CICES (*Cyberpreneurship Innovative and Creative Exact and Social Science*)
DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

Klasterisasi Tingkat Kemiskinan Menggunakan *Fuzzy c-means* dengan Optimalisasi *Entropy weight* dan PCA

Jihan Octavia Salsabillah Hidayat*[1], Ani Dijah Rahajoe[2], Muhammad Muharrom Al Haromainy[3]
[1,3]Program Studi Informatika, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur, [2]Magister Teknologi Informasi, Fakultas Ilmu Komputer, UPN "Veteran" Jawa Timur
E-mail: *[1]**21081010034@student.upnjatim.ac.id**, [2]anidijah.if@upnjatim.ac.id, [3]muhammad.muharrom.if@upnjatim.ac.id

Abstrak

Kemiskinan merupakan permasalahan utama di Provinsi Jawa Timur, yang memiliki tingkat kemiskinan tertinggi ketiga di Indonesia. Untuk memahami pola dan karakteristik kemiskinan, penelitian ini melakukan clustering tingkat kemiskinan kabupaten/kota menggunakan metode Fuzzy c-means dengan optimalisasi entropy weight dan PCA. Analisis diawali dengan seleksi fitur menggunakan entropy weight, yang dari 14 variabel menghasilkan 6 variabel utama: jumlah Penduduk miskin, persentase penduduk miskin, P1, P2, garis kemiskinan, dan pengeluaran per kapita. Reduksi dimensi dengan PCA menghasilkan dua komponen utama (PC1 dan PC2) yang menjelaskan 86,01% informasi dataset. Clustering dengan FCM menghasilkan 3 cluster optimal berdasarkan silhouette score sebesar 0,553: cluster 1 (10 kabupaten/kota dengan tingkat kemiskinan terendah), cluster 2 (21 kabupaten/kota berstatus ekonomi menengah), dan cluster 3 (7 kabupaten/kota dengan kemiskinan tertinggi). Evaluasi model menunjukkan nilai silhouette score 0,55, yang menandakan clustering layak, dengan cluster 1 memiliki keanggotaan paling seragam, sedangkan cluster 2 dan 3 memiliki penyebaran sosial-ekonomi yang lebih luas.

Kata Kunci— *Fuzzy c-means, Entropy weight, Principal Component Analysis, Clustering, Poverty.*

Abstract

Poverty is a major issue in East Java Province, which has the third-highest poverty rate in Indonesia. To understand the patterns and characteristics of poverty, this study performs clustering of poverty levels in districts/cities using the Fuzzy c-means method with entropy weight and PCA optimization. The analysis begins with feature selection using the entropy weight method, which selects 6 main variables from 14 initial ones: number of poor population, percentage of poor population, P1, P2, poverty line, and per capita expenditure. Dimensionality reduction using PCA yields two principal components (PC1 and PC2), explaining 86.01% of the total dataset information. Clustering using FCM results in 3 optimal Clusters based on a silhouette score of 0.553: cluster 1 (10 districts/cities with the lowest poverty level), cluster 2 (21 districts/cities with a middle economic status), and cluster 3 (7 districts/cities with the highest poverty level).

Model evaluation shows a silhouette score of 0.55, indicating that the clustering is appropriate, with cluster 1 having the most consistent membership, while clusters 2 and 3 exhibit wider socio-economic distribution.

Keywords— *Fuzzy c-means, Entropy weight, Principal Component Analysis, Clustering, Poverty.*

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

1. PENDAHULUAN

Seseorang maupun kelompok dianggap berada dalam kemiskinan jika mereka tidak mampu membayar kebutuhan pokok seperti makanan, pakaian, transportasi, dan pendidikan. Ketidakmampuan ini dapat disebabkan oleh keterbatasan sumber daya atau kurangnya akses terhadap pendidikan dan pekerjaan yang memadai [1]. Kemiskinan di negara berkembang seperti Indonesia merupakan tantangan ekonomi yang kompleks. Sebagai negara berkembang dengan jumlah penduduk terpadat keempat di dunia, Indonesia menghadapi persoalan serius dalam pengentasan kemiskinan [2].

Meski memiliki potensi besar dari segi letak geografis dan sumber daya, Indonesia masih menghadapi masalah ekonomi yang belum terselesaikan, termasuk kemiskinan [3]. Berdasarkan alinea keempat Pembukaan Undang – Undang Dasar 1945, Indonesia berkomitmen mencapai masyarakat adil dan makmur [4]. BPS melaporkan bahwasannya di tahun 2022, Pulau Jawa menunjukkan konsentrasi kemiskinan tertinggi, dengan Jawa Timur menempati urutan ketiga dengan tingkat kemiskinan mencapai 10,38% atau setara dengan 4.181.290 orang yang hidup dalam kemiskinan, setelah DI Yogyakarta di perangkat pertama dan Jawa Tengah di peringkat kedua [5].

Beragai kebijakan pengentasan kemiskinan telah diterapkan, namun hasilnya masih belum optimal. Oleh karena itu, penting dilakukan penelitian untuk mengidentifikasi indikator-indikator yang memengaruhi tingkat kemiskinan. Salah satu pendekatannya adalah penerapan metode *Fuzzy c-means* (FCM), yang sebelumnya telah digunakan oleh berbagai peneliti.

[6] menerapkan FCM untuk mengelompokkan kemiskinan kabupaten/kota di Jawa Tengah menjadi lima *cluster*. Hasilnya menunjukkan efektivitas FCM dalam menangkap variasi kemiskinan yang kompleks. [7] menggunakan FCM untuk klasterisasi kemiskinan di Aceh dan menghasilkan tiga *cluster*. Evaluasi model menunjukkan hasil yang cukup baik, membuktikan keunggulan FCM dalam menangani karakteristik fuzzy dalam data sosial ekonomi.

Selain FCM, metode lain seperti K-Means dan K-Medoids juga digunakan. [8] menggunakan K-Means untuk mengelompokkan provinsi berdasarkan 12 indikator, menghasilkan tiga kategori. Namun, K-Means memiliki keterbatasan karena kurang sensitif terhadap perbedaan kecil antar wilayah. [9] menerapkan K-Medoids di Jawa Barat, yang lebih stabil terhadap outlier dibanding K-Means.

Rinaldo dan Miftakhul menyimpulkan bahwa FCM lebih efektif dibanding K-Means dan K-Medoids karena dapat menangkap variasi kemiskinan yang tidak tegas batasnya. FCM memungkinkan satu daerah memiliki derajat keanggotaan dalam beberapa *cluster*, sehingga lebih fleksibel dan realistis. Keunggulan FCM terletak pada kemampuannya menangani ketidakpastian dalam data sosial-ekonomi, yang sifatnya bertahap dan berkelanjutan.

Namun, untuk hasil yang lebih optimal, FCM perlu dikombinasikan dengan metode lain seperti *entropy weight* yang memberi bobot objektif pada indikator berdasarkan variasi datanya. Indikator yang lebih bervariasi akan diberi bobot lebih besar, mengurangi subjektivitas dalam analisis.

Selain itu, untuk mengatasi tingginya dimensi data kemiskinan, digunakan *Principal Component Analysis* (PCA) yang mereduksi jumlah indikator dengan tetap mempertahankan informasi penting. PCA menggabungkan indikator yang berkorelasi ke dalam komponen utama dan mencegah terjadinya multiko*linear*itas, sehingga analisis menjadi lebih efisien.

