

Klaster Daerah Kesejahteraan pada Masa Pandemi *Covid-19* di Jawa Timur dengan Metode *Fuzzy C-Means Clustering*

YUMNA SALSABILA FIRDAUS¹, RISANG NARENDRA², ARDHI SANWIDI³,
RACHMADANIA AKBARITA⁴

^{1,2,3,4}Program Studi Matematika, FIE, Universitas Nahdlatul Ulama Blitar
e-mail: yumnafirdaus@outlook.co.id

ABSTRAK

Pembukaan UUD 1945 alinea ke-4 menyebutkan bahwa salah satu tujuan utama negara Indonesia adalah kesejahteraan. Namun, sejak pandemi *Corona Virus Disease* atau COVID-19 yang melanda Indonesia memberikan dampak pada peningkatan kemiskinan, termasuk di Provinsi Jawa Timur. Akibat meningkatnya kemiskinan, berbagai bidang di dalam indikator kesejahteraan di Provinsi Jawa Timur menurun. Upaya meningkatkan kesejahteraan, dibutuhkan data yang tepat sasaran dan akurat, yaitu dengan melakukan penelitian analisis kluster salah satunya dengan menggunakan metode *Fuzzy C-Means Clustering*. Dengan tersebut, akan dicari jumlah kluster untuk data indikator kesejahteraan rakyat di Jawa Timur. Namun, dalam menjalankan algoritma metode *Fuzzy C-Means Clustering*, peneliti harus menentukan jumlah kluster terlebih dahulu, apabila peneliti asal dalam menentukannya maka akan berdampak pada hasil kluster yang tidak optimal, serta mengakibatkan pengambilan keputusan dan kesimpulan yang salah. Hal tersebut menjadi kelemahan dari metode *Fuzzy C-Means Clustering*. Untuk mengatasinya kelemahan dari metode *Fuzzy C-Means*, jumlah kluster yang optimal ditentukan dengan *Cluster Validity Index*. Agar mendapatkan jumlah kluster yang optimal, algoritma pada metode *Fuzzy C-Means Clustering* harus dijalankan berkali-kali untuk setiap jumlah kluster yang telah ditentukan oleh peneliti, kemudian hasil atau outputnya diterapkan ke *Cluster Validity Index*. *Index* dalam penelitian ini adalah *Partition Entropy Index*, *Partition Coefficient Index*, *Xie Beni Index*, *Modified Partition Index*, dan *Fukuyama Sugeno Index*. Berdasarkan hasil analisis data, seluruh *Cluster Validity Index* menunjukkan jumlah kluster sebanyak 2 kluster sebagai jumlah kluster yang paling optimal untuk data indikator kesejahteraan rakyat di Jawa Timur pada tahun 2020.

Kata Kunci: Indikator Kesejahteraan Rakyat, Fuzzy C-Means, Cluster Validity Index.

ABSTRACT

In the opening of the 4th paragraph of the 1945 Constitution, Welfare is one of the main goals of the government in Indonesia. However, the Corona Virus Disease or COVID-19 pandemic in Indonesia since the beginning of 2020 has had an impact on increasing poverty, including in East Java. As a result of the increase in poverty, various areas in the welfare indicators in East Java have decreased in 2020. Efforts to improve welfare require targeted and accurate data, namely by conducting cluster analysis research, one of which is using the *Fuzzy C-Means Clustering* method. However, in carrying out the *Fuzzy C-Means* method, researchers must be determine the number of clusters, so if the researcher is inconsequent while determining, it will have an impact on non-optimal cluster results, and result in wrong decision making and conclusions by researchers, That is a result of the weakness in *Fuzzy C-Means Clustering* method. The optimal number of clusters is determined by the Cluster Validity Index. To get the optimal number of clusters, the algorithm in the *Fuzzy C-Means* method must be run many times for each number of clusters determined by the researcher, then the results or output are applied to the Cluster Validity Index. The Cluster Validity Index used in this research is *Pantition Entropy Index*, *Partition Coefficient Index*, *Xie Beni Index*, *Modified Partition Coefficient Index* and *Fukuyama-Sugeno Index*. Based on the results of data analysis, all Cluster Validity Indexes show 2 clusters as the most optimal number of clusters for data on indicators of people's welfare in East Java in 2020.

Keyword: Indicators Of People's Welfare, Fuzzy C-Means, Cluster Validity Index.

