

Penerapan Algoritma *Fuzzy Possibilistic C-Means* (FPCM) pada Pengelompokan Kabupaten/Kota di Indonesia Berdasarkan Indikator Indeks Pembangunan Manusia Tahun

Ghia Fauziah Aghyari*, Abdul Kudus

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*ghiafauziah01@gmail.com, akudus69@unisba.ac.id

Abstract. Human resources are a crucial factor in human development and a key component in achieving prosperity in every country. The success of development is measured in various ways, one of the most popular being the calculation of the Human Development Index (HDI). The classification of districts and cities in Indonesia is necessary as a reference for government program planning and evaluation to enhance human development in those areas. Partitioning clustering is one of the clustering techniques that aims to partition data into several groups or partitions, with the number of groups usually predetermined. One of the algorithms used in partitioning clustering is Fuzzy Possibilistic C-Means (FPCM). Fuzzy Possibilistic C-Means (FPCM) is an extension of two algorithms, namely Fuzzy C-Means (FCM) and Possibilistic C-Means (PCM). FPCM combines fuzzy and possibilistic concepts to address the weaknesses of the previous algorithms. Therefore, the Fuzzy Possibilistic C-Means (FPCM) algorithm is applied to cluster the districts and cities in Indonesia based on the indicators of the Human Development Index. Based on the results of the Modified Partition Coefficient (MPC) index, the optimal number of clusters is determined to be four clusters. Cluster 1 contains 146 districts and cities, cluster 2 contains 97 districts and cities, cluster 3 contains 141 districts and cities, and cluster 4 contains 130 districts and cities.

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

Abstrak. Sumber daya manusia adalah faktor penting dalam pembangunan manusia yang menjadi komponen utama dalam mencapai kemakmuran di setiap negara. Keberhasilan pembangunan diukur dengan berbagai cara, salah satunya yang paling populer melalui perhitungan Indeks Pembangunan Manusia (IPM) atau *Human Development Index* (HDI). Pengelompokan wilayah Kabupaten/Kota di Indonesia perlu dilakukan sebagai acuan dalam perencanaan dan evaluasi program pemerintah untuk meningkatkan pembangunan manusia di daerah tersebut. *Partitioning clustering* adalah salah satu teknik pengelompokan yang mencoba mempartisi data ke dalam beberapa kelompok (*partition*) dan jumlah kelompok yang akan dibuat biasanya telah ditentukan sebelumnya dan salah satu algoritma pada *partitioning clustering* adalah *Fuzzy Possibilistic C-Means* (FPCM). Algoritma *Fuzzy Possibilistic C-Means* (FPCM) yang merupakan perluasan dari dua algoritma yaitu algoritma *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). *Fuzzy Possibilistic C-Means* (FPCM) menggabungkan konsep *fuzzy* dan *possibilistic* untuk mengurangi kelemahan dari algoritma sebelumnya. Oleh karena itu diterapkan algoritma *Fuzzy Possibilistic C-Means* (FPCM) untuk mengelompokan Kabupaten/Kota di Indonesia berdasarkan indikator indeks pembangunan manusia. Berdasarkan hasil indeks *Modified Partition Coefficient* (MPC) jumlah klaster yang optimal adalah empat klaster. Pada klaster 1 berisi 146 Kabupaten/Kota, klaster 2 berisi 97 Kabupaten/Kota, klaster 3 berisi 141 Kabupaten/Kota, dan klaster 4 berisi 130 Kabupaten/Kota.

Kata Kunci: *Indeks Pembangunan Manusia, Pengelompokan, Fuzzy Possibilistic C-Means.*

A. Pendahuluan

Sumber daya manusia merupakan faktor yang berperan penting dalam pembangunan manusia dan menjadi komponen utama dalam mencapai kemakmuran di setiap negara. Kemajuan suatu negara dapat dipacu dengan mudah apabila sumber daya manusia yang dimilikinya berkualitas. Maka dari itu, terdapat sebuah rancangan baru untuk mengukur kemajuan suatu negara yang berpusat pada kualitas sumber daya manusia (Farah Nurul Andini & Lisnur Wachidah, 2023). Keberhasilan pembangunan diukur dengan berbagai cara, di antaranya melalui perhitungan *Human Development Index* (HDI) atau disebut juga dengan Indeks Pembangunan Manusia (IPM), yang memiliki tiga dimensi dasar yaitu pengetahuan, kehidupan yang layak, serta panjang umur dan sehat.