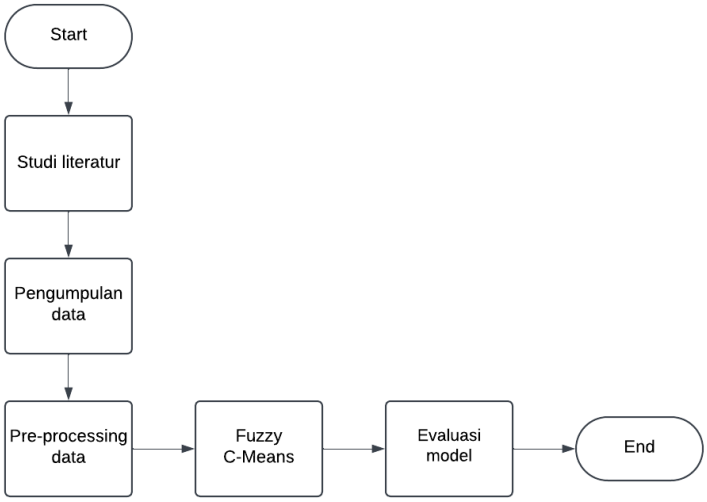
Kombinasi ketiga metode FCM, *entropy weight*, dan PCA dipilih karena saling melengkapi. Oleh karena itu, dilakukan penelitian berjudul "Klasterisasi Tingkat Kemiskinan Menggunakan *Fuzzy c-means* Dengan Optimalisasi *Entropy weight* dan PCA" untuk mengidentifikasi kemiskinan di kabupaten/kota di Jawa Timur. PCA menyederhanakan data, *entropy weight* memberikan bobot objektif, dan FCM mengelompokkan wilayah berdasarkan kemiskinan. Diharapkan pendekatan ini dapat memberikan gambaran yang lebih akurat dalam mendukung kebijakan pengentasan kemiskinan di Jawa Timur.

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

2. METODE PENELITIAN

2.1. Tahapan Penelitian

Adapun tahapan yang dilakukan untuk menyelesaikan penelitian ini bisa terlihat dalam gambar 1.



Gambar 1. Tahapan Penelitian

2.2. Studi Literatur

Sumber-sumber ilmiah yang kredibel, termasuk tesis, disertasi, makalah, dan jurnal nasional dan internasional, menjadi tinjauan literatur yang digunakan dalam penelitian ini. Pemilihan literatur dilakukan secara sistematis untuk membangun landasan teori, memahami konteks permasalahan, dan merujuk pada temuan terkini. Referensi yang digunakan mayoritas berasal dari publikasi lima tahun terakhir, termasuk setidaknya satu jurnal internasional terindeks atau jurnal nasional bereputasi (minimal SINTA). Selain menghimpun referensi, penelitian ini juga menganalisis kontribusi studi terdahulu untuk memperkuat argumen, metodologi, dan validitas hasil penelitian.

2.3. Pengumpulan Data

Dua sumber data utama dari Badan Pusat Statistik (BPS) digunakan dalam penelitian ini: publikasi data dan informasi kemiskinan kabupaten/kota dan laporan Indeks Pembangunan Manusia (IPM) Provinsi Jawa Timur, yang keduanya diakses melalui situs resmi BPS. Seluruh variabel dari publikasi kemiskinan digunakan secara menyeluruh, sementara dari IPM hanya variabel yang secara signifikan berpengaruh terhadap kemiskinan, sesuai arahan BPS. Dataset yang digunakan mencakup wilayah Provinsi Jawa Timur periode 2020–2023, terdiri atas 153 baris dan 14 variabel. Berikut ditampilkan parameter pada tabel 1.

Tabel 1. Parameter dan Deskripsi dari Dataset

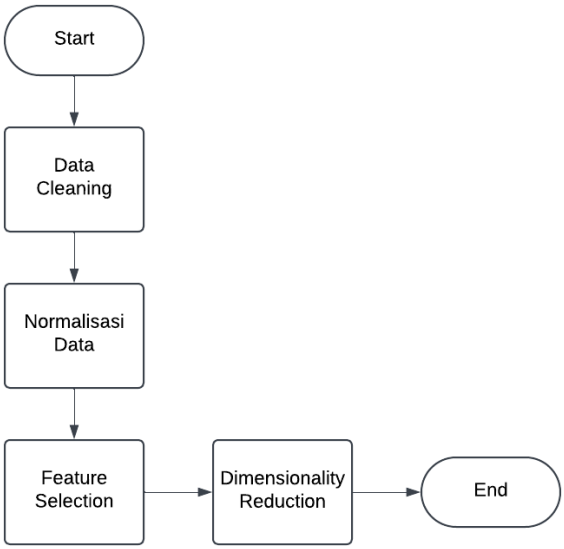
Parameter	**Deskripsi**
Kab/Kota	Kabupaten/Kota
Tahun	Tahun
Jmlh_pend_miskin	Jumlah Penduduk Miskin
Persentase_pend_miskin	Persentase Penduduk Miskin

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

P1	Indeks Kedalaman Kemiskinan
P2	Indeks Keparahan Kemiskinan
Garis Kemiskinan	Garis Kemiskinan (Rp/Kap/Bulan)
UHH	Umur Harapan Hidup
HLS	Harapan Lama Sekolah
RLS	Rata-rata Lama Sekolah
Pengeluaran_pkt	Pengeluaran per Kapita per Tahun
TPT	Tingkat Pengangguran Terbuka
TPAK	Tingkat Partisipasi Angkatan Kerja
IPM	Indeks Pembangunan Manusia

2.4. Preprocessing Data

Pada proses *preprocessing* data, sejumlah langkah dilakukan untuk mengolah data agar dapat digunakan secara optimal. Adapun tahapan *preprocessing* ini bisa terlihat dalam gambar 2.



Gambar 2. Diagram Tahapan *Preprocessing* Data

2.4.1. Data Cleaning

Sebuah langkah penting dalam analisis data yang meningkatkan kualitas data dengan menemukan dan menghilangkan data hilang, salah, atau tidak konsisten. Proses ini mencakup penanganan *missing values* melalui imputasi rata-rata, median, atau mode, koreksi *outliers*, serta konversi tipe data agar sesuai dengan analisis yang diinginkan. Pembersihan data yang baik memastikan hasil analisis lebih akurat, andal, dan mengurangi risiko kesalahan dalam pengambilan keputusan berbasis data.

2.4.2. Normalisasi Data

Metode normalisasi yang disebut normalisasi *min-max* menerapkan perubahan *linear* pada data asli. Dengan menggunakan teknik ini, nilai data asli digeser agar berada di antara 0 dan 1 [10].

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

$$x'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (1)$$

Keterangan:

- x_{ij}: Nilai asli dari variabel ke-*j* untuk ke-*j*.
- $\min(x_j)$: Nilai minimum dari variabel ke-*j*.
- $\max(x_j)$: Nilai maksimum dari variabel ke-*j*.
- x'_{ij}: Nilai normalisasi variabel ke-*j* untuk kabupaten dan kota ke-*j*.

2.4.3. Entropy weight

Entropy weight Method (EWM) merupakan teknik pembobotan untuk menilai tingkat distribusi informasi dari berbagai sumber dalam pengambilan keputusan, banyak digunakan dalam studi evaluasi menyeluruh. Indikator dengan nilai entropi lebih rendah menunjukkan distribusi lebih tinggi dan mendapat bobot lebih besar. EWM diperkenalkan oleh Shannon dan Weaver pada 1947 dan dikembangkan oleh Zeleny pada 1982. Metode ini menggunakan teori probabilitas untuk menghitung ketidakpastian informasi tanpa mempertimbangkan preferensi pengambil keputusan. Informasi dengan bobot lebih tinggi dianggap lebih bermanfaat [11]. Langkah – langkahnya meliputi:

1. Menyusun matriks keputusan berdasarkan nilai awal dari setiap indikator.
2. Standarisasi nilai dari setiap indikator menggunakan persamaan:

$$p_{ij} = \frac{x_{ij}}{\sum_j^n x_{ij}} \quad (2)$$

Di mana x_{ij} adalah nilai dari indeks ke-i dalam sampel ke-j. Langkah ini memastikan bahwa nilai dari setiap indikator dapat dibandingkan pada skala yang sama.