1. PENDAHULUAN

Pada pembukaan UUD 1945 alinea ke-4 menjelaskan bahwa kesejahteraan merupakan salah satu tujuan utama pemerintah Indonesia. Akan tetapi, terjadinya Pandemi Covid-19 di Indonesia pada awal tahun 2020 memiliki dampak yang terlihat yaitu pada bidang kemiskinan. Pada berbagai penelitian yang telah diamati oleh Badan Pusat Statistik (BPS) Provinsi Jawa Timur, kemiskinan berpengaruh atau berdampak pada bidang di dalam Indikator Kesejahteraan Jawa Timur (BPS Provinsi Jawa Timur, 2021). Menurut laporan Bank Indonesia, dampak akibat pandemi COVID-19 pada bidang ketenagakerjaan dan kemiskinan sangat terlihat jelas, pada bidang ketenagakerjaan terjadi peningkatan persentase Tingkat Pengangguran Terbuka (TPT) dari 3,92% di tahun 2019 menjadi 5,84% di tahun 2020. Selain itu peningkatan jumlah penduduk miskin di Jawa Timur per September 2020 juga mengalami peningkatan menjadi 4.585,97 ribu jiwa dari 4.056,00 pada September 2019, dan tingkat kemiskinan pada Provinsi Jawa Timur sebesar 11,46% . Angka tersebut lebih tinggi dari tingkat kemiskinan Nasional yang sebesar 10,19% (Bank Indonesia, 2021). Perbedaan lokasi geografis, sumber daya manusia dan alam di Provinsi Jawa Timur, mengakibatkan tingkat pencapaian dari kebijakan dan pembangunan pemerintah juga berbeda antara kota dan kabupatennya. Untuk membantu pemerintah Provinsi Jawa Timur agar mencapai keberhasilan dalam meningkatkan kesejahteraan, dibutuhkan data yang akurat dan tepat sasaran (Mashfuufah and Istiawan, 2018), yaitu dengan melakukan penelitian analisis kluster.

Analisis kluster merupakan bagian dari Statistik Multivariat yang dapat digunakan untuk mengelompokkan data berdasarkan kesamaan dan karakteristik yang dimiliki (Karti and Irhamah, 2013). Metode *Fuzzy C-Means Clustering* dipilih dalam penelitian ini karena dalam beberapa penelitian, metode ini memiliki keunggulan dalam mengelompokkan data dari pada penggunaan metode K-Means (Febrianti et al., 2016; Rizal & Hakim, 2015). Selain lebih unggul dari metode K-Means, metode analisis kluster *Fuzzy C-Means* memiliki kelebihan, yaitu stabil terhadap data outlier, mampu mendeteksi kluster dengan baik, dan memiliki ketepatan dalam menentukan pusat kluster (Mashfuufah and Istiawan, 2018). Namun, dalam menjalankan algoritma pada metode *Fuzzy C-Means Clustering*, peneliti harus menentukan jumlah kluster terlebih dahulu, apabila peneliti asal dalam menentukannya maka hasil kluster yang tidak mencapai optimal, peneliti akan salah dalam menentukan keputusan dan kesimpulan penelitian tidak tepat (Wang and Zhang, 2007; Zarandi, Sotudian and Castillo, 2021). Hal tersebut menjadi kelemahan dari metode *Fuzzy C-Means* (Zarandi, Sotudian and Castillo, 2021).

Dalam melakukan analisis kluster, penentuan jumlah kluster berperan penting untuk menghasilkan dugaan dan kesimpulan yang baik (Wu et al., 2015). Untuk mendapatkan jumlah kluster yang optimal, algoritma pada metode *Fuzzy C-Means* harus dijalankan berkali-kali untuk setiap jumlah kluster yang telah ditentukan oleh peneliti, kemudian hasil atau outputnya diterapkan ke Indeks Validitas Kluster atau *Cluster Validity Indices* (CVI) untuk menentukan jumlah kluster yang optimal, dan jumlah kluster yang optimal yang akan diperoleh dari CVI akan menjadi kesimpulan dari penelitian (Zarandi, Sotudian and Castillo, 2021). Penelitian ini bertujuan untuk mendapatkan jumlah kluster yang optimal menggunakan *Cluster Validity Index*.

