



**CLUSTERING KABUPATEN/KOTA DI JAWA TIMUR BERDASARKAN
INDIKATOR IPM MENGGUNAKAN SOM DAN FCM**

***DISTRICT/CITY CLUSTERING IN EAST JAVA BASED ON HUMAN
DEVELOPMENT INDEX INDICATORS USING THE SELF-ORGANIZING MAPS (SOM)
AND FUZZY C-MEANS (FCM) METHODS***

Choirul Fatihin, Prodi Matematika FMIPA UNY

Sri Andayani*, Prodi Matematika FMIPA UNY

*e-mail: andayani@uny.ac.id

Abstrak. Salah satu upaya pemerintah dalam mengukur pembangunan manusia yaitu melalui Indeks Pembangunan Manusia (IPM). IPM di Indonesia mengalami peningkatan namun tidak lebih besar dari tahun sebelumnya akibat dari covid-19. Setiap daerah tentunya memiliki tingkat keberhasilan pembangunan manusia yang berbeda, sehingga perlu dilakukan analisis hasil pembangunan manusia, salah satunya menggunakan analisis *clustering*. Hasil analisis dapat dimanfaatkan pemerintah sebagai bahan evaluasi dan perencanaan sasaran program pemerintah yang sesuai di setiap daerah. Penelitian ini bertujuan untuk mengetahui hasil *clustering* kabupaten/kota di Jawa Timur berdasarkan indikator IPM menggunakan metode *Self-Organizing Maps* (SOM) dan *Fuzzy C-Means* (FCM), dan mengetahui karakteristik masing-masing kelompok dari metode terbaik. Penelitian ini menggunakan data indikator IPM Provinsi Jawa Timur tahun 2021 yang diperoleh dari laman resmi Badan Pusat Statistik Provinsi Jawa Timur. Data yang digunakan terdiri dari 38 kabupaten/kota dan 4 variabel. Metode validasi kluster yang digunakan adalah *Davies-Bouldin Index* (DBI) dan *Calinski-Harabasz Index* (CHI). Berdasarkan DBI diperoleh metode *clustering* terbaik adalah SOM dengan jumlah kluster 3, *learning rate* 0,1 dan normalisasi min-max dengan indeks sebesar 0,74823. Berdasarkan CHI diperoleh metode *clustering* terbaik adalah FCM dengan jumlah kluster 3, bobot pangkat 1,5, dan normalisasi min-max dengan indeks sebesar 43,9344.

Kata kunci: *indeks pembangunan manusia, analisis clustering, self-organizing maps, fuzzy c-means.*

Abstract. One of the government's efforts to measure human development is through the human development index (HDI). HDI in Indonesia has increased but not greater than the previous year as a result of co-19. Each region certainly has a different level of success in human development, so it is necessary to analyze the results of human development, one of which is using clustering analysis. The results of the analysis can be used by the government as material for evaluating and planning appropriate government program targets in each region. This study aims to determine the results of district/city clustering in East Java based on HDI indicators using the Self-Organizing Maps (SOM) and Fuzzy C-Means (FCM) methods and to determine the characteristics of each group from the best method. This study uses data from the East Java Province HDI indicator for 2021 obtained from the official website of the Central Bureau of Statistics for East Java Province. The data used consists of 38 districts/cities and 4 variables. The cluster validation methods used are the Davies-Bouldin Index (DBI) and the Calinski-Harabasz Index (CHI). Based on DBI, the best clustering method is SOM with 3 clusters, 0,1 learning rate and min-max normalization with an index of 0,74823. Based on CHI, the best clustering method obtained is FCM with a number of clusters of 3, weights to the power of 1,5, and min-max normalization with an index of 43,9344.

Keywords: *human development index, clustering analysis, self-organizing maps, fuzzy c-means.*

PENDAHULUAN

Pembangunan merupakan upaya terencana pemerintah dalam meningkatkan kesejahteraan masyarakat di berbagai bidang. Suatu bangsa perlu adanya pembangunan untuk mencapai tujuan nasional dalam meningkatkan taraf hidup masyarakat. Pembangunan yang dijalankan pemerintah dapat berhasil dilakukan membutuhkan dukungan dan partisipasi dari masyarakat. Pembangunan ini dilakukan untuk mensejahterakan masyarakat berdasarkan sumber daya wilayah tersebut yang dapat ditingkatkan. Salah satu upaya pemerintah dalam mengukur pembangunan yaitu melalui indeks pembangunan manusia atau yang lebih dikenal dengan istilah IPM.