3. Perhitungan Entropi: Nilai entropi E_i untuk setiap indikator dihitung menggunakan persamaan:

$$E_i = \frac{\sum_{j=1}^n p_{ij} \cdot \ln p_{ij}}{\ln n} \quad (3)$$

Dalam implementasi EWM, jika $p_{ij} = 0$, maka $p_{ij} \cdot \ln p_{ij}$ ditetapkan sebagai nol untuk mempermudah perhitungan. Nilai entropi ini mengukur tingkat ketidakpastian atau keragaman dalam distribusi nilai dari setiap indikator.

4. Penentuan Bobot: Setelah nilai entropi diperoleh, bobot untuk setiap indikator dihitung menggunakan rumus:

$$w_{ij} = \frac{1 - E_i}{\sum_i^m (1 - E_i)} \quad (4)$$

Rentang nilai entropi E_i berada dalam interval [0, 1]. Semakin besar nilai E_i, semakin besar bobot indikator tersebut karena menunjukkan bahwa indikator memiliki distribusi informasi yang lebih signifikan.

2.4.4. Principal Component Analysis (PCA)

Metode PCA bertujuan menyederhanakan data kompleks dan mengatasi multiko*linear*itas dengan mereduksi dimensi data melalui pembentukan komponen baru yang mempertahankan total variasi variabel asli. Komponen baru tersebut saling independen dan mewakili informasi dari variabel-variabel dalam data. Penentuan jumlah komponen utama didasarkan pada beberapa kriteria [12]:

1) Scree plot: Komponen utama dipilih ketika garis kurva mulai melandai.

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

2) Kumulatif proporsi varians: Komponen utama dianggap cukup jika proporsi kumulatif varians mencapai ≥ 80%.
3) Eigenvalue: Komponen dengan eigenvalue > 1 dipilih sebagai komponen utama.

Langkah-langkah metode PCA:

1. Standarisasi Data: Jika variabel memiliki skala berbeda, data distandarisasi agar memiliki rata-rata 0 dan varians 1, dengan rumus:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (5)$$

di mana x_{ij} adalah nilai data, μ_j adalah rata-rata variabel, dan σ_j adalah deviasi standar variabel.

2. Menghitung Matriks Kovarians atau Korelasi

$$\sum = \frac{1}{n-1} Z^T Z \quad (6)$$

di mana Σ adalah matriks kovarians, Z adalah matriks data standar, dan n adalah jumlah observasi.

3. Menghitung Nilai Eigen dan Vektor Eigen
Nilai eigen (λ) dan vektor eigen (v) diperoleh dengan menyelesaikan persamaan:

$$\sum v = \lambda v \quad (7)$$

Nilai eigen menunjukkan varians yang dijelaskan oleh setiap komponen utama, sedangkan vektor eigen menentukan arah komponen utama.

4. Membentuk Komponen Utama
Data asli diproyeksikan ke vektor eigen untuk membentuk komponen utama:

$$PC_1 = Z \cdot v_1 \quad (8)$$

Proses ini dapat diulang untuk membentuk komponen utama berikutnya PC_2, PC_3 dan seterusnya.

2.4.5. Fuzzy c-means (FCM)

Jim Bezdek pertama kali mempresentasikan teknik pengelompokan *Fuzzy c-means* (FCM) pada tahun 1981. Pendekatan ini menilai tingkat hubungan setiap titik data dengan sebuah *cluster*, seperti yang ditunjukkan oleh derajat keanggotaannya [13]. Ketika membuat sistem inferensi fuzzy, *Fuzzy c-means* (FCM) menghasilkan hasil yang mencakup derajat keanggotaan dengan pusat *cluster* untuk tiap titik data [14].

Posisi awal dari pusat *cluster* sering kali kurang tepat, sehingga membutuhkan perbaikan berulang hingga titik optimal tercapai. Akibatnya, setiap titik data akan memiliki tingkat asosiasi yang unik dengan setiap *cluster* [16]. FCM menawarkan keuntungan yang signifikan dalam kemampuannya untuk memposisikan pusat *cluster* secara akurat jika dibandingkan dengan metode pengelompokan lainnya [14]. Berikut ialah algoritma *fuzzy c-means*:

1. Inisialisasi parameter: termasuk jumlah *Cluster* (*c*), pangkat pembobot (*m*), jumlah iterasi maksimum, dan batas toleransi error (ξ).
2. Membangkitkan matriks partisi awal: secara acak dengan derajat keanggotaan μ_{ik}, di mana jumlah keanggotaan dalam satu baris adalah 1.
3. Menghitung pusat *Cluster* (V_{kj}) dengan:

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

$$V_{kj} = \frac{\sum_{i=1}^{n}(\mu_{ik})^m X_{kj}}{\sum_{i=1}^{n}(\mu_{ik})^m} \quad (9)$$

Dimana :

V_{kj} = Pusat klister

X_{ij} = Atribut

W = Pangkat

4. Menghitung fungsi objektif (P_t) untuk mengevaluasi jarak titik data terhadap pusat *Cluster*:

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right](\mu_{ik})^m\right) \quad (10)$$

Dimana P_t adalah nilai fungsi objektif iterasi ke-t

5. Memperbarui matriks partisi (μ_{ik}) dengan:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2\right]^{\frac{-1}{m-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(X_{ij}-V_{kj})^2\right]^{\frac{-1}{m-1}}} \quad (11)$$

Dimana i = 1,2,…,n dan k = 1,2,…,c.

6. Memeriksa kondisi konvergensi berdasarkan perubahan nilai fungsi objektif ($|P_t - P_{t-1}| < \xi$) atau hingga jumlah iterasi maksimum tercapai.

2.5. Silhouette score

Alat penilaian internal yang disebut *silhouette score* diterapkan guna mengevaluasi kualitas pengelompokan dengan mengukur kedekatan tiap titik data terhadap *cluster*nya dibandingkan *cluster* lain. Nilai score berkisar antara -1 hingga 1, di mana nilai tinggi menunjukkan keterkaitan erat dengan pusat *cluster*nya, sedangkan nilai negatif mengindikasikan potensi masalah dalam pengelompokan [15].

Tabel 2. Tabel kriteria *Silhouette Coefficient*

Nilai Sillhouette Coefficient	Definisi
0.71 s.d 1.00	*Cluster* yang kuat telah terbentuk
0,51 s.d. 0,70	*Cluster* sudah sesuai atau praktis
0,26 s.d. 0,50	*Cluster* yang terbentuk dengan lemah
≤ 0,25	Tidak dapat diklasifikasikan sebagai *cluster*

Tabel 2 menunjukkan bahwa nilai koefisien *silhouette* mendekati 1 mengindikasikan struktur *Cluster* yang kuat dan hasil pengelompokan yang akurat. Sebaliknya, nilai mendekati 0 menunjukkan kemungkinan *misclustering* atau data yang menyulitkan pengelompokan.

3. HASIL DAN PEMBAHASAN

3.1 Pre-processing

Tahapan *preprocessing* mencakup data cleaning, normalisasi data, *feature selection, dimensionality reduction.* Proses ini memastikan bahwa data yang digunakan dalam analisis siap untuk diklasterisasi, sehingga hasil yang diperoleh lebih akurat dan representatif.

