



Perbandingan Metode Fuzzy C-Means Dan Fuzzy Probabilistic C-Means Dalam Pengelompokan Provinsi Di Indonesia Berdasarkan Indeks Pembangunan Manusia Tahun 2022

Isah¹, Mutia Nur Estri², Niken Larasati³

^{1,3}Departemen Matematika, Universitas Jenderal Soedirman

¹isah@mhs.unsoed.ac.id

³niken.larasati@unsoed.ac.id

²Departemen Matematika, Universitas Jenderal Soedirman

Corresponding author email: mutia.estri@unsoed.ac.id

Abstract: One of the problems in Indonesia is the uneven increase in people's welfare, as seen from the differences in the Human Development Index (HDI) levels between provinces. The right strategy through clustering analysis is needed to overcome this inequality. This paper discusses a comparative analysis of the Fuzzy C-Means (FCM) and Fuzzy Possibilistic C-Means (FPCM) methods in grouping provinces in Indonesia based on the Human Development Index (HDI) indicator in 2022. This paper aims to determine the most effective clustering method between FCM and FPCM in this case. Based on the results of the analysis, the best method in this case is the FCM method with a weighting value of 2 and the number of clusters of 7. The evaluation values obtained are Partition Entropy (PE) of 0.3505417, Partition Coefficient (PC) of 0.8170827, and Modified Partition Coefficient (MPC) of 0.7561103. The clustering results obtained using the FCM method are 1 province has a very high HDI category, 4 provinces have a high HDI category, 4 provinces have a fairly high HDI category, 8 provinces have a moderate HDI category, 6 provinces have a fairly low HDI category, 7 provinces have a low HDI category, and 4 provinces have a very low HDI category. The results of the clustering can be used to plan a more effective development equity strategy based on the characteristics of each cluster.

Keywords: Human Development Index, clustering, Fuzzy C-Means, Fuzzy Possibilistic C-Means.

Abstrak: Salah satu permasalahan yang ada Indonesia adalah peningkatan kesejahteraan masyarakat yang belum merata, terlihat dari perbedaan tingkat IPM antar provinsi. Strategi yang tepat melalui analisis klusterisasi diperlukan untuk mengatasi ketidakmerataan tersebut. Makalah ini membahas analisis perbandingan metode Fuzzy C-Means (FCM) dan Fuzzy Possibilistic C-Means (FPCM) dalam pengelompokan provinsi di Indonesia berdasarkan indikator Indeks Pembangunan Manusia (IPM) tahun 2022. Makalah ini bertujuan untuk menentukan metode klusterisasi yang paling efektif antara FCM dan FPCM dalam kasus tersebut. Berdasarkan hasil analisis, diperoleh metode terbaik dalam kasus ini adalah metode FCM dengan nilai pembobotan sebesar 2 dan jumlah kluster sebanyak 7. Nilai evaluasi yang diperoleh yaitu Partition Entropy (PE) sebesar 0,3505417, Partition Coefficient (PC) sebesar 0,8170827, dan Modified Partition Coefficient (MPC) sebesar 0,7561103. Hasil klusterisasi yang diperoleh menggunakan metode FCM adalah 1 provinsi memiliki IPM dengan kategori sangat tinggi, 4 provinsi memiliki IPM dengan kategori tinggi, 4 provinsi memiliki IPM dengan kategori cukup tinggi, 8 provinsi memiliki IPM dengan kategori sedang, 6 provinsi memiliki IPM dengan kategori cukup rendah, 7 provinsi memiliki IPM dengan kategori rendah, dan 4 provinsi memiliki IPM dengan kategori sangat rendah. Hasil pengelompokan tersebut dapat digunakan untuk merencanakan strategi pemerataan pembangunan yang lebih efektif berdasarkan karakteristik masing-masing kluster.

Kata kunci: Indeks Pembangunan Manusia, pengelompokan, Fuzzy C-Means, Fuzzy Possibilistic C-Means.

I. PENDAHULUAN

Kesejahteraan masyarakat merupakan kondisi yang menunjukkan bagaimana keadaan kehidupan masyarakat yang dapat dilihat dari standar kehidupan Masyarakat [1]. Salah satu parameter dari kesejahteraan masyarakat dapat dilihat melalui Indeks Pembangunan Manusia (IPM). Indeks Pembangunan Manusia adalah suatu metode untuk mengukur seberapa besar dampak yang ditimbulkan dari upaya peningkatan kemampuan modal dasar manusia. Indikator dari IPM dengan metode baru meliputi Umur Harapan Hidup (UHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran per Kapita (PPK) [2].