Tinggi rendahnya IPM di Indonesia hanya dilihat melalui indeks komposit, tanpa menunjukkan indikator apa yang paling berpengaruh terhadap peningkatan IPM tersebut. Padahal, perlu dilakukan evaluasi terhadap nilai dari setiap indikator pembentuk IPM untuk mengetahui pencapaian masing-masing indikator. Selain itu, pentingnya mengelompokkan wilayah Kabupaten/Kota di Indonesia agar dapat menjadi acuan dalam perencanaan dan penilaian program-program pemerintah untuk meningkatkan pembangunan manusia di daerah tersebut. Tujuan dari pengelompokan wilayah adalah untuk memecah wilayah menjadi beberapa kelompok dengan karakteristik yang mirip pada setiap kelompoknya dan berbeda antar kelompok.

Clustering adalah salah satu teknik pada analisis data yang berguna untuk mengelompokkan sekumpulan data ke dalam beberapa kelompok atau klaster berdasarkan kemiripan atau kesamaan diantara data tersebut (Jafar & Sivakumar, 2012). *Clustering* bermula dari kebutuhan untuk mengelompokkan data dengan cepat dan efisien dalam jumlah yang besar. Teknik *clustering* diperkenalkan pertama kali pada tahun 1950-an oleh beberapa pakar statistik seperti Robert L. Thorndike, Lee A. Feldman, dan Maurice G. Kendall. *Partitioning clustering* adalah salah satu teknik pengelompokan yang mencoba mempartisi data ke dalam beberapa kelompok (*partition*) dan jumlah kelompok yang akan dibuat biasanya telah ditentukan sebelumnya (Kusumadewi, dkk. 2006). Tujuan utama dari *partitioning clustering* adalah mengelompokkan data sehingga objek dalam satu kelompok memiliki kemiripan yang tinggi satu sama lain, sedangkan objek antar kelompok memiliki perbedaan yang signifikan. *Partitioning clustering* memiliki kelebihan seperti kemampuan untuk menangani jumlah data yang besar, skalabilitas yang baik, dan hasil klaster yang jelas. Dalam praktiknya, *partitioning clustering* digunakan dalam berbagai bidang, termasuk analisis data, pengenalan pola, bioinformatika, pengelompokan pelanggan, dan sebagainya. Metode ini dapat memberikan wawasan yang berharga tentang struktur dan pola dalam data, serta membantu dalam pengambilan keputusan yang lebih baik dalam berbagai konteks. Metode *partitioning clustering* terdiri dari beberapa algoritma salah satunya yaitu *Fuzzy Possibilistic C-Means*.

Algoritma *Fuzzy Possibilistic C-Means* (FPCM) yang merupakan perluasan dari dua algoritma, yaitu algoritma *Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM). Pada *Fuzzy Possibilistic C-Means* (FPCM) menggabungkan konsep *fuzzy* dan *possibilistic* untuk mengurangi kelemahan dari kedua algoritma tersebut (Grover, 2014). Metode ini memungkinkan objek memiliki keanggotaan ganda pada beberapa kelompok dan memberikan kemungkinan adanya objek yang tidak tergabung dalam kelompok manapun. FPCM menggunakan dua parameter utama, yaitu parameter *fuzziness* dan parameter *possibilistic*. Parameter *fuzziness* menunjukkan sejauh mana sebuah objek dapat dimasukkan ke dalam beberapa kelompok, sedangkan parameter *possibilistic* menunjukkan sejauh mana sebuah objek dapat dimasukkan ke dalam kelompok tertentu.