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

3.1.1 Data Cleaning

Dalam tahapan ini, dilakukan data cleaning dengan menghapus baris atau kolom kosong atau bernilai NaN. Namun, karena data dari BPS sudah bersih, tidak dilakukan penghapusan atau imputasi. Data tahunan 2020–2023 yang terpisah digabung melalui agregasi rata-rata per variabel pada tingkat kabupaten/kota untuk memperoleh representasi yang lebih stabil. Tujuannya adalah mengurangi fluktuasi tahunan akibat faktor eksternal agar hasil *clustering* mencerminkan pola kemiskinan struktural. Jika dilakukan per tahun, pola dapat berubah karena dinamika sosial-ekonomi. Oleh karena itu, agregasi dipilih agar analisis lebih fokus pada tren jangka panjang. Kolom tahun tidak disertakan dan dihapus, sehingga tiap nilai variabel merupakan rata-rata seluruh tahun. Hasil akhir ditampilkan pada gambar 3 dan digunakan dalam tahap berikutnya.

	Kab/Kota	Jmlh_pend_miskin	Persentase_pend_miskin	P1	P2	Garis Kemiskinan	UHH	HLS	RLS	Pengeluaran_pkt	TPT	TPAK	IPM
2	Bangkalan	203,185	20,23	3,5075	0,925	453228,75	72,95	11,8025	5,9675	8923	7,7675	70,2525	65,88
3	Banyuwangi	125,7075	7,745	1,1725	0,2525	406142,5	73,59	13,0325	7,5	12374,25	5,1925	73,79	72,8525
4	Blitar	106,2625	9,095	1,185	0,245	351123	74,78	12,595	7,635	10977,75	4,46	71,89	71,8975
5	Bojonegoro	158,5675	12,6325	1,86	0,4225	395593,75	74,3925	12,7075	7,3975	10360,25	4,765	73,2225	70,9575
6	Bondowoso	109,06	13,9275	2,0075	0,4575	443815,5	72,8725	13,3	6,1125	10851,5	4,265	74,495	69,7025
7	Gresik	157,4475	11,71	2,115	0,67	530746,75	73,8825	13,8575	9,655	13445	7,7175	68,595	77,4075
8	Jember	243,5675	9,85	1,3125	0,2875	396956	73,695	13,445	6,4975	9705,25	4,6575	69,6875	69,6075
9	Jombang	121,52	9,5325	1,405	0,3075	442218	74,09	13,5025	8,655	11558,25	6,175	69,87	74,3875
10	Kediri	176,265	11,1025	1,5225	0,3325	349778	74,5	13,455	8,1425	11411	5,7525	69,365	73,7475
11	Kota Batu	7,975	3,77	0,4925	0,105	552675	74,7875	14,3125	9,465	13102	6,3625	74,1425	77,8075
12	Kota Blitar	10,9225	7,585	0,9375	0,2025	508185	74,6325	14,445	10,4725	14038,75	5,98	69,8825	79,77
13	Kota Kediri	21,73	7,455	1,1225	0,255	531355,75	75,38	15,355	10,305	12659	5,255	68,7775	79,755
14	Kota Madiun	8,71	4,8925	0,6225	0,1375	542197	75,0425	14,42	11,5	16432,75	7,1775	67,91	82,7325
15	Kota Malang	38,9325	4,4225	0,8225	0,2225	602325,25	74,98	15,6975	10,555	16843,75	8,43	66,165	83,085
16	Kota Mojokerto	7,9975	6,095	0,7825	0,1525	522069,5	75,455	14,0175	10,6425	13896,25	5,8475	69,3025	79,845
17	Kota Pasuruan	13,4875	6,6275	1,03	0,24	479890	74,3	13,6375	9,475	13672	6,095	71,3525	77,3275
18	Kota Probolinggo	16,9125	7	1,065	0,23	568514,5	73,7575	13,6475	9,125	12498,75	5,5875	69,445	75,755
19	Kota Surabaya	143,185	4,905	0,755	0,1875	643628,25	75,5275	14,8225	10,55	18234,75	8,4625	68,5175	83,2525
20	Lamongan	158,13	13,165	2,3175	0,605	436464,75	74,52	13,82	8,1575	11658,25	5,385	71,4775	74,3825
21	Lumajang	99,1775	9,4675	1,365	0,29	348218,25	74,0775	11,9675	6,77	9369,25	3,8775	67,8375	68,135
22	Madiun	77,4625	11,3	1,605	0,3525	396837	74,2575	13,185	7,88	11834,75	5,1925	71,1	73,385
23	Magetan	64,495	10,1625	1,35	0,2775	388504,25	74,845	14,0475	8,4825	12033,75	4,0225	74,7625	75,4975

Gambar 3. Dataset setelah Data *Cleaning*

3.1.2 Normalisasi Data

Setelah tahap pembersihan data, langkah berikutnya adalah normalisasi menggunakan metode *min-max scaling*. Data tingkat kabupaten dan kota untuk sementara tidak disertakan, karena normalisasi hanya dilakukan pada variabel numerik. Gambar 4 menunjukkan hasil normalisasi melalui penskalaan *min-max*.

	Jmlh_pend_miskin	Persentase_pend_miskin	P1	P2	Garis Kemiskinan	UHH	HLS	RLS	Pengeluaran_pkt	TPT	TPAK	IPM
2	0,769694819	0,879861018	0,78721	0,65339	0,408799563	0,02919	0	0,15405	0	0,74992	0,28017	0,03177
3	0,464208264	0,212481625	0,17755	0,11753	0,262594287	0,27024	0,31579	0,38838	0,370633877	0,41188	0,50881	0,42037
4	0,387538443	0,284645196	0,18081	0,11155	0,091755853	0,71846	0,20347	0,40902	0,220662067	0,31572	0,38601	0,36715
5	0,593772179	0,473740478	0,35705	0,25299	0,229839865	0,5725	0,23235	0,37271	0,154348001	0,35576	0,47213	0,31476
6	0,398568725	0,542964052	0,39556	0,28088	0,379570928	0	0,38447	0,17622	0,207103928	0,29012	0,55437	0,24481
7	0,589356123	0,424428705	0,42363	0,4502	0,64949702	0,38041	0,5276	0,71789	0,485623003	0,74335	0,17305	0,67424
8	0,928919249	0,325003341	0,2141	0,14542	0,234069722	0,30979	0,42169	0,23509	0,084006766	0,34165	0,24366	0,23952
9	0,447697342	0,308031538	0,23825	0,16135	0,374610607	0,45857	0,43646	0,56498	0,283002658	0,54086	0,25545	0,50592
10	0,66355177	0,391955098	0,26893	0,18127	0,087579557	0,61299	0,42426	0,48662	0,267189304	0,4854	0,22281	0,47025
11	0	0	0	0	0,717585387	0,72128	0,64442	0,68884	0,448787822	0,56547	0,53159	0,69653
12	0,011621718	0,203928906	0,11619	0,07769	0,579441603	0,6629	0,67843	0,84289	0,549386528	0,51526	0,25626	0,80591
13	0,054234682	0,196979821	0,16449	0,11952	0,651387997	0,94444	0,91207	0,81728	0,401213521	0,42009	0,18484	0,80507
14	0,002898036	0,060002673	0,03394	0,0259	0,685050647	0,81733	0,67202	1	0,806481059	0,67246	0,12878	0,97102
15	0,122062535	0,034879059	0,08616	0,09363	0,871752018	0,79379	1	0,8555	0,850618842	0,83689	0,016	0,99066
16	8,87154E-05	0,124281705	0,07572	0,03785	0,622553704	0,97269	0,56868	0,86888	0,534083282	0,49787	0,21878	0,81009
17	0,021735273	0,152746225	0,14034	0,10757	0,491584143	0,53766	0,47112	0,69037	0,510000805	0,53036	0,35127	0,66978
18	0,035239729	0,172658025	0,14948	0,0996	0,766767866	0,33333	0,47368	0,63685	0,384004081	0,46373	0,22799	0,58214
19	0,533120416	0,060670854	0,06854	0,06574	1	1	0,77535	0,85474	1	0,84116	0,16804	1
20	0,592047157	0,502204998	0,4765	0,39841	0,356746464	0,62053	0,51797	0,48891	0,293741778	0,43715	0,35935	0,50564
21	0,359602949	0,304556996	0,22781	0,14741	0,082736452	0,45386	0,04236	0,27676	0,047923323	0,23925	0,12409	0,15745
22	0,27398273	0,402512361	0,29047	0,19721	0,233700221	0,52166	0,35494	0,44648	0,312696325	0,41188	0,33495	0,45005
23	0,222853087	0,341707871	0,22389	0,13745	0,207826595	0,74294	0,57638	0,53861	0,334067173	0,25829	0,57166	0,56779
24	1	0,328344247	0,17689	0,10558	0,130246859	0,72505	0,3896	0,39908	0,150777244	0,49032	0,28244	0,37216