Mengutip data dari Badan Pusat Statistik (BPS) Indonesia tahun 2022, Indonesia dibagi menjadi 34 provinsi dengan tingkat IPM yang berbeda. Permasalahan yang timbul di Indonesia adalah peningkatan IPM yang tidak sejalan dengan pemerataan kesejahteraan masyarakat. Oleh karena itu, langkah yang bisa diambil adalah merencanakan upaya yang dapat memitigasi dampak dari ketidakmerataan IPM tersebut untuk mencapai kesuksesan dalam pembangunan. Analisis mengenai pengelompokan provinsi berdasarkan indikator IPM penting untuk dilakukan agar pemerintah dapat lebih efektif mencapai kesuksesan dalam pemerataan pembangunan. Salah satu metode yang dapat digunakan untuk mengelompokkan adalah analisis kluster. Metode klusterisasi yang dapat digunakan dalam analisis kluster adalah metode *Fuzzy C-Means* (FCM) dan *Fuzzy Possibilistic C-Means* (FPCM). *Fuzzy C-Means* yaitu suatu teknik klusterisasi data yang mana keberadaan tiap-tiap data dalam suatu kluster ditentukan oleh derajat keanggotaan [3], sedangkan FPCM adalah pengembangan dari dua algoritma, yaitu FCM dan *Possibilistic C-Means* (PCM). Metode tersebut sesuai dengan karakter data IPM yang menunjukkan bahwa provinsi memiliki karakteristik IPM yang tidak secara jelas tergolong dalam satu kelompok tertentu.

Beberapa penelitian yang berkaitan dengan metode FCM dan FPCM telah dilakukan. Penelitian tersebut dilakukan oleh Ozdemir dan Kaya [4] yang memperoleh kesimpulan bahwa metode FPCM dan PFCM memberikan hasil yang lebih baik dibandingkan metode FCM dan PCM. Sementara itu, Mulyaningsih [5] telah melakukan penelitian yang memperoleh kesimpulan bahwa metode FCM merupakan metode terbaik. Selain itu, David dkk [6] melakukan penelitian yang memperoleh kesimpulan bahwa pengelompokan terbaik juga dipengaruhi oleh penentuan nilai parameter awal.

Berdasarkan uraian di atas, penulis tertarik untuk mendapatkan hasil pengelompokan terbaik provinsi di Indonesia berdasarkan indikator IPM tahun 2022 dengan membandingkan dua metode klusterisasi *fuzzy*, yaitu FCM dan FPCM. Pemilihan metode ini didasarkan pada karakteristik data penelitian. Data IPM terdiri dari indikator yang bersifat kuantitatif yang menunjukkan distribusi kesejahteraan antara provinsi-provinsi di Indonesia, sehingga metode yang mampu menangani ketidakpastian seperti FCM dan FPCM diperlukan. Evaluasi kinerja dari FCM dan FPCM dilakukan dengan mempertimbangkan tingkat akurasi menggunakan indeks validitas, yaitu *Partition Entropy* (PE), *Partition Coefficient* (PC), dan *Modified Partition Coefficient* (MPC).

II. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah studi kasus. Penelitian ini menerapkan metode FCM dan FPCM untuk klusterisasi IPM provinsi di Indonesia tahun 2022 [7]. Hasil dari kedua metode tersebut kemudian dibandingkan keakuratannya berdasarkan nilai indeks PE, PC, dan MPC. Setelah itu, dilakukan penentuan jumlah kluster yang optimal dan analisis karakteristik hasil klusterisasi tersebut.

Data yang digunakan dalam penelitian ini merupakan data sekunder, yakni data indikator IPM provinsi di Indonesia tahun 2022 yang diperoleh dari *website* BPS Indonesia. Variabel data yang digunakan dalam penelitian ini adalah UHH, HLS, RLS, dan PPK.

Prosedur yang dilakukan untuk menganalisis data indikator IPM provinsi di Indonesia tahun 2022 adalah sebagai berikut:

1. Menyiapkan data yang akan diteliti, yaitu data indikator IPM provinsi di Indonesia tahun 2022;
2. Membuat plot data awal untuk melihat analisis deskriptifnya guna melihat gambaran awal dari data indikator IPM provinsi di Indonesia tahun 2022;
3. Melakukan analisis kluster dengan metode FCM;
4. Melakukan analisis kluster dengan metode FPCM;