Indeks Pembangunan Manusia (IPM) pertama kali diperkenalkan oleh United Nations Development Programme (UNDP) pada tahun 1990 melalui Human Development Report (HDR) yang dipublikasikan dalam laporan tahunan secara berkala. IPM menjelaskan bagaimana masyarakat dapat memperoleh pendapatan, pendidikan, kesehatan, dan lainnya melalui hasil pembangunan. Pembangunan manusia bukan hanya tentang pertumbuhan ekonomi dimana masyarakat hidup, namun bagaimana peluang masyarakat dalam mengembangkan kemampuannya dan membangun keterampilan yang dimiliki.

Berdasarkan data Badan Pusat Statistik, capaian IPM di Indonesia mengalami peningkatan setiap tahun, namun peningkatan tersebut tidak lebih besar dari peningkatan pada tahun sebelumnya. Tahun 2020, IPM di Indonesia mengalami pertumbuhan hanya sebesar 0,03 persen, pertumbuhan ini terjadi seiring dengan sebaran covid-19 yang semakin luas dengan pemberlakuan pembatasan kegiatan masyarakat. Seiring berjalan waktu hingga memasuki tahun 2021, masyarakat Indonesia dapat beradaptasi dengan pandemi dari tahun sebelumnya. Adanya vaksinasi serta berbagai kebiasaan baru yang disesuaikan dengan keadaan dapat mengendalikan penyebaran covid-19. Akibatnya IPM Indonesia tumbuh lebih baik dari pada tahun 2020, yaitu sebesar 0,49 persen.

Provinsi di Pulau Jawa yang memiliki IPM terendah pada tahun 2021 diduduki oleh Jawa Timur sebesar 72,14. Namun, Jawa Timur juga menduduki selisih IPM tertinggi di Pulau Jawa, yaitu sebesar 0,43 poin. Sehingga terlihat peningkatan yang baik di Jawa Timur dari tahun 2020 ke tahun 2021 (Badan Pusat Statistik, 2021).

Rendahnya IPM di Jawa Timur dibandingkan dengan lima provinsi yang ada di Pulau Jawa dikarenakan angka pendidikan rendah yang diakibatkan oleh sebagian besar masyarakat Jawa Timur yang hanya menempuh pendidikan sampai dengan Sekolah Menengah Pertama dan gizi masyarakat yang kurang baik di beberapa daerah Jawa Timur (Cahyanti, Muchtolifah, & Sishadiyati, 2021). Meningkatnya IPM pada tahun 2021 dibandingkan dengan tahun sebelumnya menunjukkan adanya pemulihan pembangunan manusia di Jawa Timur pasca pandemi covid-19. Secara logika, hal ini terjadi karena masyarakat yang mulai dapat menyesuaikan diri dengan kondisi covid-19 dari tahun sebelumnya yang mana terjadi pembatasan sosial.

IPM menjadi dasar tolak ukur bagi pemerintah dalam menentukan keberhasilan pembangunan manusia di setiap daerah. Namun setiap daerah tentunya memiliki tingkat keberhasilan pembangunan manusia yang berbeda sehingga perlu dilakukan analisis hasil pembangunan manusia. Hasil analisis ini nantinya dapat dimanfaatkan pemerintah sebagai bahan evaluasi dan perencanaan sasaran program pemerintah yang sesuai di setiap daerah. Salah satu hal yang dapat dilakukan untuk memperoleh hasil analisis tersebut yaitu dengan melakukan analisis *clustering* kabupaten/kota di Jawa Timur berdasarkan indikator IPM menggunakan metode *Self-Organizing Maps* (SOM) dan *Fuzzy C-Means* (FCM).

SOM merupakan *unsupervised learning* yang mampu memproses informasi dalam dimensi yang besar dengan visualisasi yang efektif dari dimensi yang tinggi ke dimensi yang rendah (Nawaratne et al., 2019). SOM merupakan salah satu jenis jaringan syaraf tiruan yang dapat digunakan secara efisien dalam analisis data statistik dan visual, terutama untuk data

bervolume tinggi dan tidak seragam (Faradonbeh et al., 2020). Akan tetapi SOM lambat konvergen menuju titik pusat klaster dan hasil klaster terkadang tidak tepat (Nugroho, Hendrawan, & Hafidz, 2012).