Berdasarkan penjelasan pada latar belakang, untuk rumusan masalah dalam penelitian ini adalah “Bagaimana hasil pengelompokan wilayah kabupaten/kota di Indonesia berdasarkan indikator indeks pembangunan manusia tahun 2022 dengan menggunakan algoritma *Fuzzy Possibilistic C-Means*”. Selanjutnya, tujuan dari penelitian ini adalah untuk mengetahui hasil pengelompokan wilayah kabupaten/kota di Indonesia berdasarkan indikator indeks pembangunan manusia tahun 2022 dengan menggunakan algoritma *Fuzzy Possibilistic C-Means*.

B. Metodologi Penelitian

Data yang digunakan merupakan data sekunder yang merujuk pada data yang diperoleh dari sumber yang sudah ada dan dikumpulkan dan dicatat oleh pihak lain. Data yang digunakan pada skripsi ini adalah indikator dari Indeks Pembangunan Manusia dari 514 Kabupaten/Kota di seluruh Indonesia pada tahun 2022, yang diperoleh dari situs resmi Badan Pusat Statistik. Variabel bebas yang digunakan berjumlah 4 variabel.

Metode Analisis Data

Pertama, Analisis kluster. Analisis kluster adalah sebuah prosedur untuk mengelompokkan data berdasarkan kemiripan ciri-ciri yang dimiliki sehingga membentuk kelompok-kelompok dengan karakteristik yang serupa. Data-data yang memiliki ciri-ciri yang mirip akan dikelompokkan menjadi satu kelompok, sedangkan data yang memiliki perbedaan ciri-ciri akan dikelompokkan ke kelompok lain (Siyamto, 2017). Analisis kluster termasuk dalam kategori *unsupervised learning* pada ilmu data mining. Hal ini berarti bahwa pendekatan ini tidak memerlukan data latih atau tahap pembelajaran pada data yang ada (Kusuma & Nazori, 2015).

Kedua, Logika Fuzzy. Logika Fuzzy adalah suatu teknik dalam matematika yang yang dirancang untuk menangani ketidakpastian informasi dalam pengambilan keputusan dengan menggunakan himpunan *fuzzy*. Lotfi A. Zadeh (1965) memperkenalkan himpunan *fuzzy*, yang menunjukkan bahwa teori probabilitas bukan satu-satunya cara untuk menangani masalah ketidakpastian dalam pengambilan keputusan. Namun, himpunan *fuzzy* tidak bisa dikatakan sebagai pengganti teori probabilitas (Kusumadewi, dkk., 2006). Himpunan *fuzzy* menggambarkan nilai derajat keanggotaan yang dinyatakan dalam bilangan riil.

Atribut pada himpunan *fuzzy* dibagi menjadi dua jenis, yaitu atribut linguistik dan atribut numeris (Kusumadewi dan Purnomo, 2010).

1. Atribut linguistik adalah penamaan suatu grup yang mewakili suatu keadaan tertentu menggunakan bahasa alami.
2. Atribut numeris adalah suatu nilai yang mewujudkan ukuran dari suatu variabel. Dalam himpunan *fuzzy*, atribut numeris digunakan ketika variabel input atau output dapat diukur dengan angka atau bilangan riil

Ketiga, Algoritma Fuzzy Possibilistic C-Means. *Fuzzy Possibilistic C-Means* (FPCM) adalah variasi dari algoritma *clustering Fuzzy C-Means* (FCM) dan *Possibilistic C-Means* (PCM) yang menggabungkan sifat *fuzziness* dan *possibilistic*. FPCM menghitung derajat keanggotaan objek yang dilambangkan dengan (μ_{ik}) ke dalam suatu kelompok seperti halnya FCM, namun juga memperhitungkan derajat kemungkinan objek dilambangkan dengan (t_{ik}) yang menyatakan seberapa mungkin suatu objek termasuk ke dalam suatu kelompok (Jafar dan Sivakumar, 2012). Nilai derajat keanggotaan ini digunakan untuk melihat kecenderungan titik data terhadap suatu kelompok tertentu, di mana semakin tinggi nilai derajat keanggotaan pada suatu kelompok, maka semakin besar kecocokan data itu dengan karakteristik kelompok tersebut. Sedangkan untuk derajat kemungkinan digunakan untuk mengekspresikan tingkat keyakinan atau ketidakpastian terhadap keanggotaan suatu data pada kelompok tertentu, di mana semakin tinggi nilai derajat kemungkinan pada suatu kelompok, maka semakin besar kemungkinan data tersebut termasuk pada kelompok tersebut.