5. Mengukur tingkat akurasi metode FCM dan FPCM menggunakan indeks validitas, yaitu *Partition Entropy* (PE), *Partition Coefficient* (PC), dan *Modification Partition Coefficient* (MPC) untuk menentukan metode terbaik dan jumlah kluster optimal;

a. *Partition Entropy*

Nilai ini mengevaluasi keteracakan data dalam kluster. *Partition Entropy* adalah evaluasi yang mengukur tingkat keaburan partisi kluster. *Partition Entropy* mengukur seberapa baik partisi kluster mampu mengelompokkan data ke dalam kelompok yang homogen. Kluster optimal didapatkan apabila nilai yang diperoleh semakin kecil atau mendekati angka 0 [6]. Indeks PE didefinisikan dengan persamaan (1):

$$PE = -\frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik} \ln(\mu_{ik}) \quad (1)$$

dengan

n : jumlah objek penelitian (jumlah data),

c : jumlah kluster,

μ_{ik} : derajat keanggotaan objek data ke- i pada kluster ke- k .

b. *Partition Coefficient*

Bezdek (1981) mengusulkan validitas dengan menghitung koefisien partisi sebagai evaluasi derajat keanggotaan data pada setiap kluster. Nilai PC berkisar dari 0 hingga 1, dengan nilai yang lebih besar menunjukkan partisi yang lebih baik [8]. Indeks PC didefinisikan oleh persamaan (2):

$$PC = \frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^2. \quad (2)$$

c. *Modification Partition Coefficient*

Modified Partition Coefficient adalah indeks validitas hasil perbaikan indeks PC yang memperhitungkan jumlah klusternya [8]. Indeks MPC didefinisikan oleh persamaan (3):

$$MPC = 1 - \frac{c}{c-1} (1 - PC). \quad (3)$$

Suatu kluster dikatakan optimal jika nilai MPC mendekati 1 atau yang paling besar. Semakin kecil nilai MPC, semakin kabur keakuratannya [9].

6. Melakukan analisis karakteristik masing-masing kluster dari hasil klusterisasi provinsi di Indonesia berdasarkan indikator IPM tahun 2022 dengan metode terbaik dengan menggunakan *Sum of Squares Between clusters* (SSB) dan *Sum of Squares Total* (SST);

Rasio antara SSB dan SST memberikan perspektif tentang kemampuan kluster dalam membedakan data sesuai dengan karakteristik yang digunakan dalam klusterisasi. Nilai rasio antara SSB dan SST yang tinggi menunjukkan nilai keragaman antar kluster yang tinggi pula. Hal ini mengindikasikan semakin berkualitas kluster yang terbentuk. Adapun rumus untuk menghitung rasio antara SSB dan SST tercantum pada persamaan (4) hingga persamaan (6) [10]:

$$R = \frac{SSB}{SST} \cdot 100\% \quad (4)$$



$$SSB = \sum_{k=1}^c \sum_{j=1}^m n_k (\bar{x}_{jk} - \bar{x}_j)^2 \quad (5)$$

$$SST = \sum_{k=1}^c \sum_{j=1}^m (x_{ij} - \bar{x}_j)^2 \quad (6)$$

dengan

- R : rasio antara SSB dan SST,
- SSB : *Sum of Squares Between clusters*,
- SST : *Sum of Squares Total*,
- n_k : jumlah data pada kluster ke- k ,
- \bar{x}_{jk} : nilai rata-rata data pada atribut ke- j kluster ke- k ,
- \bar{x}_j : nilai rata-rata data pada atribut ke- j .

7. Mengambil kesimpulan terkait pengelompokan provinsi di Indonesia menggunakan data indikator IPM tahun 2022 berdasarkan analisis yang telah dilakukan.

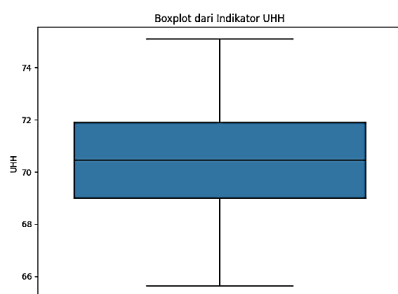
III. HASIL DAN PEMBAHASAN

3.1. Deskripsi Data

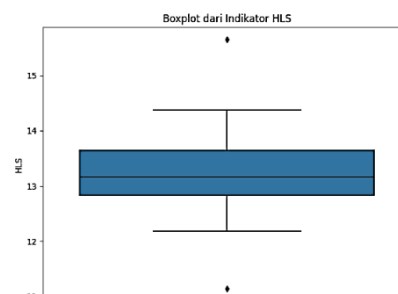
Deskriptif data dilakukan untuk mengetahui gambaran umum dari data yang akan diteliti. Statistika deskriptif dari data indikator IPM provinsi di Indonesia tahun 2022 [7] dapat dilihat pada Tabel 1. Berdasarkan Gambar 1 tidak terdapat data *outlier* pada indikator UHH. Data pada indikator UHH menyebar dengan baik. Sedangkan pada indikator HLS terdapat 2 data *outlier* yang berada di atas dan di bawah *whisker*. Berdasarkan Gambar 2, pada indikator RLS dan PPK juga terdapat data *outlier* yang berjumlah masing-masing 1 yang berada di atas *whisker*.