FCM merupakan metode yang mampu mengelompokkan data yang tersebar secara tidak teratur, karena ketidakteraturan suatu data memungkinkan data tersebut mempunyai karakteristik tertentu dari klaster lainnya (Poerwanto & Ali, 2019). Akan tetapi FCM sensitif terhadap nilai awal, sehingga mudah terjebak solusi optimal lokal (Dhanachandra & Chanu, 2020).

Adapun penelitian terdahulu untuk merujuk beberapa penelitian yang menggunakan metode SOM dan FCM. Deni, Midyanti, & Hidayati (2022) menggunakan metode SOM dengan jumlah klaster sebanyak 3 untuk penentuan pencemaran air. Herlinda, Darwis, dan Dartono (2021) menggunakan metode FCM dengan jumlah klaster sebanyak 2 untuk rekredensialing fasilitas kesehatan.

Perhitungan dengan metode SOM dan FCM membutuhkan penentuan awal jumlah klaster yang dibutuhkan karena jika peneliti sembarang menentukan jumlah klaster maka hasil *clustering* akan menjadi tidak optimal atau terjadi inkonsistensi (Deana, Suhaedi, & Harahap, 2022). Sehingga dibutuhkan sebuah metode validasi klaster agar jumlah klaster yang diperoleh optimal dengan data yang digunakan. Pada penelitian ini untuk validasi klaster menggunakan *davies-bouldin index* (DBI) dan *calinski-harabasz index* (CHI) sebagai metode yang membandingkan hasil *clustering* dari metode SOM dan FCM.

Berdasarkan uraian di atas, penelitian ini akan melakukan analisis clustering menggunakan metode SOM dan FCM untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan indikator indeks pembangunan manusia tahun 2021. Hasil *clustering* dapat digunakan sebagai bahan evaluasi pemerintah dalam perencanaan sasaran program pemerintah yang sesuai di setiap daerah.

METODE

Penelitian ini menggunakan data indikator Indeks Pembangunan Manusia (IPM) Kabupaten/Kota di Jawa Timur tahun 2021 yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Timur pada laman resmi (<https://jatim.bps.go.id>). Data terdiri dari 38 Kabupen/Kota Jawa Timur dan 4 variabel berupa indikator IPM yang dapat dilihat pada Tabel 1.

Tabel 1. Variabel Penelitian

Indikator	Definisi
Angka Harapan Hidup (AHH)	Angka harapan hidup adalah rata-rata perkiraan waktu (dalam tahun) untuk seseorang menjalani hidupnya (BPS, 2021:11).
Harapan Lama Sekolah (HLS)	Harapan lama sekolah adalah lamanya (dalam tahun) diharapkan akan dilalui anak usia 7 tahun untuk bersekolah (BPS, 2021:11).
Rata-rata Lama Sekolah (RLS)	Rata-rata lama sekolah adalah jumlah tahun yang dilalui seseorang usia 25 tahun ke atas dalam menempuh pendidikan formal (BPS, 2021:11).
Pengeluaran Riil Per Kapita Disesuaikan (PKR)	Pengeluaran riil per kapita disesuaikan adalah kemampuan daya beli masyarakat dalam periode waktu tertentu (BPS, 2021:22).

Langkah-langkah dalam penelitian ini secara umum, yaitu pengumpulan data, *preprocessing* data, *clustering* menggunakan metode *Self-Organizing Maps* (SOM) dan *Fuzzy C-Means* (FCM), validasi klaster, dan analisis karakteristik. Tahap pertama diawali dengan pengumpulan data, *preprocessing* data dengan mengecek data kosong dan melakukan

normalisasi data menggunakan *min-max* (Persamaan 1) dan *z-score* (Persamaan 2).

$$X = \frac{x' - \min(x)}{\max(x) - \min(x)} \quad (1)$$

$$X = \frac{x' - \mu}{\sigma} \quad (2)$$

Keterangan:

X : atribut data hasil normalisasi

x' : atribut data

min(x) : nilai minimal dari x

max(x) : nilai maksimal dari x

μ : rata-rata data (mean)

σ : deviasi standar data (standart deviation)

Selanjutnya, yaitu *clustering* menggunakan metode *Self-Organizing Maps* (SOM) dan *Fuzzy C-Means* (FCM). Adapun algoritma SOM menurut Fausett (1994:170-172) sebagai berikut:

1. Inisialisasi bobot dan tentukan jumlah variabel (m), jumlah data (n), *learning rate* (α), jumlah kluster (k), dan jumlah maksimum iterasi ($t = MaxIter$).
2. Langkah 3-8 jika kondisi berhenti belum terpenuhi.
3. Langkah 4-6 untuk setiap data input x .
4. Untuk setiap j , hitunglah jarak Euclidean antara data dan bobot menggunakan Persamaan 3.

$$D(j) = \sum_{i=1}^m (w_{ij} - x_i)^2 \quad (3)$$

Keterangan:

D(j) : jarak euclidean pada kluster ke-j

w_{ij} : bobot variabel ke-i kluster ke-j

x_i : data variabel ke-i

5. Temukan indeks J sedemikian sehingga D(J) minimum, yang mana indeks J menunjukkan bobot kluster pemenang.
6. Perbarui bobot pemenang dan tetangganya menggunakan Persamaan 4.

$$w_{ij}(\text{baru}) = w_{ij}(\text{lama}) + \alpha[x_i - w_{ij}(\text{lama})] \quad (4)$$

7. Perbarui *learning rate*.
8. Cek kondisi berhenti.
9. Simpan bobot yang sudah konvergen.

Adapun algoritma FCM menurut Bezdek (1981) sebagai berikut:

1. Tentukan jumlah kluster ($k(1 \leq k \leq c)$), bobot pangkat (w dengan $w > 1$), jumlah iterasi maksimum ($MaxIter$), fungsi objektif awal ($J_0 = 0$), iterasi awal ($t = 1$) dan *error* terkecil yang diharapkan (ε (nilai positif yang kecil)).
2. Menentukan matriks partisi awal dengan mengikuti syarat pada Persamaan 5, 6, dan 7.

$$\mu_{ik}(x) \in [0,1] \text{ dengan } (2 \leq k \leq i) \quad (5)$$

$$\sum_{k=1}^c \mu_{ik} = 1, \forall k \in \{1,2, \dots, c\} \quad (6)$$

$$0 < \sum_{i=1}^n \mu_{ik} < n, \forall i \in \{1,2, \dots, n\} \quad (7)$$

3. Menghitung *centroid* menggunakan Persamaan 8.

$$c_{ij} = \frac{\sum_{i=1}^m (u_{il})^w x_{lj}}{\sum_{i=1}^m (u_{il})^w} \quad (8)$$

Keterangan:

c_{ij} : *centroid* data ke-i kluster ke-j
 u_{il} : derajat keanggotaan pada data ke-i
 x_{lj} : data pada kluster ke-j
 m : jumlah data

4. Menghitung fungsi objektif menggunakan Persamaan 9.

$$J = \sum_{i=1}^m \sum_{j=1}^k (u_{ij})^w D(x_i, c_j)^2 \quad (9)$$

Keterangan:

$D(x_i, c_j)$: jarak antar data dan *centroid*
 J : fungsi objektif

5. Memperbaiki derajat keanggotaan pada matriks menggunakan Persamaan 10.

$$u_{ij} = \frac{\frac{1}{D(x_i, c_j)^{\frac{2}{w-1}}}}{\sum_{l=1}^k \frac{1}{D(x_i, c_l)^{\frac{2}{w-1}}}} \quad (10)$$

Setelah diperoleh hasil *clustering* dari metode SOM dan FCM, dilakukan validasi kluster menggunakan metode *Davies-Bouldin Index* (DBI) dan *Calinski-Harabasz Index* (CHI). DBI merupakan salah satu metode validasi kluster yang memaksimalkan jarak antar kluster dan meminimalkan jarak antar titik dalam sebuah kluster (Sitompul, Sitompul, & Sihombing, 2019). Nilai DBI akan semakin bagus jika nilai DBI mendekati nol, namun nilainya tidak negatif. Adapun persamaan yang digunakan seperti pada Persamaan 11.

$$DBI = \frac{1}{K} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (11)$$

Keterangan:

K : jumlah kluster yang digunakan

$R_{i,j}$ adalah perhitungan rasio untuk menghitung nilai perbandingan antara kluster ke-i dan kluster ke-j yang diformulasikan pada Persamaan 12.

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (12)$$

Keterangan:

SSW_i : SSW dari kluster ke-i, $i=1, \dots, k-1$

SSW_j : SSW dari kluster ke-j, $j=2, \dots, k$

$SSB_{i,j}$: SSB dari kluster ke-i dan ke-j

SSW (*Sum of square within cluster*) atau kohesi adalah keterikatan antar data dalam kluster yang diformulasikan pada Persamaan 13.