Fuzzy Possibilistic C-Means mengoptimalkan fungsi tujuan atau fungsi objektif yang terdiri dari tiga bagian yaitu jarak antara titik data dan pusat kluster, derajat keanggotaan *fuzzy* (μ_{ik}) , dan derajat kemungkinan (t_{ik}) . Persamaan dari fungsi objektif pada *Fuzzy Possibilistic C-Means* sebagai berikut:

$$J_{w,\eta}(U, T, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik}^w + t_{ik}^{\eta}) d_{ik}^2 \quad (1)$$

Dengan derajat keanggotaan dan derajat kemungkinan harus memenuhi kondisi berikut:

$$\sum_{i=1}^c \mu_{ik} = 1 ; \forall k \quad (2)$$

$$\sum_{k=1}^n t_{ik} = 1 ; \forall i \quad (3)$$

Di mana:

$J_{w,\eta}(U, T, V)$ = fungsi objektif terhadap U, T dan V

c = jumlah kluster/kelompok

n = jumlah objek

w = pembobot/pangkat pada matriks kekhasan relatif U

η = pembobot/pangkat pada matriks kekhasan absolut T

μ_{ik} = elemen- elemen matriks partisi U , atau fungsi keanggotaan data ke- k pada kluster ke- i

d_{ik} = fungsi jarak, jarak titik data ke- k dan pusat kluster ke- i

U = matriks partisi awal

Ukuran jarak yang digunakan adalah sebagai berikut:

$$(d_{ik})^2 = \|x_k - v_i\|^2 \quad (4)$$

Di mana:

d_{ik} = jarak titik data ke- k dan pusat kluster ke- i

v_i = pusat kluster ke- i

x_k = data pada objek ke- k

Keempat, Validitas Kluster. Validitas kluster adalah metode yang digunakan untuk mengevaluasi kualitas dari hasil *clustering* dan menentukan apakah kluster tersebut dapat menggambarkan seluruh data dengan baik atau tidak (Xie, dkk 2011). *Modified Partition Coefficient* (MPC) adalah salah satu metrik evaluasi untuk mengukur kualitas partisi hasil pengelompokan pada data mining. MPC sering digunakan pada pengelompokan berbasis partisi. Nilai MPC berada antara 0 sampai 1. Semakin besar nilai MPC yang diperoleh (mendekati 1) maka semakin bagus atau semakin optimal hasil *clustering* (Amirah, dkk., 2017). *Modified Partition Coefficient* (MPC) ini dapat digunakan untuk memilih jumlah kluster yang optimal dalam analisis kluster. Tujuannya untuk menemukan jumlah kluster yang dapat memberikan hasil pengelompokan yang baik dan berarti. Berikut merupakan persamaan untuk menghitung *Modified Partition Coefficient* (MPC):

$$MPC(c) = 1 - \frac{c}{c-1} (1 - PC(c)) \quad (5)$$

Di mana nilai indeks PC diperoleh menggunakan persamaan berikut:

$$PC(c) = \frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^2 \quad (6)$$

Di mana:

n = banyak objek penelitian

c = banyak kluster

μ_{ik} = nilai keanggotaan objek ke- i kluster ke- k

C. Hasil Penelitian dan Pembahasan

Pengelompokan dengan Algoritma *Fuzzy Possibilistic C-Means*

Sebelum melakukan pengelompokan menggunakan algoritma *Fuzzy Possibilistic C-Means*, harus dilakukan terlebih dahulu proses pengelompokan menggunakan algoritma *Fuzzy C-Means* untuk mendapatkan nilai akhir dari pusat kluster dan derajat keanggotaan yang selanjutnya akan digunakan pada pengelompokan data dengan menggunakan algoritma *Fuzzy Possibilistic C-Means*. Pada algoritma ini dilakukan penentuan parameter awal yaitu nilai pembobot (w) yang digunakan adalah 2, nilai pembobot (η) yang digunakan adalah 2, maksimum iterasi yang