Tabel 1 Statistika deskriptif indikator IPM provinsi di Indonesia tahun 2022

	UHH	HLS	RLS	PPK
<i>Mean</i>	70,42	13,241	8,839	11.079,559
<i>Median</i>	70,445	13,165	8,835	10.990
<i>Std. deviation</i>	2,452	0,741	0.923	2.246,728
<i>Minimum</i>	65,63	11,14	7,02	7.146
<i>Maximum</i>	75,08	15,65	11,31	18.927

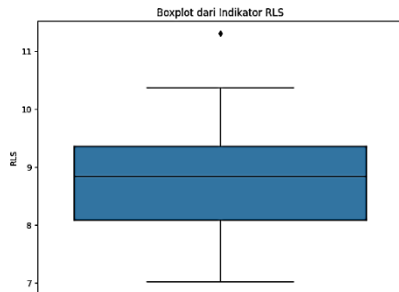


Gambar 1a.

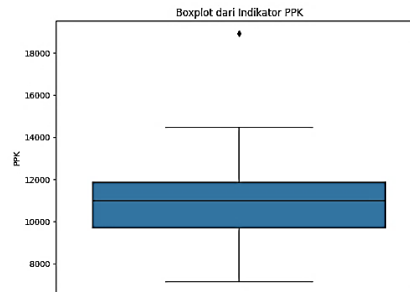


Gambar 1b.

Gambar 1. (1a) boxplot dari indikator UHH (1b) boxplot dari indikator HLS



Gambar 2a.



Gambar 2b.

Gambar 2. (2a) boxplot dari indikator RLS (2b) boxplot dari indikator PPK

3.2. Analisis Kluster dengan Metode Fuzzy C-Means

Langkah-langkah klusterisasi [8] pada iterasi ke-1 menggunakan metode *Fuzzy C-Means* adalah:

- Memasukkan data yang akan dikluster ke dalam matriks berukuran $n \times m$ dengan n adalah jumlah data yaitu 34, m adalah jumlah variabel/atribut setiap data yaitu 4. Elemen matriksnya yaitu x_{ij} , dengan i menunjukkan urutan data dan j menunjukkan urutan atribut dari data;
- Menetapkan nilai parameter awal, yaitu pangkat/pembobot (w) = 2, error/galat terkecil (ϵ) = 10^{-9} , maksimum iterasi (MaxIter) = 1000, jumlah kluster (c) = 2, nilai obyektif awal (P_0) = 0, dan iterasi awal (t) = 1;
- Membangkitkan nilai acak μ_{ik} sebagai elemen matriks partisi awal. Matriks partisi awal U dibentuk secara acak dengan syarat jumlah nilai elemen matriks pada setiap baris harus 1;
- Menghitung matriks pusat kluster V dengan rumus pada persamaan (7):

$$v_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w \cdot x_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (7)$$

dengan v_{kj} adalah elemen matriks pusat kluster.

Pusat kluster V yang terbentuk pada iterasi pertama adalah

$$V = \begin{bmatrix} 70,755 & 13,372 & 8,896 & 11351,944 \\ 70,241 & 13,261 & 8,937 & 10893,460 \end{bmatrix};$$

- Menghitung nilai obyektif pada iterasi ke-1 dengan rumus (8):

$$P_t = \sum_{i=1}^n \sum_{k=1}^c ((d_{ik})^2 (\mu_{ik})^w) \quad (8) \quad d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - v_{kj})^2} \quad (9)$$

dengan d_{ik} adalah jarak antara titik data ke- i dengan pusat kluster ke- k . Nilai obyektif pada iterasi pertama adalah $P_1 = 95.785.384,339$;

- Menghitung perubahan elemen matriks partisi (U) menggunakan rumus (10):



$$\mu_{ik} = \frac{[(d_{ik})^2]^{w-1}}{\sum_{k=1}^c [(d_{ik})^2]^{w-1}}; \quad (10)$$

g. Mengecek kondisi berhenti:

$|P_1 - P_0| = 95.785.384,339 > \varepsilon = 10^{-9}$ dan iterasi = 1 < $MaxIter = 1000$, maka dilanjutkan ke iterasi 2 ($t = 2$) dengan mengulangi dari langkah ke-4.