2. METODE PENELITIAN

Data Penelitian

Data yang digunakan pada Penelitian ini adalah data sekunder yaitu Data Indikator pada tahun 2020 Kesejahteraan Rakyat Provinsi Jawa Timur dan bersumber dari Badan Pusat Statistik (BPS), dengan menggunakan 6 indikator sebagai variabel penelitian yaitu persentase penduduk miskin (X1), gini rasio (X2), indeks kedalaman kemiskinan (X3), indeks keparahan kemiskinan (X4), tingkat pengangguran terbuka (X5), dan tingkat partisipasi angkatan kerja (X6).

Metode Penelitian

Metode *Fuzzy C-Means Clustering*

Algoritma metode *Fuzzy C-Means Clustering* yang diberikan oleh Bezdek adalah menggunakan Fungsi Objektif (*Objective Function*, J_{FCM}) seperti pada Formula 1 berikut ini (Kusumadewi and Purnomo, 2013):

$$J_{FCM}(X; U, V) = \sum_{k=1}^n \sum_{k=1}^c \left(u_{ij}^m \cdot \left(\sum_{j=1}^p (x_j - v_i)^2 \right) \right) \quad \dots(1)$$

Pada formula (1) m adalah *weighting exponent* atau eksponen pembobot (yang merupakan bilangan real) dengan nilai $m > 1$. Di beberapa penelitian sebelumnya, tidak ada ketentuan pasti untuk nilai m , namun nilai m yang sering diterapkan dalam penelitian adalah $m = 2$ seperti pada penelitian yang dilakukan oleh Dunn (Bezdek, 1981; Cebeci, 2020). Selain m atau eksponen pembobot, terdapat U dan V . U merupakan matriks keanggotaan berdimensi $n \times c$; sedangkan V merupakan matriks pusat kluster. Berikut urutan algoritma metode *Fuzzy C-Means Clustering* :

1) Tentukan:

- Matriks X berukuran $n \times p$, dengan n = jumlah data yang akan dikluster; dan p = parameter/jumlah variabel.
- Jumlah kluster yang akan dibentuk = $c (\geq 2)$.
- Eksponen pembobot, $m > 1$.
- Maksimum iterasi.
- Kriteria berhenti iterasi atau error yang diharapkan = ξ (nilai *error* positif yang amat kecil), dalam berbagai penelitian menggunakan metode *Fuzzy C-Means Clustering* nilai ξ yang digunakan adalah 10^{-5} (Bezdek, 1981; Azizah, Yuniarti and Goejantoro, 2018; Mashfuufah and Istiawan, 2018; Cebeci, 2020)
- Fungsi objektif awal (J_0) = 0
- Iterasi awal, $t = 1$

2) Bentuk matriks keanggotaan awal U_0 dengan bilangan acak untuk u_{ij} , $i = 1,2,3,\dots,n$; $j = 1,2,3,\dots,c$, syarat $u_{ij} \in [0,1], \forall i,j; \sum_{i=1}^c u_{ij} = 1$.

$$U_0 = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1c} \\ u_{21} & u_{22} & \dots & u_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nc} \end{bmatrix}$$

3) Cari nilai pusat kluster (V) untuk setiap kluster dengan $i = 1,2,3,\dots,c$ dan $j = 1,2,3,\dots,p$ (Kusumadewi and Purnomo, 2013)

$$v_i = \frac{\sum_{j=1}^n ((u_{ij})^m \cdot x_j)}{\sum_{j=1}^n (u_{ij})^m} \quad \dots(2)$$

$$V = \begin{bmatrix} v_{11} & \dots & v_{1p} \\ \vdots & \ddots & \vdots \\ v_{c1} & \dots & v_{cp} \end{bmatrix}$$

4) Derajat keanggotaan seluruh data masing-masing kluster diperbaiki (menggunakan formula (3)), dengan catatan menghitung fungsi objektif (formula (1)) terlebih dahulu, kemudian menggunakan persamaan sebagai berikut (Kusumadewi and Purnomo, 2013):

$$u_{ij} = \frac{\left[\sum_{j=1}^p (x_j - v_i)^2 \right]^{-\frac{1}{m-1}}}{\sum_{i=1}^c \left[\sum_{j=1}^p (x_j - v_i)^2 \right]^{-\frac{1}{m-1}}} \quad \dots(3)$$