Gambar 4. Dataset setelah Normalisasi

3.1.3 Feature Selection

Setelah proses normalisasi data, tahap selanjutnya adalah seleksi fitur menggunakan metode *entropy weight* yang bertujuan menentukan bobot masing-masing variabel berdasarkan tingkat entropinya. Dataset dinormalisasi terlebih dahulu agar seluruh nilai berada pada skala yang sama, dilanjutkan dengan perhitungan probabilitas relatif, entropi tiap fitur, dan bobot

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

yang dinormalisasi. Hasil seleksi fitur ditampilkan pada gambar 5, dan bobot ditunjukkan dalam tabel 3. Variabel jmlh_pend_miskin, persentase_pend_miskin, P1, P2, garis kemiskinan, dan pengeluaran_pkt memiliki bobot lebih tinggi dan dipilih untuk proses *clustering* karena kontribusinya lebih besar dalam analisis kemiskinan.

Tabel 3. Bobot Setiap Variabel

Variabel	Bobot
Jmlh_pend_miskin	0,106
Persentase_pend_miskin	0,084
P1	0,1
P2	0,132
Garis Kemiskinan	0,088
UHH	0,066
HLS	0,064
RLS	0,053
Pengeluaran_pkt	0,121
TPT	0,059
TPAK	0,062
IPM	0,066

```
Fitur yang dipilih untuk clustering: ['Jmlh_pend_miskin', 'Persentase_pend_miskin', 'P1', 'P2', 'Garis Kemiskinan', 'Pengeluaran_pkt']
```

Gambar 5. Hasil Feature Selection

3.1.3 Dimensionality Reduction

Setelah seleksi fitur, dilakukan reduksi dimensi menggunakan PCA untuk menyederhanakan kompleksitas data tanpa kehilangan informasi yang signifikan. Hasil PCA menunjukkan bahwa PC1 menjelaskan 66,01% varians, sedangkan PC2 menjelaskan 20%, sehingga secara keseluruhan dua komponen utama (PC1 dan PC2) mampu merepresentasikan 86,01% informasi dataset. Pada gambar 6 menunjukkan variabel yang paling berpengaruh dalam komponen utama yang dipilih, di mana persentase penduduk miskin memiliki bobot tertinggi dalam PC1 (0,474) dan garis kemiskinan memiliki bobot tertinggi dalam PC2 (0,701). Hal ini menunjukkan bahwa kedua variabel tersebut memainkan peran penting dalam menentukan distribusi kemiskinan di kabupaten/kota Jawa Timur.

```
Variabel yang Paling Berpengaruh pada Setiap Komponen Utama:
PC1: Persentase_pend_miskin dengan nilai bobot 0.4737
PC2: Garis Kemiskinan dengan nilai bobot 0.7012
```

Gambar 6. Variabel yang paling Berpengaruh

3.2 Pemilihan Jumlah Cluster

Dalam menentukan jumlah *cluster* yang optimal, digunakan *silhouette score*, yang mengukur seberapa baik setiap data dikelompokkan dalam *cluster* masing-masing dibandingkan

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

dengan *cluster* lain. Hasil pengujian untuk jumlah *cluster* 2 hingga 5 menunjukkan bahwa nilai *silhouette* tertinggi diperoleh pada *cluster* 3 dengan *silhouette score* sebesar 0,55, sehingga tiga *cluster* menjadi konfigurasi terbaik. Pada tabel 4 menampilkan nilai *silhouette score* untuk berbagai jumlah *cluster*, di mana terlihat bahwa *cluster* 3 memiliki nilai tertinggi, sementara jumlah *cluster* yang lebih besar justru menunjukkan nilai *silhouette* yang lebih rendah. Gambar 9 menampilkan visualisasi hasil *silhouette analysis*, yang menunjukkan bagaimana objek-objek dalam setiap *cluster* dikelompokkan berdasarkan nilai *silhouette* mereka.

Tabel 4. Nilai *Silhouette* untuk Setiap Nilai Jumlah K *Cluster*

Jumlah *Cluster*	***Silhouette score***
2	0.4898
3	**0.5533**
4	0.4351
5	0.4174

3.3 Fuzzy c-means

Metode *Fuzzy c-means* (FCM) digunakan untuk mengelompokkan kabupaten/kota berdasarkan karakteristik kemiskinan. Hasil *clustering* membentuk tiga *cluster* utama, yang memiliki karakteristik sebagai berikut:

1. *Cluster* 1 (10 kabupaten/kota) memiliki tingkat kemiskinan terendah dan pengeluaran per kapita tertinggi, mencerminkan daerah dengan kesejahteraan relatif tinggi.
2. *Cluster* 2 (21 kabupaten/kota) memiliki karakteristik ekonomi menengah, yang mencerminkan ketimpangan dalam distribusi kesejahteraan.
3. *Cluster* 3 (7 kabupaten/kota) memiliki tingkat kemiskinan tertinggi dan pengeluaran per kapita terendah, mencerminkan keterbatasan akses terhadap sumber daya ekonomi dan sosial.

Pada tabel 5 menunjukkan daftar kabupaten/kota dalam setiap *cluster* berdasarkan hasil FCM, yang mengonfirmasi pola kemiskinan di berbagai wilayah di Jawa Timur.

Tabel 5. Hasil *Clustering*

Kab/Kota	*Cluster*_1	*Cluster*_2	*Cluster*_3	*Cluster*
Kota Batu	0,936738897	0,048131592	0,015129511	1
Kota Blitar	0,974750233	0,019924819	0,005324948	1
Kota Kediri	0,9364555	0,051052596	0,012491904	1
Kota Madiun	0,931745802	0,049300187	0,018954011	1
Kota Malang	0,879384699	0,08251627	0,038099031	1
Kota Mojokerto	0,971255758	0,022295745	0,006448497	1
Kota Pasuruan	0,910957782	0,072270649	0,016771569	1

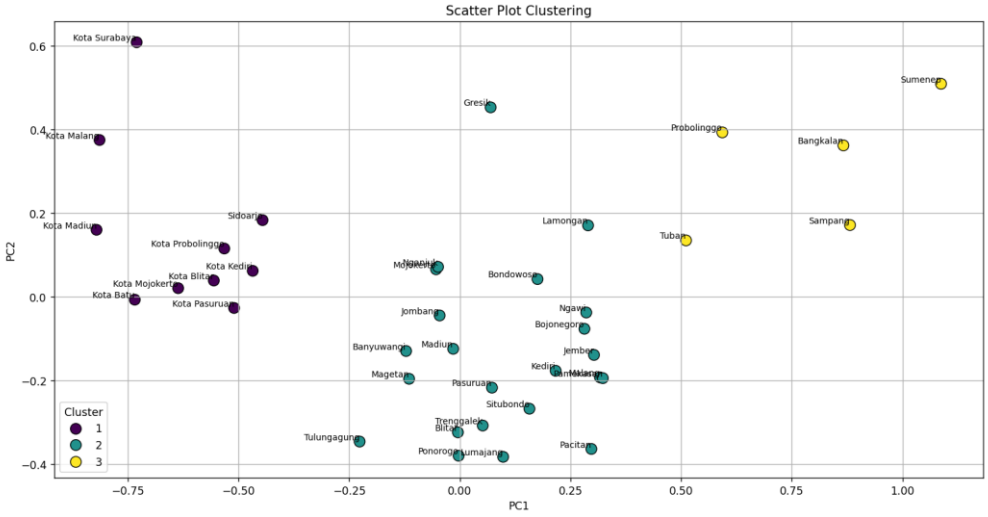
DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

