/agr.1210.
- [7] Nurjanah, A. Farmadi, and F. Indriani, “Implementasi Metode Fuzzy C-Means Pada Sistem Clustering Data Varietas Padi,” *Kumpul. J. Ilmu Komput.*, vol. 01, no. 01, pp. 23–32, 2014.
- [8] P. R. Arum and I. M. Nur, “Clustering The Level Of Education Quality In The Central Java Province Using C-Means Method,” pp. 508–513, 2020, doi: 10.4108/eai.28-9-2019.2291056.
- [9] 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 J.* ..., vol. III, pp. 1–8, 2022, [Online]. Available: <https://journal.ubpkarawang.ac.id/mahasiswa/index.php/ssj/article/view/411>
- [10] J. Tamaela, E. Sedyono, and A. Setiawan, “Cluster Analysis Menggunakan Algoritma Fuzzy C-means dan K-means Untuk *Clustering* dan Pemetaan Lahan Pertanian di Minahasa Tenggara,” *J. Buana Inform.*, vol. 8, no. 3, pp. 151–160, 2017, doi: 10.24002/jbi.v8i3.1317.
- [11] N. Ulinnuha, “Provincial Clustering in Indonesia Based on Plantation Production Using Fuzzy C-Means,” *J. Ilm. Teknol. dan Inf.*, vol. 9, no. 1, pp. 8–12, 2020.
- [12] Z. Zhou, Y. Zang, Y. Li, Y. Zhang, P. Wang, and X. Luo, “Rice plant-hopper infestation detection and classification algorithms based on fractal dimension values and fuzzy C-means,” *Math. Comput. Model.*, vol. 58, no. 3–4, pp. 701–709, 2013, doi: 10.1016/j.mcm.2011.10.028.
- [13] B. P. B. Statistik, “Data BPS.”
- [14] S. Agung Priambodo and A. Zakki Falani, “Pemanfaatan Data Mining Untuk *Clustering* Potensi Produksi Beras Di Kabupaten Blitar Dengan Menggunakan Metode Fuzzy C-Means,” *Nopember*, vol. 12, no. 2, pp. 30–36, 2020, [Online]. Available: <https://www.blitarkab.go.id/%0Ahttp://jurnal.stmik-yadika.ac.id/index.php/spirit/article/view/181>
- [15] S. Mulyani, B. N. Sari, and A. A. Ridha, “Clustering Productivity of Rice in Karawang Regency Using the Fuzzy C-Means Method,” *Indones. J. Artif. Intell. Data Min.*, vol. 3, no. 2, pp. 103–112, 2020, [Online]. Available: <https://ejournal.uin-suska.ac.id/index.php/IJAIDM/article/view/10415>
- [16] M. Orisa and A. Faisol, “Journal of Computer Networks , Architecture and High Performance Computing Researchers Productivity Level Clustering Based On H-Index and Citation Using The Fuzzy C-Means Algorithm Journal of Computer Networks , Architecture and High Performance Computing,” vol. 5, no. 1, pp. 18–26, 2023.

- [17] S. Solimun and A. A. R. Fernades, “Cluster Analysis on Various Cluster Validity Indexes with Average Linkage Method and Euclidean Distance (Study on Compliant Paying Behavior of Bank X Customers in Indonesia 2021),” *Math. Stat.*, vol. 10, no. 4, pp. 747–753, 2022, doi: 10.13189/ms.2022.100405.
- [18] D. A. I. C. Dewi and D. A. K. Pramita, “Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali,” *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- [19] F. I. Dwiana, W. D. Utami, and A. Hamid, “Implementasi K-Means pada *Clustering* Jenis Disabilitas,” vol. 6, no. 1, pp. 92–101, 2023.
- [20] M. F. Febrian, “PENERAPAN FUZZY C-MEANS DAN FUZZY TIME SERIES UNTUK ANALISIS TINGKAT NILAI SAHAM BEI TAHUN 2015-2020,” 2021.
- [21] S. Mashfuufah and D. Istiawan, “Penerapan Partition Entropy Index, Partition Coefficient Index dan Xie Beni Index untuk Penentuan Jumlah *Cluster* Optimal pada Algoritma Fuzzy C-Means dalam Pemetaan Tingkat Kesejahteraan Penduduk Jawa Tengah,” *Proceeding of The URECOL*, pp. 51–60, 2018.