Dibutuhkan iterasi sebanyak 83 kali menggunakan bantuan *Software R 4.3.2* untuk memperoleh solusi optimal.

3.3. Analisis Kluster dengan Metode Fuzzy Possibilistic C-Means

Langkah-langkah klusterisasi [11] pada iterasi ke-1 menggunakan metode *Fuzzy Possibilistic C-Means* adalah:

- memasukkan data yang akan dikluster ke dalam matriks berukuran $n \times m$;
- menetapkan nilai parameter awal, yaitu pangkat/pembobot (w dan η) = 2, error/galat terkecil (ε)= 10^{-9} , maksimum iterasi =1000, jumlah kluster (c)=2, nilai objektif awal (P_0)=0, iterasi awal (t) = 1;
- memanggil matriks kekhasan relatif U dan matriks pusat kluster V dari hasil akhir algoritma FCM;
- menghitung elemen matriks kekhasan absolut S , dengan persamaan (11), lalu memperbaiki matriks pusat kluster V dengan rumus pada persamaan (12).

$$s_{ik} = \frac{[d_{ik}^2]^{-1}}{\sum_{i=1}^n [d_{ik}^2]^{-1}}, \quad (11) \quad v_{kj} = \frac{\sum_{i=1}^n (\mu_{ik}^w + s_{ik}^\eta) x_{ij}}{\sum_{i=1}^n (\mu_{ik}^w + s_{ik}^\eta)}. \quad (12)$$

Pusat kluster V yang terbentuk pada iterasi pertama adalah:

$$V = \begin{bmatrix} 69,804 & 13,231 & 8,669 & 10116,214 \\ 72,178 & 13,429 & 9,419 & 13842,162 \end{bmatrix};$$

- Memperbaiki matriks kekhasan relatif U dengan rumus pada persamaan (13) lalu memperbaiki matriks kekhasan absolut S mengacu pada (11). Kemudian menghitung nilai objektif pada iterasi ke-1 dengan persamaan (14).

$$\mu_{ik} = \frac{[d_{ik}^2]^{w-1}}{\sum_{k=1}^c [d_{ik}^2]^{w-1}}, \quad (13) \quad P_t = \sum_{i=1}^n \sum_{k=1}^c (\mu_{ik}^w + s_{ik}^\eta) (d_{ik})^2 \quad (14)$$

f. Mengecek kondisi berhenti:

$|P_1 - P_0| = 30.691,937 > \varepsilon = 10^{-9}$ dan iterasi = 1 < $MaxIter = 1000$, maka dilanjutkan ke iterasi 2 ($t = 2$) dengan mengulangi dari langkah ke-5.

Dibutuhkan iterasi sebanyak 65 kali menggunakan bantuan *Software R 4.3.2* untuk memperoleh solusi optimal.

3.4. Analisis Perbandingan Hasil Metode FCM dengan Metode FPCM

Hasil indeks validitas kluster PE pada metode FCM dan FPCM disajikan dalam Tabel 2.



Tabel 2. Hasil indeks validitas kluster PE

Jumlah kluster	FCM		FPCM			
	w = 2	w = 3	w = 2 $\eta = 2$	w = 2 $\eta = 3$	w = 3 $\eta = 2$	w = 3 $\eta = 3$
2	0,29498	0,52324	0,07679	0,31342	0,08430	0,34599
3	0,33433	0,65007	0,07265	0,29894	0,06922	0,27524
4	0,35054	0,86860	0,05697	0,30931	0,12833	0,43233
5	0,41146	0,90782	0,12414	0,48154	0,12154	0,55739
6	0,48817	1,06481	0,18672	0,66424	0,21899	0,74473
7	0,40187	1,13843	0,13022	0,50639	0,15004	0,61147

Berdasarkan analisis hasil indeks validitas kluster PE, dalam kasus metode FCM, hasil analisis menunjukkan bahwa kluster terbaik diperoleh ketika $w = 2$ dibandingkan dengan $w = 3$. Hal ini diindikasikan oleh nilai PE yang lebih rendah. Hasil analisis untuk metode FPCM menunjukkan bahwa $\eta = 2$ cenderung memberikan hasil yang lebih baik daripada $\eta = 3$. Berdasarkan hasil indeks PE, metode FPCM dengan $w = 2$ dan $\eta = 2$ merupakan pilihan terbaik dengan jumlah kluster 2, 3, dan 4 dengan pembobotan $w = 2$ dan $\eta = 2$.