Kota Probolinggo	0,988884053	0,008638737	0,00247721	1
Kota Surabaya	0,771450129	0,145856494	0,082693377	1
Sidoarjo	0,929054963	0,054883112	0,016061926	1
Banyuwangi	0,149239529	0,806097705	0,044662766	2
Blitar	0,059223054	0,90771784	0,033059106	2
Bojonegoro	0,041684978	0,870245573	0,088069449	2
Bondowoso	0,06870936	0,83221544	0,0990752	2
Gresik	0,28528053	0,400373027	0,314346443	2
Jember	0,03899601	0,878717687	0,082286303	2
Jombang	0,100953063	0,855966603	0,043080334	2
Kediri	0,015141044	0,963713674	0,021145282	2
Lamongan	0,105386284	0,55439612	0,340217596	2
Lumajang	0,054870666	0,901171038	0,043958296	2
Madiun	0,040146042	0,939585746	0,020268212	2
Magetan	0,123949534	0,8353603	0,040690167	2
Malang	0,041140767	0,874499838	0,084359395	2
Mojokerto	0,197764119	0,721063307	0,081172574	2
Nganjuk	0,197686821	0,718893269	0,08341991	2
Ngawi	0,050878275	0,832533581	0,116588143	2
Pacitan	0,059760874	0,84847215	0,091766975	2
Pamekasan	0,04282137	0,867421878	0,089756753	2
Pasuruan	0,006034878	0,989452334	0,004512788	2
Ponorogo	0,080369521	0,873606582	0,046023897	2
Situbondo	0,0159322	0,967698184	0,016369616	2
Trenggalek	0,034100229	0,942561649	0,023338121	2
Tulungagung	0,26290462	0,669547699	0,067547681	2
Bangkalan	0,00588985	0,015034887	0,979075264	3
Probolinggo	0,021252284	0,056996688	0,921751028	3
Sampang	0,015208017	0,046446562	0,938345421	3
Sumenep	0,041947419	0,087877558	0,870175022	3
Tuban	0,053715794	0,257516502	0,688767703	3

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

Hasil *Clustering* divisualisasikan dalam bentuk *scatter plot* sebagaimana ditunjukkan pada gambar 7. Visualisasi ini mempermudah pemahaman persebaran data dalam masing-masing *cluster* berdasarkan dua *principal component* utama. Terlihat bahwa *cluster* 1, 2, dan 3 terpisah cukup jelas, meskipun terdapat beberapa titik yang berada dekat batas antar *cluster*. Hal ini mencerminkan adanya kemiripan karakteristik antar daerah yang berada di perbatasan dua kelompok.

Cluster 1 menunjukkan distribusi yang lebih rapat, menandakan kesamaan karakteristik yang tinggi antar daerah di dalamnya. Sementara itu, *cluster* 2 lebih tersebar, mencerminkan variasi karakteristik ekonomi yang lebih besar. Adapun *cluster* 3 memiliki beberapa titik yang jauh dari pusat *cluster*, yang mengindikasikan keberadaan daerah dengan karakteristik yang lebih unik dalam kelompok tersebut.

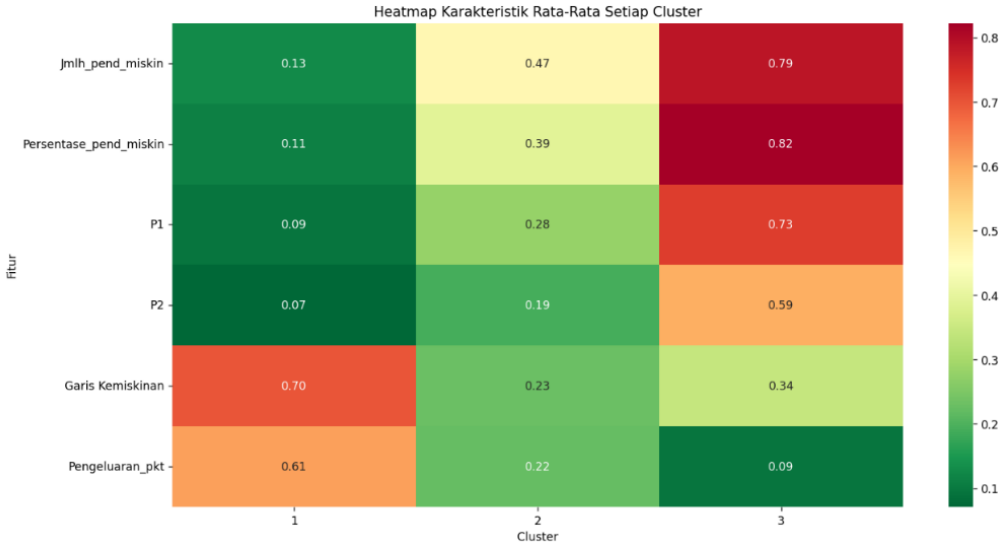


Gambar 7. *Scatter Plot*

3.4 Karakteristik Cluster

Analisis karakteristik dilakukan dengan menghitung rata-rata setiap karakteristik dalam masing-masing *cluster*, lalu data diurutkan dari terendah ke tertinggi berdasarkan *cluster* dan informasi kabupaten/kota. Dalam memahami perbedaan antar *cluster*, digunakan visualisasi heatmap yang terlihat pada gambar 8, di mana warna hijau menandakan nilai rendah, hijau muda hingga kuning nilai sedang, dan oranye hingga merah nilai tinggi. *cluster* 1 didominasi warna hijau (nilai rendah), *cluster* 3 warna oranye-merah (nilai tinggi), dan *cluster* 2 warna hijau muda-kuning (nilai sedang). Interpretasi ini menunjukkan bahwa heatmap menggambarkan pengaruh variabel terhadap karakteristik tiap *cluster*, memperkuat bahwa *cluster* 1 mencerminkan kemiskinan rendah, *cluster* 2 sedang, dan *cluster* 3 tinggi.