digunakan adalah 100, *error* terkecil yang diharapkan (ϵ) sebesar 10^{-5} , fungsi objektif awal yang digunakan adalah 0 dan iterasi awal dimulai pada iterasi 1. Untuk menentukan jumlah kluster pada algoritma *Fuzzy Possibilistic C-Means*, dilakukan uji coba dengan membandingkan nilai validasi kluster dari jumlah kluster 2 sampai 5. Setelah itu, jumlah kluster yang paling optimal akan dipilih untuk digunakan dalam proses pengelompokan data. Jumlah kluster optimal dapat ditentukan dengan melihat nilai *Modified Partition Coefficient* (MPC) tertinggi. Berikut adalah hasil nilai MPC pada FPCM untuk menentukan jumlah kluster optimal.

Tabel 1. Penentuan Jumlah Kluster Optimal.

Jumlah Kluster	MPC
2	0,3355122
3	0,3085459
4	0,358196
5	0,2795686

Berdasarkan Tabel 1, nilai *Modified Partition Coefficient* (MPC) tertinggi pada metode *Fuzzy Possibilistic C Means* adalah 0,358196 dengan jumlah kluster empat. Oleh karena itu, kluster yang optimal untuk digunakan adalah sebanyak empat kluster.

Pembuatan Profil Hasil Analisis Kluster

Anggota dari hasil kluster Kabupaten/Kota di Indonesia dengan metode *Fuzzy Possibilistic C-Means* (FPCM) disajikan pada Tabel 2.

Tabel 2. Hasil Kluster

Kluster	Jumlah Kabupaten/Kota
1	146
2	97
3	141
4	130

Berdasarkan Tabel 2, kluster 1 memiliki jumlah anggota sebanyak 146 Kabupaten/Kota, kluster 2 memiliki jumlah anggota sebanyak 97 Kabupaten/Kota, kluster 3 memiliki jumlah anggota sebanyak 141 Kabupaten/Kota, dan kluster 4 memiliki jumlah anggota sebanyak 130 Kabupaten/Kota. Kluster yang memiliki jumlah anggota terbanyak yaitu kluster 1 dan kluster yang memiliki jumlah anggota yang paling sedikit yaitu kluster 2.

Selanjutnya yaitu identifikasi karakteristik kluster, yang merupakan proses akhir dalam membentuk kluster. Untuk mengidentifikasi karakteristik dapat dilihat melalui nilai rata-rata setiap variabel pada setiap klasternya.

Tabel 3. Nilai Rata-Rata Kluster Pada Setiap Variabel

Pusat Kluster	Variabel/Atribut				Kategori
	Umur Harapan Hidup	Rata-Rata Lama Sekolah	Harapan Lama Sekolah	Pengeluaran Riil Per Kapita	
Kluster 1	71,46233	8,297671	12,85719	10854,21	Tinggi
Kluster 2	67,13268	6,809485	11,94309	8273,814	Rendah
Kluster 3	67,71468	8,223759	13,01688	9193,809	Sedang
Kluster 4	72,69531	10,49531	14,28015	13746,67	Sangat Tinggi

Berdasarkan Tabel 3 setiap klaster memiliki tafsiran nilai yang dihasilkan dari masing-masing variabel untuk menunjukkan karakteristik pada masing-masing klaster. Dari keempat klaster yang terbentuk, maka dibagi menjadi empat kategori yaitu kategori dengan rata-rata klaster rendah, sedang, tinggi, dan sangat tinggi. Berikut merupakan penjelasan untuk masing-masing klaster.