Hasil perhitungan indeks PC pada metode FCM dan FPCM disajikan dalam Tabel 3.

Tabel 3 Hasil indeks validitas kluster PC

Jumlah kluster	FCM		FPCM			
	w = 2	w = 3	w = 2 $\eta = 2$	w = 2 $\eta = 3$	w = 3 $\eta = 2$	w = 3 $\eta = 3$
2	0,82369	0,65056	0,03146	0,05882	0,02910	0,05882
3	0,81863	0,63244	0,05440	0,08824	0,05637	0,08824
4	0,81708	0,54492	0,09727	0,11765	0,06902	0,11765
5	0,80229	0,55707	0,09289	0,14706	0,09373	0,14706
6	0,76258	0,50300	0,09340	0,17647	0,08352	0,17647
7	0,81458	0,49195	0,14473	0,20588	0,13663	0,49195

Berdasarkan analisis hasil indeks validitas kluster PC, diperoleh bahwa dengan nilai $w = 2$ maupun $w = 3$, hasil kluster terbaik untuk semua jumlah kluster adalah menggunakan metode FCM. Hal ini terindikasi dari nilai PC yang tinggi. Ketika menggunakan metode FCM, hasil analisis menunjukkan bahwa kluster terbaik diperoleh ketika $w = 2$ dibandingkan dengan $w = 3$. Selanjutnya, untuk metode FPCM, hasil analisis menunjukkan bahwa nilai $\eta = 3$ menghasilkan kualitas kluster yang lebih baik dibandingkan $\eta = 2$. Nilai PC tertinggi didapatkan ketika menggunakan metode FCM dengan $w = 2$ dan jumlah kluster 2. Hasil indeks validitas kluster MPC pada metode FCM dan FPCM disajikan dalam Tabel 4.

Tabel 4. Hasil indeks validitas kluster MPC

Jumlah kluster	FCM		FPCM			
	w = 2	w = 3	w = 2 $\eta = 2$	w = 2 $\eta = 3$	w = 3 $\eta = 2$	w = 3 $\eta = 3$
2	0,64738	0,30112	-0,93708	-0,88235	-0,94180	-0,88235
3	0,72794	0,44866	-0,41841	-0,36765	-0,41545	-0,36765
4	0,75611	0,39322	-0,20364	-0,17647	-0,24130	-0,17647
5	0,75287	0,44634	-0,13389	-0,06618	-0,13283	-0,06618
6	0,71510	0,40360	-0,08792	0,01176	-0,09978	0,01176
7	0,78368	0,40727	0,00218	0,07353	-0,00727	0,07353

Berdasarkan analisis hasil indeks validitas kluster MPC, diperoleh bahwa metode FCM dengan $w = 2$ memberikan hasil terbaik dengan jumlah kluster 3, 4, 5, 6, dan 7. Secara khusus, pada $w = 2$, FCM mencapai nilai MPC tertinggi saat jumlah klasternya 7. Sedangkan pada metode FPCM, nilai



pembobotan $\eta = 3$ memberikan hasil yang lebih baik dibandingkan $\eta = 2$, terutama ketika jumlah kluster lebih banyak. Kesimpulannya, FCM memberikan nilai MPC yang lebih baik daripada FPCM. Hasil MPC tertinggi diperoleh menggunakan metode FCM dengan $w = 2$ dan jumlah kluster 7.

3.5. Penentuan Metode dan Jumlah Kluster Optimal

Selain mempertimbangkan indeks validitas kluster, dalam upaya menentukan jumlah kluster optimal dapat digunakan rasio antara SSB dan SST.

Tabel 5 Hasil rasio SSB dan SST

Jumlah kluster	FCM		FPCM			
	$w = 2$	$w = 3$	$w = 2$ $\eta = 2$	$w = 2$ $\eta = 3$	$w = 3$ $\eta = 2$	$w = 3$ $\eta = 3$
2	55,31%	48,78%	51,37%	51,38%	46,83%	46,83%
3	77,37%	75,00%	78,29%	78,21%	47,76%	47,72%
4	91,13%	80,79%	86,20%	86,34%	56,56%	55,37%
5	85,89%	84,49%	79,42%	79,22%	88,46%	57,70%
6	86,24%	84,99%	85,69%	85,65%	97,47%	97,47%
7	98,23%	87,47%	90,29%	90,32%	97,97%	97,97%

Berdasarkan analisis rasio SSB dan SST, dapat disimpulkan bahwa pada jumlah kluster 7, semua metode dan pembobotan memberikan hasil yang sangat bagus, FCM dengan $w = 2$ mencapai nilai tertinggi. Klasterisasi dapat dengan baik direpresentasikan dalam 7 kluster. Metode FCM dengan $w = 2$ tetap menjadi pilihan unggul dalam kasus ini. Rangkuman nilai semua indeks validitas kluster serta rasio SSB dan SST menggunakan metode FCM dengan $w = 2$ untuk memudahkan menentukan jumlah kluster optimal termuat dalam Tabel 6. Selanjutnya, akan dieksplorasi lebih lanjut hasil pengelompokan dengan jumlah kluster 7.