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \quad (13)$$

Keterangan:

m_i : jumlah data dalam kluster ke-i
 $D(x_j, c_i)$: jarak data ke *centroid*
 x_j : data pada kluster tersebut
 c_i : *centroid* kluster ke-i

SSB (*Sum of square between cluster*) atau separasi adalah jarak antar kluster yang diformulasikan pada Persamaan 14.

$$SSB_{i,j} = d(c_i, c_j) \quad (14)$$

CHI merupakan salah satu metode validasi kluster yang menghitung perbandingan antara nilai SSB dan SSW yang dikalikan dengan faktor normalisasi, yaitu selisih antara jumlah data dan jumlah kluster dibanding dengan jumlah kluster dikurang satu (Khairati, Adlina, & Handari, 2019). Semakin besar nilai CHI, maka semakin baik hasil *clustering* yang diperoleh. Adapun persamaan yang digunakan seperti pada Persamaan 15 (Aliyatussya'ni & Saputro, 2022).

$$CHI = \frac{SSB}{SSW} \times \frac{N - k}{k - 1} \quad (15)$$

SSW (*Sum of square within cluster*) atau kohesi adalah keterikatan antar data dalam kluster yang diformulasikan pada Persamaan 16 (Wang & Xu, 2019).

$$SSW = \sum_{l=1}^k \sum_{x_l \in C_l} \|x_l - \bar{x}_l\|^2 \quad (16)$$

Keterangan:

\bar{x}_l : rata-rata data (titik pusat) pada setiap kluster ke-l
 C_l : kluster ke-l
 x_l : titik data ke-i pada kluster ke-l

SSB (*Sum of square between cluster*) atau separasi adalah jarak antar kluster yang diformulasikan pada Persamaan 17 (Wang & Xu, 2019).

$$SSB = \sum_{l=1}^k N_l \| \bar{x}_l - \bar{x} \|^2 \quad (17)$$

Keterangan:

\bar{x} : rata-rata data (titik pusat) keseluruhan
 N_l : jumlah titik data pada kluster ke-l

Tahap terakhir yaitu analisis karakteristik yang dilakukan dengan menghitung rata-rata variabel dari setiap kluster. Rata-rata tersebut menunjukkan karakteristik dari setiap kluster yang diperoleh.

HASIL DAN PEMBAHASAN

Hasil

1. *Preprocessing Data*

Sebelum data digunakan kedalam penelitian, data akan diperiksa terlebih dahulu apakah terdapat data kosong (*missing value*) seperti pada Gambar 1. Pemeriksaan data kosong dilakukan karena data kosong dapat mempengaruhi hasil analisis data.

```

Kabupaten/Kota    0
AHH                0
HLS                0
PKR                0
RLS                0
IPM                0
dtype: int64

```

Gambar 1. Jumlah Data Kosong

Selanjutnya dilakukan normalisasi data menggunakan *min-max* dan *z-score*. Perhitungan dilakukan menggunakan Persamaan 1 untuk *min-max* dan Persamaan 2 untuk *z-score*, sebagai contoh akan ditunjukkan salah satu variabel sebelum dan sesudah normalisasi pada Tabel 2.

Tabel 2. Normalisasi Data

Kabupaten/Kota	AHH		
	Data Asli	Min-Max	Z-Score
Pacitan	72,07	0,710562	0,18070
Ponorogo	72,85	0,817558	0,58101
Trenggalek	73,86	0,956104	1,09935
Tulungagung	74,16	0,997257	1,25331
Blitar	73,61	0,921811	0,97105

2. Clustering dengan Metode *Self-Organizing Maps* (SOM)

Pada algoritma SOM, ditentukan terlebih dahulu jumlah kluster, *learning rate*, dan jumlah iterasi yang akan digunakan pada penelitian. Pada penelitian ini digunakan jumlah kluster adalah 2 sampai 7 dan *learning rate* adalah 0,1 sampai 0,8. Penentuan jumlah iterasi dilakukan dengan uji coba (*trial & error*) berdasarkan nilai DBI yang sudah stabil atau tidak berubah pada jumlah iterasi tertentu seperti pada Tabel 3.