1. Klaster 1, Klaster 1 memiliki nilai rata-rata untuk variabel umur harapan hidup sebesar 71,46233 tahun, rata-rata lama sekolah sebesar 8,297671 tahun, harapan lama sekolah 12,85719 tahun, dan pengeluaran riil per kapita sebesar Rp. 10.854.210,- untuk tiap individu per tahunnya. Berdasarkan ciri-ciri yang dimiliki oleh klaster 1, maka dapat disimpulkan bahwa klaster 1 dikategorikan sebagai rata-rata klaster yang tinggi.
2. Klaster 2, Klaster 2 memiliki nilai rata-rata untuk variabel umur harapan hidup sebesar 67,13268 tahun, rata-rata lama sekolah sebesar 6,809485 tahun, harapan lama sekolah 11,94309 tahun, dan pengeluaran riil per kapita sebesar Rp. 8.273.814,- untuk tiap individu per tahunnya. Berdasarkan ciri-ciri yang dimiliki oleh klaster 2, dapat dilihat bahwa klaster 2 memiliki nilai rata-rata paling rendah dibandingkan dengan klaster lainnya pada setiap variabelnya maka dapat disimpulkan bahwa klaster 2 dikategorikan sebagai rata-rata klaster yang rendah.
3. Klaster 3, Klaster 3 memiliki nilai rata-rata untuk variabel umur harapan hidup sebesar 67,71468 tahun, rata-rata lama sekolah sebesar 8,223759 tahun, harapan lama sekolah 13,01688 tahun, dan pengeluaran riil per kapita sebesar Rp. 9.193.809,- untuk tiap individu per tahunnya. Berdasarkan ciri-ciri yang dimiliki oleh klaster 3, maka dapat disimpulkan bahwa klaster 3 dikategorikan sebagai rata-rata klaster yang sedang.
4. Klaster 4, Klaster 4 memiliki nilai rata-rata untuk variabel umur harapan hidup sebesar 72,69531 tahun, rata-rata lama sekolah sebesar 10,49531 tahun, harapan lama sekolah 14,28015 tahun, dan pengeluaran riil per kapita sebesar Rp. 13.746.670,- untuk tiap individu per tahunnya. Berdasarkan ciri-ciri yang dimiliki oleh klaster 4, dapat dilihat bahwa klaster 4 memiliki nilai rata-rata paling tinggi dibandingkan dengan klaster lainnya pada setiap variabelnya maka dapat disimpulkan bahwa klaster 4 dikategorikan sebagai rata-rata klaster yang sangat tinggi.

Identifikasi karakteristik klaster dilakukan menggunakan metode rata-rata. Dari 514 Kabupaten/Kota di Indonesia, diambil rata-rata dari masing-masing variabel indikator IPM, yaitu umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil per kapita. Selanjutnya, rata-rata juga diambil untuk setiap indikator IPM dalam setiap klaster yang terbentuk. Dalam menentukan karakteristik klaster, setiap variabel yang ada dalam klaster diberi tanda (+) apabila rata-rata klaster lebih besar atau sama dengan rata-rata keseluruhan. Sedangkan jika rata-rata klaster lebih kecil daripada rata-rata keseluruhan, maka diberi tanda (-). Dengan demikian, dengan melihat tanda (+) atau (-), kita dapat mengidentifikasi karakteristik masing-masing variabel dalam setiap klaster, apakah memiliki nilai yang lebih tinggi atau lebih rendah dibandingkan dengan nilai rata-rata keseluruhan. Karakteristik tersebut disajikan pada Tabel 4.