Tabel 6. Penentuan jumlah kluster optimal

Jumlah Kluster	FCM dengan $w = 2$			
	PE	PC	MPC	SSB/SST
2	0,29498	0,82369	0,64738	55,31%
3	0,33433	0,81863	0,72794	77,37%
4	0,35054	0,81708	0,75611	91,13%
5	0,41146	0,80229	0,75287	85,89%
6	0,48817	0,76258	0,71510	86,24%
7	0,40187	0,81458	0,78368	98,23%

3.6. Karakteristik Hasil Kluster

Dibutuhkan iterasi sebanyak 59 kali untuk memperoleh solusi optimal. Pada iterasi terakhir, pusat kluster yang dihasilkan adalah:

$$v = \begin{bmatrix} 72,61627 & 13,93648 & 9,741095 & 14.247,443 \\ 67,14009 & 12,88465 & 7,913000 & 7.899,743 \\ 69,64359 & 13,24373 & 8,761832 & 9.457,757 \\ 73,31955 & 13,07991 & 11,309473 & 18.925,879 \\ 71,80221 & 13,06596 & 8,802257 & 11.271,931 \\ 71,35216 & 13,19230 & 8,959079 & 12.435,626 \\ 69,33987 & 13,38114 & 8,512053 & 10.726,226 \end{bmatrix}$$

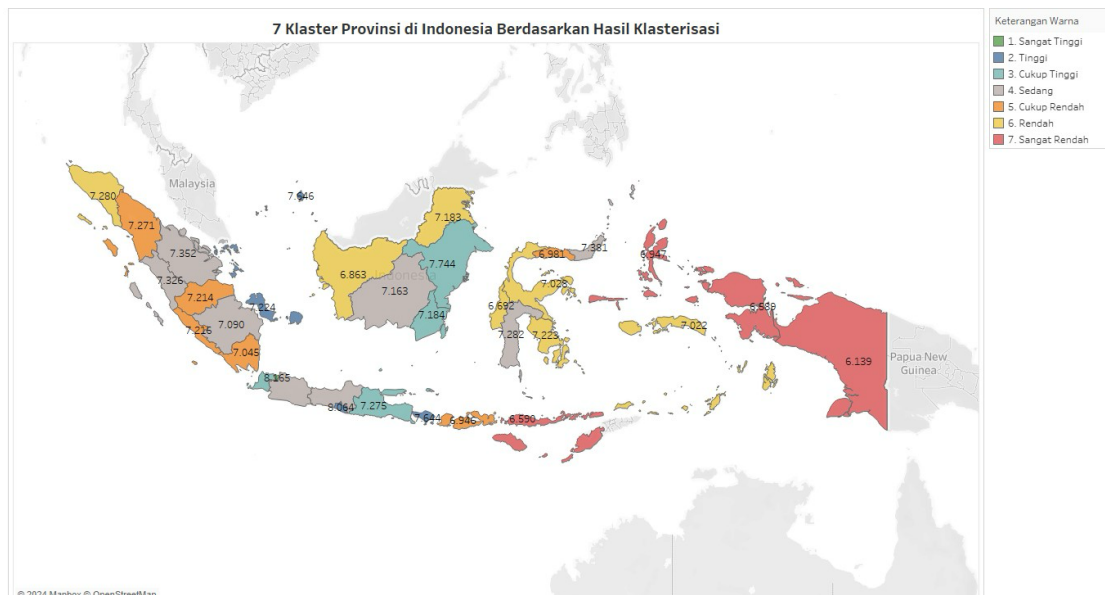
Hasil pengelompokan disajikan pada Tabel 7 dan karakteristiknya, disajikan dalam Tabel 8.