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

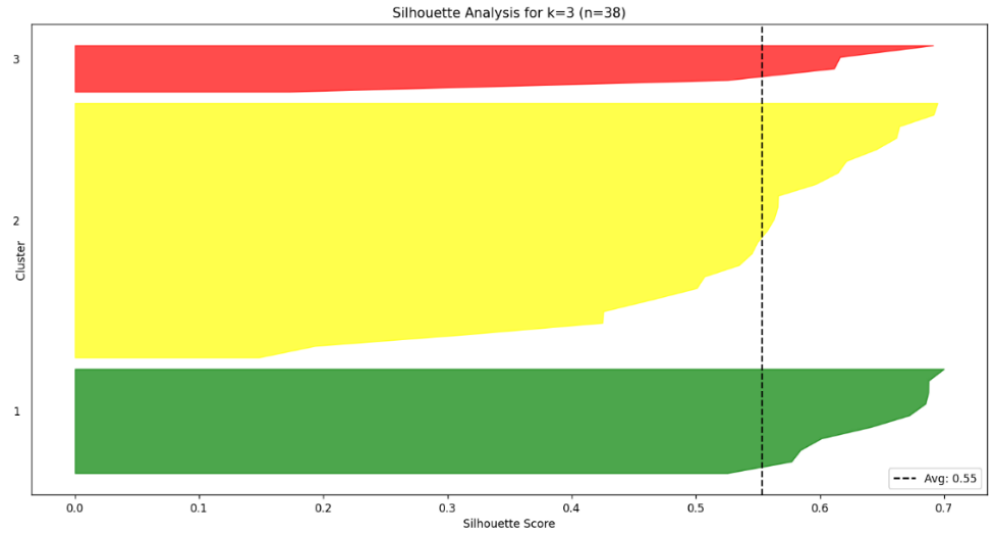


Gambar 8. *Heatmap* Karakteristik Setiap *Cluster*

3.5 Evaluasi Model

Evaluasi model *clustering* menggunakan *silhouette score* menghasilkan nilai 0.55 yang terlihat pada tabel 4 yang masuk kategori "*Cluster* telah layak atau sesuai" (0.51–0.70), menunjukkan hasil *clustering* cukup baik. Visualisasi *silhouette* analysis yang bisa terlihat pada gambar 9 menunjukkan *cluster* 2 memiliki distribusi lebih luas dibanding *cluster* 1 dan 3, mengindikasikan keragaman data lebih tinggi, namun masih dalam batas wajar kategori "layak".

Cluster 1 menunjukkan nilai *silhouette* yang lebih seragam dan konsisten, sementara *cluster* 3 memiliki rata-rata nilai lebih rendah dari *cluster* 1, tetapi tetap stabil. Garis vertikal putus-putus menandai rata-rata *silhouette score*, dan sebagian besar data berada di atas garis ini, menandakan anggota *cluster* telah tergolong dengan baik.



Gambar 9. *Silhouette Analysis*

Berdasarkan hasil ini, pengelompokan dengan metode *fuzzy c-means* cukup baik dalam membagi data ke dalam 3 *cluster* berbeda. Meski terdapat variasi kualitas antar *cluster*, hasil akhir menunjukkan pemisahan yang cukup jelas.

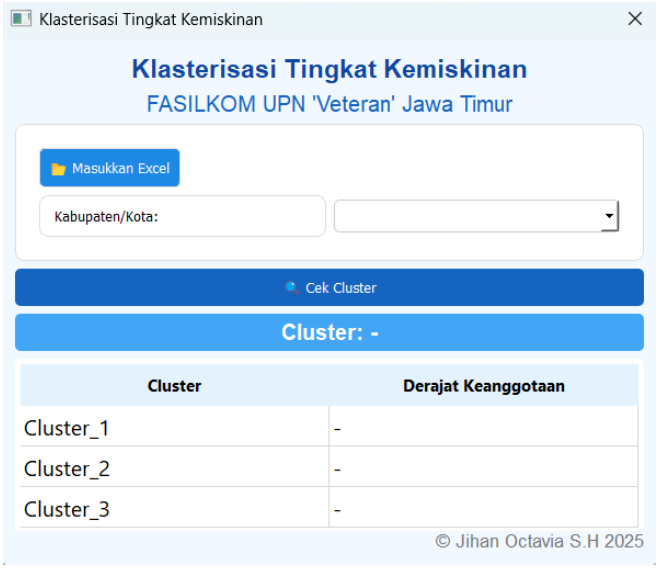
DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**

3.6 Antarmuka Pengguna (GUI)

Penelitian ini membuat *Graphical User Interface* (*GUI*) berbasis desktop (.exe) untuk memudahkan pengguna memperoleh hasil *clustering* tingkat kemiskinan secara praktis tanpa menjalankan kode secara manual. *GUI* ini dikembangkan menggunakan PyQt5, pandas, scikit-learn, dan fcmeans, dengan fitur unggah *file Excel*, pemilihan kabupaten/kota, serta tombol untuk menampilkan hasil *clustering*.

Antarmuka ini dirancang untuk memudahkan pengguna non-teknis, seperti pemerintah daerah, dalam mengakses hasil *clustering* secara cepat dan praktis, dengan output berupa nomor *cluster* dan derajat keanggotaan.

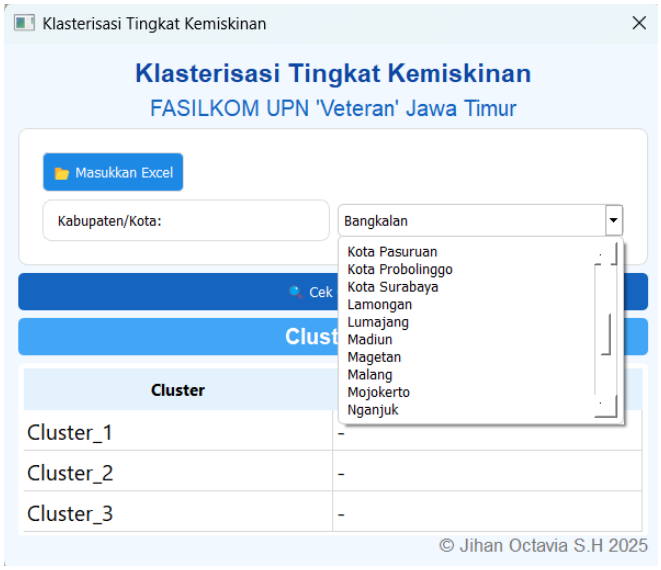
3.6.1 Tampilan GUI



Gambar 10. Tampilan Awal *GUI*

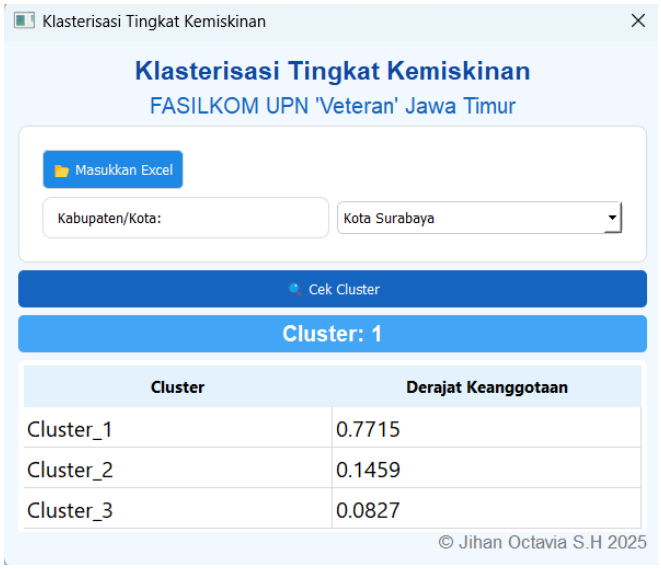
Tampilan awal *GUI* ditunjukkan pada gambar 10, memperlihatkan halaman antarmuka dengan judul "Klasterisasi Tingkat Kemiskinan" dan subjudul "FASILKOM UPN 'Veteran' Jawa Timur". Terdapat tombol "Masukkan Excel", dropdown kabupaten/kota, tombol "Cek *Cluster*", dan tabel kosong karena data belum dimasukkan. Setelah file Excel dimasukkan dan diproses, dropdown terisi daftar kabupaten/kota dari data, seperti pada gambar 11. Namun, hasil *Clustering* belum ditampilkan sebelum tombol "Cek *Cluster*" ditekan.

DOI: 10.33050/cices.v12i2 ISSN: **2356-5209**
pp. 226~242 *Online* ISSN: **2655-3058**



Gambar 11. Tampilan Setelah Upload Data

Setelah memilih kabupaten/kota dan menekan tombol "Cek *Cluster*", sistem memproses data dan menampilkan hasil *clustering* seperti pada gambar 11. Kota Surabaya terdeteksi masuk ke dalam *cluster*_1 dengan derajat keanggotaan 0.7715, 0.1459 untuk *cluster*_2, dan 0.0827 untuk *cluster*_3, menunjukkan keanggotaan tertinggi di *cluster*_1.