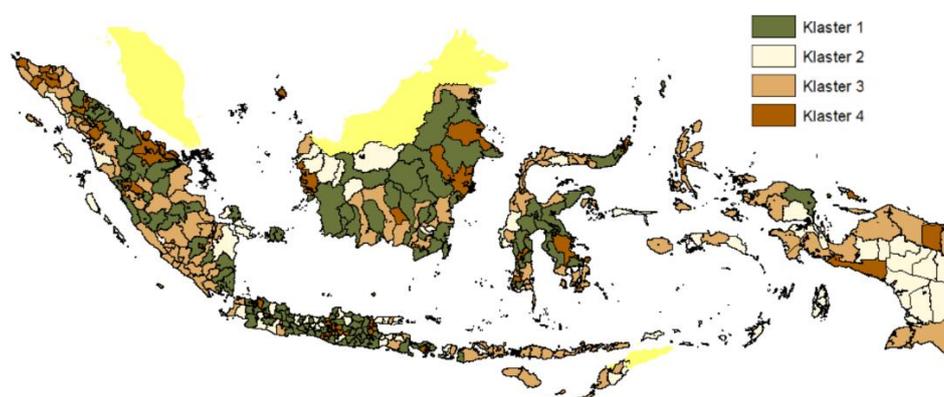
Tabel 4. Identifikasi Karakteristik Klaster

Klaster	Umur Harapan Hidup	Rata-Rata Lama Sekolah	Harapan Lama Sekolah	Pengeluaran Riil Per Kapita
1	+	-	-	+
2	-	-	-	-
3	-	-	-	-
4	+	+	+	+

Berdasarkan Tabel 4 maka dapat diidentifikasi karakteristik setiap klaster dimana untuk klaster 1 memiliki nilai rata-rata untuk variabel rata-rata lama sekolah dan harapan lama sekolah dibawah rata-rata keseluruhan, sedangkan untuk variabel umur harapan hidup dan pengeluaran

riil per kapita berada diatas rata-rata keseluruhan. Klaster 2 memiliki nilai rata-rata untuk variabel umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil perkapita dibawah rata-rata keseluruhan. Klaster 3 memiliki nilai rata-rata untuk variabel umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil perkapita dibawah rata-rata keseluruhan. Dan klaster 4 memiliki nilai rata-rata untuk variabel umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil perkapita diatas rata-rata keseluruhan.

Selanjutnya visualisasi hasil klaster dilakukan menggunakan diagram peta (kartogram) untuk menggambarkan Kabupaten/Kota di Indonesia sesuai dengan klasternya masing-masing. Diagram peta disajikan pada Gambar 1.



Gambar 1. Daerah Hasil Klaster dengan Metode *Fuzzy Possibilistic C-Means*

Berdasarkan Gambar 1 menggambarkan daerah Kabupaten/Kota di Indonesia berdasarkan hasil klaster. Dari visualisasi diagram peta tersebut, dapat dilihat bahwa Kabupaten/Kota di Indonesia berdasarkan indikator indeks pembangunan manusia terbagi menjadi empat klaster yang ditandai adanya empat perbedaan warna. Klaster 1 ditandai dengan warna hijau tua merupakan kategori rata-rata klaster tinggi. Klaster 2 ditandai dengan warna krem merupakan kategori rata-rata klaster rendah. Klaster 3 ditandai dengan warna coklat muda merupakan kategori rata-rata klaster sedang. Klaster 4 ditandai dengan warna coklat tua merupakan kategori rata-rata klaster sangat tinggi.

Indonesia memiliki potensi besar dalam jumlah penduduk dan sumber daya alam yang bisa dimanfaatkan untuk pembangunan yang optimal. Namun, pada kenyataannya, pembangunan manusia di Indonesia masih terhambat oleh kesenjangan antar wilayah yang disebabkan oleh berbagai faktor. Setiap daerah di Indonesia memiliki karakteristik unik, sehingga faktor yang memengaruhi perkembangan pembangunan manusia di setiap daerah juga berbeda-beda. Capaian pembangunan manusia dipengaruhi oleh beberapa faktor, di antaranya adalah kesehatan, pendidikan, dan ekonomi. Oleh karena itu, dibutuhkan strategi yang tepat untuk mencapai pembangunan manusia yang lebih baik dan berkelanjutan.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, dapat disimpulkan bahwa pengelompokan wilayah menggunakan algoritma *Fuzzy Possibilistic C-Means* (FPCM) telah mengelompokkan 514 Kabupaten/Kota di Indonesia berdasarkan indikator indeks pembangunan manusia menjadi 4 klaster.