Tabel 7. Hasil pengelompokan provinsi di Indonesia dengan jumlah kluster 7

Kluster	Provinsi
1	Kep. Bangka Belitung, Kep. Riau, DI Yogyakarta, Bali
2	Nusa Tenggara Timur, Maluku Utara, Papua Barat, Papua
3	Aceh, Kalimantan Barat, Kalimantan Utara, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Barat, Maluku
4	DKI Jakarta
5	Sumatera Barat, Riau, Sumatera Selatan, Jawa Barat, Jawa Tengah, Kalimantan Tengah, Sulawesi Utara, Sulawesi Selatan
6	Jawa Timur, Banten, Kalimantan Selatan, Kalimantan Timur
7	Sumatera Utara, Jambi, Bengkulu, Lampung, Nusa Tenggara Barat, Gorontalo

Tabel 8. Hasil pengelompokan dan karakteristiknya

Kluster	Rata-rata UHH	Rata-rata HLS	Rata-rata RLS	Rata-rata PPK	Rata-rata IPM
1 (tinggi)	72,29	13,575	9,405	14.063	76,45
2 (sangat rendah)	67,238	12,823	7,95	7,881	65,66
3 (rendah)	69,464	13,424	8,959	9,472	70,42
4 (sangat tinggi)	73,32	13,08	11,31	18,927	81,65
5 (sedang)	71,669	13,075	8,805	11,265	72,73
6 (cukup tinggi)	71,47	13,27	8,885	12.330	73,84
7 (cukup rendah)	69,562	13,31	8,518	10.711	71,12



Gambar 3. Peta hasil klusterisasi provinsi di Indonesia dengan jumlah kluster 7

IV. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan dalam penelitian yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut:

1. Hasil perbandingan nilai indeks validitas kluster PE, PC, dan MPC pada metode FCM dan FPCM diperoleh bahwa metode FCM dengan parameter pembobotan 2 dan jumlah kluster 7 merupakan metode yang terbaik dalam studi kasus ini. Nilai evaluasi yang diperoleh, yaitu PE sebesar 0,3505417, PC sebesar 0,8170827, dan MPC sebesar 0,7561103.



2. Hasil klasterisasi dan karakteristik dari masing-masing klaster berdasarkan hasil analisis klaster yang diperoleh dari metode FCM menggunakan data indikator IPM provinsi di Indonesia disajikan dalam Tabel 7 dan Tabel 8.

REFERENSI

1. Badrudin, R. *Ekonomika Otonomi Daerah*, 2nd ed.; UPP STIM YKPN: Yogyakarta, Indonesia, 2017.
2. Badan Pusat Statistik, *Indeks Pembangunan Manusia 2020*; Badan Pusat Statistik Indonesia: Jakarta, Indonesia, 2021.
3. Kusumadewi, S.; Purnomo, H. *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*, 2nd ed.; Graha Ilmu: Yogyakarta, Indonesia, 2010.
4. Ozdemir, O.; Kaya, A. Comparison of FCM, PCM, FPCM and PFCM Algorithms in Clustering Methods. *Afyon Kocatepe University Journal of Sciences and Engineering* **2019**, vol. 19, pp. 92–102.
5. Mulyaningsih, W. S. Implementasi Fuzzy C-Means dan Fuzzy Possibilistic C-Means untuk Pengelompokan Kabupaten/Kota di Provinsi Banten. Skripsi, Universitas Islam Indonesia, Yogyakarta, 2021.
6. David; Lauro, M. D.; Herwindiati, D. E.; Sistem Prediksi Customer Loyalty dengan Metode RFM dan Fuzzy C-Means. *Computatio: Journal of Computer Science and Information Systems* **2020**, vol. 4, pp. 33–44.
7. Badan Pusat Statistik, *Indeks Pembangunan Manusia 2022*; Badan Pusat Statistik Indonesia: Jakarta, Indonesia, 2023.
8. Xie, N.; Hu, L.; Luktarhan, N.; Zhao, K. A Classification of Cluster Validity Indexes Based on Membership Degree and Applications. *LNCS* **2011**, vol. 1, pp. 43–50.
9. Amirah, M. M. A.; Widodo, A. W.; Dewi, C. Pengelompokan Lagu Berdasarkan Emosi Menggunakan Algoritma Fuzzy C-Means. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* **2017**, vol. 1, pp. 1526–1534.
10. Howell, D. C. *Fundamental Statistics for the Behavioral Sciences*, Ninth Edition.; Cengage Learning: Boston, Amerika Serikat, 2017.
11. Kusumadewi, S.; Hartati, S.; Harjoko, A.; Wardoyo, R. *Fuzzy Multi-Attribute Decision Making (Fuzzy MADM)*; Graha Ilmu: Yogyakarta, Indonesia, 2006.