5) Tentukan kriteria berhenti/periksa kondisi berhenti atau konvergen dari perulangan, yaitu dengan melihat perubahan nilai fungsi objektif iterasi saat ini dengan iterasi sebelumnya, atau iterasi kurang dari maksimum iterasi, ditunjukkan sebagai berikut $|J_{FCM_t} - J_{FCM_{t-1}}| < \xi$ atau $t >$ Iterasi Maksimal. Jika kriteria pemberhentian belum terpenuhi, maka iterasi bertambah $t = t + 1$ kemudian ulangi langkah ke-3

Cluster Validity Index

1) Partition Entropy Index (PEI)

Nilai b pada indeks ini adalah bilangan eksponen (e). Penentuan jumlah kluster optimal pada indeks *Partition Entropy* diambil pada nilai yang paling minimal dan bergantung pada matriks derajat keanggotaan U pada iterasi akhir (Bezdek, 1973; Ferraro and Giordani, 2015), dan persamaan PEI dinyatakan sebagai berikut (Liu *et al.*, 2019):

$$PEI(U) = \left(-\frac{1}{n}\right) \left(\sum_{i=1}^c \sum_{j=1}^n u_{ij} \log_b u_{ij}\right) \quad \dots(4)$$

2) Partition Coefficient Index (PCI)

Indeks *Partition Coefficient* juga bergantung pada matriks derajat keanggotaan U pada iterasi akhir dengan mengambil nilai yang paling maksimal untuk menentukan banyak kluster yang optimal (Bezdek, 1973; Cebeci, 2020), indeks *Partition Coefficient* memiliki nilai pada interval $\frac{1}{c}$ sampai 1 (Haqiqi and Kurniawan, 2015), persamaan PCI dinyatakan sebagai berikut (Cebeci, 2020):

$$PCI(U) = \frac{1}{n} \left(\sum_{i=1}^c \sum_{j=1}^n u_{ij}^m\right) \quad \dots(5)$$

Xie Beni Index (XBI)

Indeks ini bergantung pada derajat keanggotaan U , jarak titik data (X) ke pusat kluster (V), serta jarak antara pusat kelompok. Jumlah kluster yang optimal ditunjukkan oleh nilai indeks yang minimal. Berikut persamaan dari XBI (Liu *et al.*, 2019):

$$XBI(X; V; U) = \frac{J_{FCM}(X; U, V)}{n \left(\min_{i, j = 1, \dots, c; i \neq j} \{\|v_i - v_j\|^2\}\right)} \quad \dots(6)$$

4) Modified Partition Coefficient Index (MPCI)

Indeks *Modified Partition Coefficient* bertujuan untuk mengatasi kekurangan dari indeks *partition coefficient*, dimana nilai indeks *partition coefficient* sering berubah seiring dengan berubahnya nilai c . Jumlah kluster yang optimal ditunjukkan oleh nilai indeks yang maksimal (Dave, 1996; Haqiqi and Kurniawan, 2015; Cebeci, 2020). Persamaan indeks ini dituliskan sebagai berikut:

$$MPCI(U) = \frac{(c \cdot PCI - 1)}{c - 1} \quad \dots(7)$$

5) Fukuyama-Sugeno Index (FSI)

Nilai indeks Fukuyama Sugeno yang digunakan untuk menentukan pusat kluster adalah nilai yang paling minimal. Berikut persamaan dari indeks Fukuyama Sugeno (Liu *et al.*, 2019):

$$FSI(X; V; U) = J_{FCM}(X; U, V) - \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|v_i - \bar{v}\|^2 \quad \dots(8)$$

3. HASIL DAN PEMBAHASAN

Metode Fuzzy C-Means Clustering

Analisis data yang pertama adalah meng-kluster data penelitian menggunakan metode *Fuzzy C-Means Clustering*.

1. Langkah pertama pada analisis kluster menggunakan metode *Fuzzy C-Means* adalah:

- Matriks X yang merupakan data penelitian berukuran $n \times p$, dengan $n = 38$; dan $p = 6$.
- Penentuan jumlah kluster uji dalam penelitian $c = 2 - 6$.
- Eksponen pembobot, $m = 2$ (untuk semua jumlah kluster).
- Penetapan iterasi maksimal, **Max.Iter. = 100** (untuk semua jumlah kluster).