Gambar 12. Tampilan Akhir *GUI*

4. KESIMPULAN

Penelitian ini mengelompokkan kabupaten/kota di Provinsi Jawa Timur ke dalam tiga *Cluster* menggunakan indikator kemiskinan dan metode *Fuzzy c-means* (FCM). Analisis diawali dengan seleksi fitur menggunakan bobot entropi, menghasilkan enam variabel utama: jumlah penduduk miskin, persentase penduduk miskin, P1, P2, garis kemiskinan, dan pengeluaran per kapita.

PCA digunakan untuk reduksi dimensi, dengan dua komponen utama (PC1 dan PC2) menjelaskan 86,01% total informasi, di mana persentase penduduk miskin dan garis kemiskinan memiliki bobot tertinggi pada masing-masing komponen.

Pengelompokan dilakukan dengan FCM, dengan jumlah optimal tiga *cluster* berdasarkan *silhouette score* 0,5533. *Cluster* 1 berisi 10 kabupaten/kota dengan kemiskinan terendah dan pengeluaran tertinggi. *Cluster* 2 mencakup 21 kabupaten/kota dengan karakteristik ekonomi menengah. *Cluster* 3 terdiri dari 7 kabupaten/kota dengan kemiskinan tertinggi dan pengeluaran terendah.

Evaluasi menggunakan *silhouette score* menunjukkan nilai 0,55, menandakan model layak atau sesuai. *Cluster* 1 menunjukkan distribusi keanggotaan yang seragam, sedangkan *cluster* 2 memiliki variasi karakteristik ekonomi. Secara keseluruhan, kombinasi *entropy weight*, PCA, dan FCM efektif dalam mengelompokkan daerah berdasarkan karakteristik kemiskinan, dan hasilnya dapat digunakan sebagai dasar kebijakan pengentasan kemiskinan dan perencanaan pembangunan daerah.

5. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran untuk penelitian selanjutnya, sebagai berikut:

1. Penelitian selanjutnya dapat mempertimbangkan variabel tambahan yang lebih beragam untuk meningkatkan akurasi model dan mendapatkan gambaran sosial-ekonomi yang lebih komprehensif.
2. Antarmuka pengguna (*GUI*) dalam penelitian ini dapat dikembangkan lebih lanjut agar lebih interaktif dan responsif, dengan penambahan fitur validasi input serta visualisasi hasil *clustering* yang lebih informatif.

DAFTAR PUSTAKA

[1] R. Amalia and L. Rachmawati, "*PENGARUH INFLASI, PERTUMBUHAN EKONOMI DAN PENGANGGURAN TERHADAP TINGKAT KEMISKINAN DI KOTA SURABAYA.*" [Online]. Available: https://ejournal.unesa.ac.id/index.php/independent

[2] L. Khasanah, "*Dampak Ketimpangan Pendapatan, Tata Kelola Pemerintahan dan Korupsi terhadap Tingkat Kemiskinan di Indonesia*," *Bharanomics*, vol. 1, no. 2, pp. 75–81, Apr. 2021, doi: 10.46821/bharanomics.v1i2.156.

[3] D. Novitasari and P. Putri, "*EVALUASI PROGRAM PAHLAWAN EKONOMI DALAM MENGATASI KEMISKINAN DI KOTA SURABAYA*."

[4] N. Prawoto and J. L. Selatan, "*MEMAHAMI KEMISKINAN DAN STRATEGI PENANGGULANGANNYA*," 2009.

[5] A. P. Romadhona, D. I. Agustin, S. N. Nisa, and W. D. Maulana, "*Pengaruh Indeks Pembangunan Manusia (IPM) dan Tingkat Pengangguran Terhadap Kemiskinan di Jawa Timur Periode 2020-2022*," *Inisiat. J. Ekon. Akunt. dan Manaj.*, vol. 3, no. 1, pp. 52–66, 2024, [Online]. Available: https://doi.org/10.30640/inisiatif.v3i1.1968

[6] J. Tengah, "*Implementasi fuzzy c-means dalam pengelompokan tingkat kemiskinan pada kabupaten/kota di provinsi jawa tengah 1.*," vol. 10, no. 1, pp. 137–149, 2025.

[7] M. K. Sukardi and Indah Manfaati Nur, "*Penerapan Algoritma Fuzzy C-Means Untuk Pengelompokkan Kemiskinan Di Kabupaten/kota Provinsi Aceh*," *J. Ilm. Mat. Dan Terap.*, vol. 20, no. 2, pp. 115–126, 2023, doi: 10.22487/2540766x.2023.v20.i2.16494.

[8] A. Khalif, A. N. Hasanah, M. H. Ridwan, and B. N. Sari, "*Klasterisasi Tingkat Kemiskinan di Indonesia menggunakan Algoritma K-Means*," *Gener. J.*, vol. 8, no. 1, pp.

54–62, 2024, doi: 10.29407/gj.v8i1.21470.

[9] P. N. Safitri, R. Aristawidya, and S. B. Faradilla, "*Klasterisasi Faktor-Faktor Kemiskinan Di Provinsi Jawa Barat Menggunakan K-Medoids Clustering*," *J. Math. Educ. Sci.*, vol. 4, no. 2, pp. 75–80, 2021, doi: 10.32665/james.v4i2.242.

[10] Y. K. JAIN and S. K. BHANDARE, "*Min Max Normalization Based Data Perturbation Method for Privacy Protection*," *Int. J. Comput. Commun. Technol.*, vol. 4, no. 4, pp. 233–238, 2013, doi: 10.47893/ijcct.2013.1201.

[11] R. Kumar *et al.*, "*Revealing the benefits of entropy weights method for multi-objective optimization in machining operations: A critical review*," Jan. 01, 2021, *Elsevier Editora Ltda*. doi: 10.1016/j.jmrt.2020.12.114.

[12] M. Rais, R. Goejantoro, and S. Prangga, "*Optimalisasi K-Means Cluster dengan Principal Component Analysis pada Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Tingkat Pengangguran Terbuka*," *Eksponensial*, vol. 12, no. 2, p. 129, 2021, doi: 10.30872/eksponensial.v12i2.805.

[13] W. S. Mulyaningsih, "*IMPLEMENTASI FUZZY C-MEANS DAN FUZZY POSSIBILISTIC C-MEANS UNTUK PENGELOMPOKKAN KABUPATEN/KOTA DI PROVINSI BANTEN TUGAS AKHIR*."

[14] C. Ais, A. Hamid, and D. C. R. Novitasari, "*Analysis of Livestock Meat Production in Indonesia Using Fuzzy C-Means Clustering*," *J. Ilmu Komput. dan Inf.*, vol. 15, no. 1, pp. 1–8, Feb. 2022, doi: 10.21609/jiki.v15i1.993.

[15] K. Dbscan and Y. Hasan, "*Pengukuran Silhouette Score dan Davies-Bouldin Index pada Hasil Cluster*," vol. 06, no. 01, pp. 60–74, 2024.

[16] Utama, D. N. (2021). *Logika Fuzzy Untuk Model Penunjang Keputusan*. Garudhawaca.

[17] *Badan Pusat Statistik Jawa Timur*. (n.d.). Retrieved from Publikasi Indeks Pembangunan Manusia (IPM) Jawa Timut: https://jatim.bps.go.id/id/publication?keyword=indeks+pembangunan+manusia&sort=latest

[18] *Publikasi Data dan Informasi Kemiskinan Kabupaten/kota*. (2022). Retrieved from Badan Pusat Statistik: https://www.bps.go.id/id/publication?keyword=data+dan+informasi+kemiskinan+kabupaten+kota&sort=latest