Tabel 3. DBI dari Beberapa *Trial & Error Clustering* dengan SOM

α	Jumlah Iterasi	Jumlah Kluster					
		2	3	4	5	6	7
0,1	10	0,84822	0,74823	0,77890	0,94395	0,89386	0,79093
	50	0,84822	0,74823	0,77890	0,94395	0,89386	0,79093
	100	0,84822	0,74823	0,77890	0,94395	0,89386	0,79093
0,2	10	0,92707	0,81633	0,80513	0,85316	0,91563	0,82370
	50	0,92707	0,81633	0,80513	0,85316	0,91563	0,82370
	100	0,92707	0,81633	0,80513	0,85316	0,91563	0,82370
0,3	10	0,92707	0,83193	0,82370	0,83193	0,90379	0,81633
	50	0,92707	0,83193	0,82370	0,83193	0,90379	0,81633
	100	0,92707	0,83193	0,82370	0,83193	0,90379	0,81633
0,4	10	0,92707	0,82370	0,81633	0,82436	0,91623	0,81633
	50	0,92707	0,82370	0,81633	0,82436	0,91623	0,81633
	100	0,92707	0,82370	0,81633	0,82436	0,91623	0,81633
0,5	10	0,77309	0,80513	0,80513	0,82436	0,91623	0,82436
	50	0,77309	0,80513	0,80513	0,82436	0,91623	0,82436
	100	0,77309	0,80513	0,80513	0,82436	0,91623	0,82436
0,6	10	0,77309	0,82436	0,82370	0,82436	0,97579	0,82436
	50	0,77309	0,82436	0,82370	0,82436	0,97579	0,82436
	100	0,77309	0,82436	0,82370	0,82436	0,97579	0,82436
0,7	10	0,77309	0,80513	0,82436	0,82436	0,97579	0,82436
	50	0,77309	0,80513	0,82436	0,82436	0,97579	0,82436
	100	0,77309	0,80513	0,82436	0,82436	0,97579	0,82436
0,8	10	0,77309	0,80513	0,82436	0,82436	0,97579	0,82436
	50	0,77309	0,80513	0,82436	0,82436	0,97579	0,82436

100	0,77309	0,80513	0,82436	0,82436	0,97579	0,82436
-----	---------	---------	---------	---------	---------	---------

Pada Tabel 3 terlihat bahwa pada iterasi ke-50 dan iterasi ke-100 nilai DBI memiliki kesamaan nilai DBI pada iterasi ke-10, yang berarti nilai DBI ini sudah stabil pada iterasi ke-10. Oleh karena itu pada penelitian ini jumlah iterasi yang digunakan pada metode SOM sebanyak 10 iterasi.

Pada metode SOM juga perlu dilakukan penentuan bobot awal. Nilai acak dapat diberikan untuk bobot awal. Namun, jika informasi tersedia mengenai distribusi kluster yang mungkin sesuai untuk masalah tertentu, bobot awal dapat diambil dari informasi tersebut untuk mencerminkan pengetahuan sebelumnya (Fausett, 1994:172).

3. *Clustering* dengan Metode *Fuzzy C-Means* (FCM)

Pada algoritma FCM, ditentukan terlebih dahulu jumlah kluster, bobot pangkat, dan jumlah iterasi yang akan digunakan pada penelitian. Pada penelitian ini digunakan jumlah kluster adalah 2 sampai 7 dan bobot pangkat adalah 1,5 sampai 5. Penentuan jumlah iterasi dilakukan dengan uji coba berdasarkan nilai DBI yang sudah stabil atau tidak berubah pada jumlah iterasi tertentu seperti pada Tabel 4.