- Kriteria pemberhentian iterasi atau *error* yang diharapkan, $\xi = 10^{-5}$ (untuk semua jumlah kluster).
 - Fungsi objektif awal (J_0) = 0 (untuk semua jumlah kluster).
 - Iterasi awal, $t = 1$ (untuk semua jumlah kluster).
2. Kedua, bentuk matriks keanggotaan awal atau U_0 dengan bilangan acak untuk u_{ij} , $i = 1,2,3,\dots, n$; $j = 1,2,3,\dots, c$, syarat $u_{ij} \in [0,1], \forall i,j$; $\sum_{i=1}^c u_{ij} = 1$.
3. Langkah ketiga, menghitung nilai pusat kluster pada masing-masing kluster menggunakan formula (2).

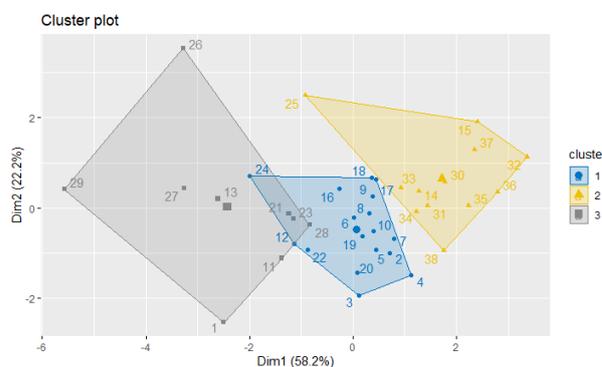
$$v_1 = \frac{\sum_{j=1}^{38} ((u_{ij})^2 \cdot x_j)}{\sum_{j=1}^{38} (u_{ij})^2}$$

4. Pada langkah keempat, derajat keanggotaan seluruh data masing-masing kluster diperbaiki untuk iterasi berikutnya menggunakan formula (3), dengan catatan menghitung fungsi objektif terlebih dahulu (formula (1)).

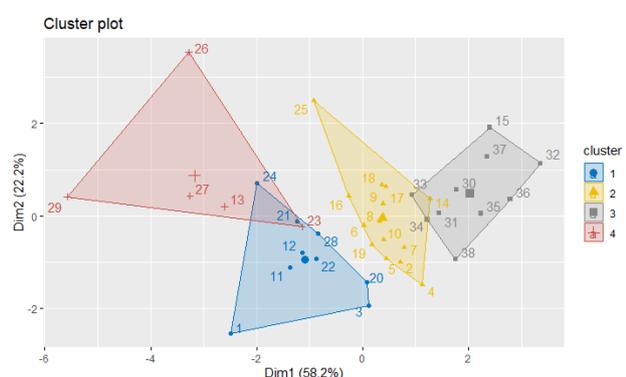
Untuk mempermudah dalam menjalankan algoritma *Fuzzy C-Means Clustering*, terlebih lagi pada algoritma ini melakukan banyak iterasi atau perulangan, maka penelitian ini menggunakan software pemrograman RStudio untuk memudahkan dan meminimalisir kesalahan dalam proses pengolahan data. Dari pengolahan data menggunakan RStudio diperoleh hasil plot penempatan kluster wilayah kota dan kabupaten (setelah iterasi terakhir):