1. Klaster 1 merupakan kelompok Kabupaten/Kota dengan rata-rata klaster yang tinggi. Klaster ini ditandai dengan nilai rata-rata lama sekolah dan harapan lama sekolah yang berada di bawah rata-rata, sementara umur harapan hidup dan pengeluaran riil per kapita berada di atas rata-rata. Terdapat 146 Kabupaten/Kota yang termasuk dalam Klaster 1.
2. Klaster 2 adalah kelompok Kabupaten/Kota dengan rata-rata klaster yang rendah. Klaster

ini memiliki nilai rata-rata untuk variabel umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil per kapita yang berada di bawah rata-rata. Terdapat 97 Kabupaten/Kota yang termasuk dalam Klaster 2.

3. Klaster 3 merupakan kelompok Kabupaten/Kota dengan rata-rata klaster yang sedang. Klaster ini memiliki nilai rata-rata untuk variabel umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil per kapita yang berada di bawah rata-rata, namun masih di atas rata-rata Klaster 2. Terdapat 141 Kabupaten/Kota yang termasuk dalam Klaster 3.
4. Klaster 4 adalah kelompok Kabupaten/Kota dengan rata-rata klaster yang sangat tinggi. Klaster ini memiliki nilai rata-rata untuk variabel umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran riil per kapita yang berada di atas rata-rata. Terdapat 130 Kabupaten/Kota yang termasuk dalam Klaster 4.

Acknowledge

Puji dan syukur penulis panjatkan kehadirat Allah SWT karena berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan penelitian ini. Terimakasih kepada ayah, ibu, kakak-kakak dan keluarga yang selalu mendo'akan dan memberi dukungan baik moral maupun materi kepada penulis. Terimakasih kepada Bapak Abdul Kudus, S.Si., M.Si., Ph.D. yang telah memberi bimbingan kepada penulis hingga penelitian ini selesai. Dosen-dosen Program Studi Statistika Universitas Islam Bandung yang telah banyak memberikan ilmu pengetahuan. Sahabat dan teman-teman serta semua pihak yang telah hingga penelitian ini selesai.

Daftar Pustaka

- [1] Amirah, M. M., Widodo, A. W., & Dewi, C. (2017). *Pengelompokan Lagu Berdasarkan Emosi Menggunakan Algoritma Fuzzy C-Means*. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer.
- [2] Grover, N. (2014). *A Study of Various Fuzzy Clustering Algorithms*. International Journal of Engineering Research. 3(3): 177-181.
- [3] Jafar, M., & R. Sivakumar. (2012). *A Study on Possibilistic and Fuzzy Possibilistic C-Means Clustering Algorithms for Data Clustering*. International Conference on Emerging Trends in Science, Engineering and Technology.
- [4] Kusuma, D. T. dan Nazori, A. (2015). *Prototipe Komparasi Model Clustering Menggunakan Metode K-Means dan FCM Untuk Menentukan Strategi Promosi: Studi Kasus Sekolah Tinggi Teknik PLN Jakarta*. Jurnal TICOM, 1-10.
- [5] Kusumadewi, S., Hartati, S., Harjoko, A., dan Wardoyo, R. (2006). *Fuzzy Multi Attribute Decision Making (FMADM)*. Yogyakarta: Graha ilmu.
- [6] Kusumadewi, S. & Purnomo, H. (2010). *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan*. Yogyakarta: Graha Ilmu.
- [7] Siyamto, Y. (2017). *Pemanfaatan Data Mining dengan Metode Clustering untuk Evaluasi Biaya Dokumen Ekspor di PT Winstar Batam*. Media Informatika Budidarma. 1(2): 28-31.
- [8] Xie, N., Hu, L., Luktarhan, N., dan Zhao, K. (2011). *Web Information Systems and Mining*. Springer.
- [9] Farah Nurul Andini, & Lisnur Wachidah. (2023). Penerapan Regresi Spasial Panel Random Effect pada Kasus Kemiskinan di Provinsi Jawa Tengah Tahun 2011-2020. Jurnal Riset Statistika, 61–70. <https://doi.org/10.29313/jrs.v3i1.1885>