Tabel 4. DBI dari Beberapa *Trial & Error Clustering* dengan FCM

w	Jumlah Iterasi	Jumlah Kluster					
		2	3	4	5	6	7
1,5	10	0,90036	0,83193	0,97579	1,04617	1,11945	0,95227
	50	0,84822	0,83193	1,02040	1,07894	1,04171	0,96914
	100	0,84822	0,83193	1,02040	1,07894	1,04171	0,96914
2	10	0,91428	0,83193	0,97579	1,13262	1,16254	0,98797
	50	0,91428	0,83193	0,99488	1,01639	1,08531	0,98110
	100	0,91428	0,83193	0,99488	1,01639	1,08531	0,98110
2,5	10	0,92707	0,85316	1,01788	1,04052	1,15129	1,01652
	50	0,92707	0,85315	0,99488	1,04052	1,19925	0,99063
	100	0,92707	0,85315	0,99488	1,04052	1,19925	0,99063
3	10	0,93588	0,85995	1,02211	1,09564	1,21480	1,19102
	50	0,93588	0,85995	1,02674	1,05555	1,15129	0,98441
	100	0,93588	0,85995	1,02674	1,05555	1,15129	0,98441
3,5	10	0,93588	0,85995	1,12138	1,12138	1,03499	1,19578
	50	0,93588	0,85995	1,02674	0,98686	1,01494	1,01541
	100	0,93588	0,85995	1,02674	0,98686	1,01494	1,01541
4	10	0,93588	0,85995	1,12138	1,22331	1,24314	1,05244
	50	0,93588	0,85995	1,10543	0,98686	0,95403	0,99290
	100	0,93588	0,85995	1,10543	0,98686	0,95403	0,99290
4,5	10	0,93588	0,89125	1,12138	1,21393	1,21794	1,17849
	50	0,93588	0,85995	1,10543	0,99008	0,92777	0,99290
	100	0,93588	0,85995	1,10543	0,99008	0,92777	0,99290
5	10	0,93588	0,89125	1,13056	1,37466	1,25185	0,98439
	50	0,93588	0,85995	1,10543	0,99008	0,92777	0,93747
	100	0,93588	0,85995	1,10543	0,99008	0,92777	0,93747

Pada Tabel 4 terlihat bahwa pada iterasi ke-50 nilai DBI memiliki kesamaan nilai DBI pada iterasi ke-100, namun beberapa nilai pada iterasi ke-50 nilai DBI memiliki perbedaan dengan nilai DBI pada iterasi ke-10 seperti pada 4 kluster hingga 7 kluster dengan bobot pangkat 1,5 dan lainnya, yang berarti nilai DBI ini sudah stabil pada iterasi ke-50. Oleh karena itu pada penelitian ini jumlah iterasi maksimum yang digunakan pada metode *fuzzy c-means* sebanyak 50 iterasi.

4. Validasi Kluster

Validasi kluster bertujuan untuk menentukan jumlah kluster optimal dalam analisis *clustering*. metode validasi kluster yang digunakan adalah *davies-bouldin index* (DBI) dan *calinski-harabasz index* (CHI). Hasil DBI yang diperoleh dari hasil *clustering* metode SOM dan FCM pada Tabel 5.

Tabel 5. DBI Hasil *Clustering* Metode SOM dan FCM

		SOM							
Normalisasi Data	$k \setminus \alpha$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8
<i>Min-max</i>	2	0,84822	0,92707	0,92707	0,92707	0,77309	0,77309	0,77309	0,77309
	3	0,74823	0,81633	0,83193	0,82370	0,80513	0,82436	0,80513	0,80513
	4	0,77890	0,80513	0,82370	0,81633	0,80513	0,82370	0,82436	0,82436
	5	0,94395	0,85316	0,83193	0,82436	0,82436	0,82436	0,82436	0,82436
	6	0,89386	0,91563	0,90379	0,91623	0,91623	0,97579	0,97579	0,97579
	7	0,79093	0,82370	0,81633	0,81633	0,82436	0,82436	0,82436	0,82436
	<i>Z-score</i>	2	0,90848	0,83279	0,79069	0,79069	0,75653	0,75653	0,75653
3		0,95225	0,89591	0,85355	0,85355	0,85355	0,85355	0,85355	0,85355
4		0,99725	0,97481	0,96992	0,96992	0,96992	0,88786	0,88786	0,88786
5		1,29299	1,23366	1,00096	1,00096	1,05047	1,03435	1,03435	0,99392
6		1,12851	0,97006	1,07004	0,99721	0,99721	0,99721	1,03328	1,01491
7		1,11441	0,86289	0,87526	0,90634	0,95026	1,00662	1,00662	1,05936
		FCM							
Normalisasi Data	$k \setminus w$	1,5	2	2,5	3	3,5	4	4,5	5
<i>Min-max</i>	2	0,84822	0,91428	0,92707	0,93588	0,93588	0,93588	0,93588	0,93588
	3	0,83193	0,83193	0,85316	0,85995	0,85995	0,85995	0,85995	0,85995
	4	1,02040	0,99488	0,99488	1,02674	1,02674	1,10543	1,10543	1,10543
	5	1,07894	1,01639	1,04052	1,05555	0,98686	0,98686	0,99008	0,99008
	6	1,04171	1,08531	1,19925	1,15129	1,01494	0,95493	0,92777	0,92777
	7	0,96914	0,98110	0,99063	0,98441	1,01541	0,99290	0,99290	0,93747
	<i>Z-score</i>	2	0,79069	0,83279	0,90848	0,90848	0,90848	0,90848	0,90848
3		0,88412	0,88412	0,88412	0,88412	0,89206	0,91830	0,91830	0,91830
4		1,01282	0,98323	0,98323	0,98323	0,98323	1,04846	1,08541	1,08942
5		0,98152	0,99250	1,03763	1,07637	1,05457	1,00747	1,00977	1,00977
6		1,06476	1,17869	1,13248	0,99517	0,99517	0,99517	0,93921	0,93921
7		0,97332	0,96311	1,02913	0,94426	0,96656	0,92630	1,27185	1,03970