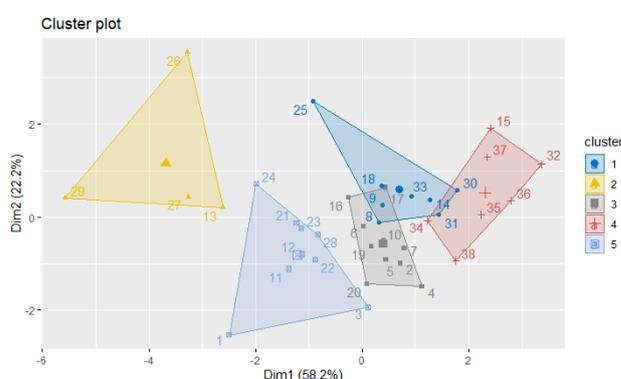
Gambar 1. Plot 2 kluster



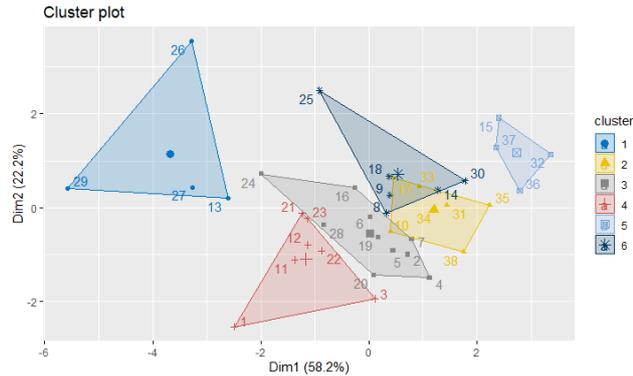
Gambar 2. Plot 3 kluster



Gambar 3. Plot 4 kluster



Gambar 4. Plot 5 kluster



Gambar 5. Plot 6 kluster

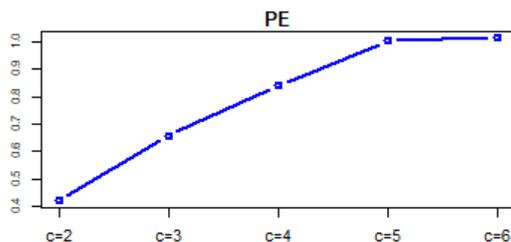
Cluster Validity Index

Cluster Validity Index pada metode *Fuzzy C-Means Clustering* bertujuan untuk menentukan jumlah kluster yang optimal pada data yang diteliti. Pada penelitian ini jumlah kluster yang diuji sebanyak 2-6 kluster dengan 5 indeks validasi kluster (*Partition Entropy Index, Partition Coefficient Index, Xie Beni Index, Modified Partition Index, dan Fukuyama Sugeno Index*). Berikut hasil pengujian menggunakan *software* RStudio:

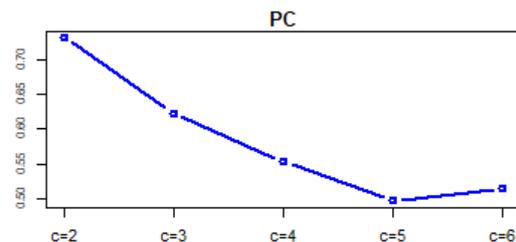
Tabel 1. Hasil olah data RStudio untuk nilai *Cluster Validity Index* 2 - 6 kluster

Jumlah Kluster	Cluster Validity Index				
	PEI	PCI	XBI	MPCI	FS
2	0.42150448	0.7324392	0.2207773	0.46487484	-105546.3
3	0.6571734	0.623253	0.2826384	0.4348794	-90145.28
4	0.8395703	0.553346	0.258673	0.4044613	-80423.01
5	0.9544183	0.5215835	0.3691068	0.4019794	-75698.21
6	1.024691	0.5145755	0.2481865	0.4174906	-74598.09

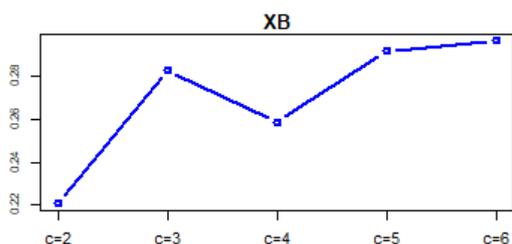
Syarat masing-masing indeks yaitu *Partition Entropy Index* dengan nilai yang paling minimal, *Partition Coefficient Index* dengan nilai yang paling maksimal, *Xie-Beni Index* dengan nilai yang paling minimal, *Modified Partition Coefficient Index* dengan nilai yang paling maksimal, dan *Fukuyama-Sugeno Index* dengan nilai yang paling minimal. Berdasarkan persyaratan masing-masing indeks, jumlah kluster yang optimal pada Data Indikator Kesejahteraan Rakyat pada tahun 2020 adalah sebanyak 2 kluster (ditunjukkan pada Tabel yang memiliki blok berwarna kuning). Berikut perbandingan nilai masing-masing *cluster validity index* dalam bentuk diagram garis.



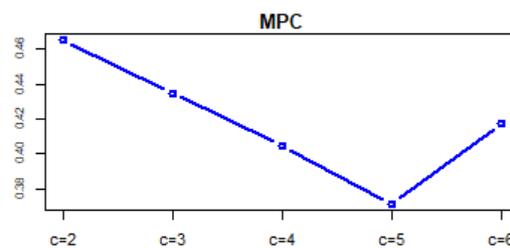
Gambar 6. Diagram nilai Partition Entropy Index c =2-6



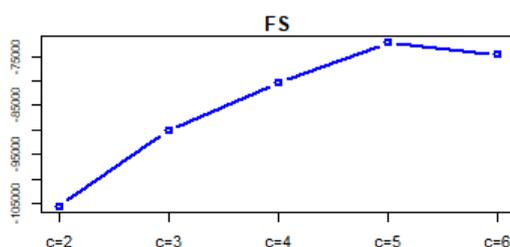
Gambar 7. Diagram nilai Partition Coefficient Index c =2-6



Gambar 8. Diagram nilai Xie-Beni Index c =2-6

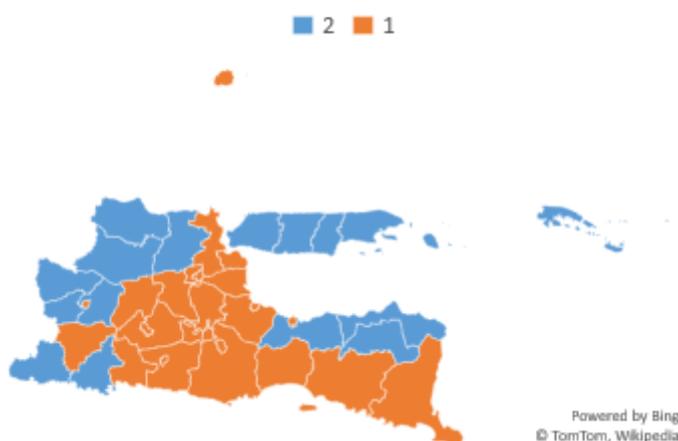


Gambar 9. Diagram nilai Modified Partition Coefficient Index c =2-6



Gambar 10. Diagram nilai Fukuyama-Sugeno Index c =2-6

Jadi, dapat disimpulkan bahwa 2 klaster adalah jumlah klaster yang optimal untuk data Indikator Kesejahteraan Rakyat Jawa Timur pada tahun 2020. Berdasarkan hasil akhir, jika digambarkan dalam peta wilayah, maka kabupaten dan kota anggota klaster 1 dan 2 akan seperti pada peta berikut:



Gambar 11. Peta klaster wilayah Jawa Timur berdasarkan Indikator Kesejahteraan Rakyat Jawa Timur Tahun 2020

Dengan jumlah sebanyak 2 klaster, klaster 1 ditunjukkan oleh warna kuning, sedangkan untuk klaster 2 ditunjukkan oleh warna biru. Anggota klaster pertama terdiri dari 14 kabupaten dan 9 kota. Ciri-ciri kalster pertama sebagai berikut: persentase penduduk miskin 8,24%; gini rasio di angka 0,34; indeks kedalaman kemiskinan sebesar 1,20; indeks keparahan kemiskinan sebesar 0,28; persentase TPT (Tingkat Pengangguran Terbuka) sebesar 6,43%; dan persentase TPAK (Tingkat Partisipasi Angkatan Kerja) sebesar 69,24%. Sedangkan ciri-ciri dari klaster kedua adalah rata-rata: persentase penduduk miskin sebesar 15,07%, gini rasio diangka 0,32; indeks kedalaman kemiskinan sebesar 2,41; indeks keparahan kemiskinan sebesar 0,59; persentase TPT (Tingkat Pengangguran Terbuka) sebesar 4,56%; dan persentase TPAK (Tingkat Partisipasi Angkatan Kerja) sebesar 72,65%. Secara sederhana, klaster 1 adalah kumpulan wilayah yang lebih sejahtera, sedangkan klaster 2 adalah wilayah yang kurang sejahtera selama pandemi COVID-19 melanda di Provinsi Jawa Timur pada tahun 2020.

4. KESIMPULAN

Dari hasil penelitian dan analisis data, diperoleh jumlah kluster sebanyak 2 kluster, dengan nilai fungsi objektif (formula 1) diperoleh $J_{FCM} = 532,0957$. Kluster pertama beranggotakan 14 kabupaten dan 9 kota, serta kluster kedua beranggotakan 15 kabupaten. Anggota dari kluster 1 adalah Kabupaten Ponorogo, Kabupaten Tulungagung, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Lumajang, Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Pasuruan, Kabupaten Sidoarjo, Kabupaten Mojokerto, Kabupaten Jombang, Kabupaten Nganjuk, Kabupaten Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu. Sedangkan anggota dari kluster 2 adalah Kabupaten Pacitan, Kabupaten Trenggalek, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Probolinggo, Kabupaten Madiun, Kabupaten Magetan, Kabupaten Ngawi, Kabupaten Bojonegoro, Kabupaten Tuban, Kabupaten Lamongan, Kabupaten Bangkalan, Kabupaten Sampang, Kabupaten Pamekasan, dan Kabupaten Sumenep. Secara sederhana, kluster 1 adalah kumpulan wilayah yang lebih sejahtera, sedangkan kluster 2 adalah wilayah yang kurang sejahtera selama masa pandemi COVID-19 di tahun 2020.

Berdasarkan hasil dari nilai keseluruhan indeks validasi kluster, jumlah kluster yang optimal pada data Indikator Kesejahteraan Rakyat tahun 2020 berjumlah 2 kluster, dibuktikan dengan seluruh nilai pada *Cluster Validity Index* memihak pada 2 kluster, yaitu dengan nilai *Partition Entropy Index* = 0,42150448; *Partition Coefficient Index* = 0,7324392; *Xie Beni Index* = 0,2207773; *Modified Partition Coefficient Index* = 0,46487484; dan *Fukuyama-Sugeno Index* = -105546,3.

DAFTAR PUSTAKA

- Azizah, N., Yuniarti, D. and Goejantoro, R. (2018) 'Penerapan Metode Fuzzy Subtractive Clustering (Studi Kasus: Pengelompokan Kecamatan di Provinsi Kalimantan Timur Berdasarkan Luas Daerah dan Jumlah Penduduk Tahun 2015)', *Jurnal EKPOENSIAL*, 9, pp. 197–206.
- Bezdek, J. C. (1973) 'Cluster validity with fuzzy sets', *Journal of Cybernetics*, 3(3), pp. 58–73. doi: 10.1080/01969727308546047.
- Bezdek, J. C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. 1st edn. New York: Plenum Press. doi: 10.1007/978-1-4757-0450-1.
- Cebeci, Z. (2020) 'fvalid: An R Package for Internal Validation of Probabilistic and Possibilistic Clustering', *Sakarya University Journal of Computer and Information Sciences*, 3(1), pp. 11–26. doi: 10.35377/saucis.03.01.664560.
- Dave, R. N. (1996) 'Validating fuzzy partitions obtained through c-shells clustering', *Pattern Recognition Letters*, 17(6), pp. 613–623. doi: 10.1016/0167-8655(96)00026-8.
- Ferraro, M. B. and Giordani, P. (2015) 'A Toolbox for Fuzzy Clustering Using the R Programming Language', *Fuzzy Sets and Systems*, 279, pp. 1–16. doi: <https://doi.org/10.1016/j.fss.2015.05.001>.
- Haqiqi, B. N. and Kurniawan, R. (2015) 'Analisis Perbandingan Metode Fuzzy C-Means Dan Subtractive Fuzzy C-Means', *Media Statistika*, 8(2), pp. 59–67. doi: 10.14710/medstat.8.2.59-67.
- Kusumadewi, S. and Purnomo, H. (2013) *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan*. 2nd edn. Yogyakarta: Graha Ilmu.
- Liu, Y. et al. (2019) 'A Validity Index for Fuzzy Clustering Based on Bipartite Modularity', *Journal of Electrical and Computer Engineering*, 9, pp. 1–9. doi: 10.1155/2019/2719617.
- Mashfuufah, S. and Istiawan, D. (2018) 'Penerapan Partition Entropy Index, Partition Coefficient Index dan Xie Beni Index untuk Penentuan Jumlah Kluster Optimal pada Algoritma Fuzzy C-Means dalam Pemetaan Tingkat Kesejahteraan Penduduk Jawa Tengah', in *The 7th University Research Colloquium 2018*. Surakarta: STIKES PKU Muhammadiyah Surakarta, pp. 51–60.
- Wu, C. H. et al. (2015) 'A New Fuzzy Clustering Validity Index With a Median Factor for Centroid-Based Clustering', *IEEE Transactions on Fuzzy Systems*, 23(3), pp. 701–718. doi: 10.1109/TFUZZ.2014.2322495.
- Zarandi, M. H. F., Sotudian, S. and Castillo, O. (2021) 'A New Validity Index for Fuzzy-Possibilistic C-Means Clustering', *Scientia Iranica*, 28(4), pp. 2277–2293. doi: 10.24200/SCI.2021.50287.1614.