Pada Tabel 5 terlihat bahwa indeks terkecil DBI adalah 0,74823, yaitu metode SOM dengan jumlah kluster 3, *learning rate* 0,1 dan normalisasi *min-max*. Maka dari itu terpilih bahwa hasil *clustering* terbaik berdasarkan indeks terkecil DBI adalah metode SOM dengan jumlah kluster 3, *learning rate* 0,1 dan normalisasi *min-max*.

Hasil CHI yang diperoleh dari hasil *clustering* metode SOM dan FCM pada Tabel 6.

Tabel 6. CHI Hasil *Clustering* Metode SOM dan FCM

		SOM							
Normalisasi Data	$k \setminus \alpha$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8
<i>Min-max</i>	2	35,5160	35,0068	35,0068	35,0068	36,2894	36,2894	36,2894	36,2894
	3	34,2814	43,4751	43,9344	43,4270	42,4131	42,9225	42,4131	42,4131
	4	35,9644	42,4131	43,4270	43,4751	42,4131	43,4270	42,9225	42,9225
	5	30,6402	43,9041	43,9344	42,9225	42,9225	42,9225	42,9225	42,9225
	6	26,0184	33,5482	38,0533	35,8320	35,8320	38,5104	38,5104	38,5104

	7	30,1217	43,4270	43,4751	43,4751	42,9225	42,9225	42,9225	42,9225
<i>Z-score</i>	2	36,8969	37,9078	39,1598	39,1598	39,2380	39,2380	39,2380	39,2380
	3	37,3222	41,0275	41,9120	41,9120	41,9120	41,9120	41,9120	41,9120
	4	31,4534	37,0088	38,0740	38,0740	38,0740	35,9640	35,9640	35,9640
	5	23,3278	27,3695	31,8665	31,8665	31,9828	33,6042	33,6042	34,0755
	6	19,3592	24,3971	27,7033	30,3412	30,3412	30,3412	31,3548	33,5478
	7	17,0309	21,6247	22,6868	24,2078	26,4642	28,6122	28,6122	31,0256
	FCM								
Normalisasi Data	$k \setminus w$	1,5	2	2,5	3	3,5	4	4,5	5
<i>Min-max</i>	2	35,5160	35,1885	35,0068	34,8520	34,8520	34,8520	34,8520	34,8520
	3	43,9344	43,9344	43,9041	43,5151	43,5151	43,5151	43,5151	43,5151
	4	38,4828	37,7927	37,7927	38,1098	38,1098	35,6321	35,7084	35,7084
	5	35,5152	34,0838	32,1110	31,3222	31,6321	31,6321	30,1257	30,1257
	6	34,3449	31,5350	30,3744	29,5721	30,3508	29,9657	29,2210	29,2210
	7	32,1690	30,5302	30,7303	29,3819	28,8264	28,0229	28,0229	24,4082
	<i>Z-score</i>	2	39,1598	37,9078	36,8969	36,8969	36,8969	36,8969	36,8969
3		43,0433	43,0433	43,0433	43,0433	42,7451	41,8075	41,8075	41,8075
4		38,0449	37,5255	37,5255	37,5255	37,5255	35,1264	35,2935	34,5518
5		34,9803	34,0128	31,7391	31,8614	31,6804	31,6620	29,3573	29,3573
6		32,8622	29,4188	29,9572	30,0920	30,0920	30,0920	28,8443	28,8443
7		30,8847	30,0125	29,8055	29,5709	29,3188	28,4516	24,6882	24,5074

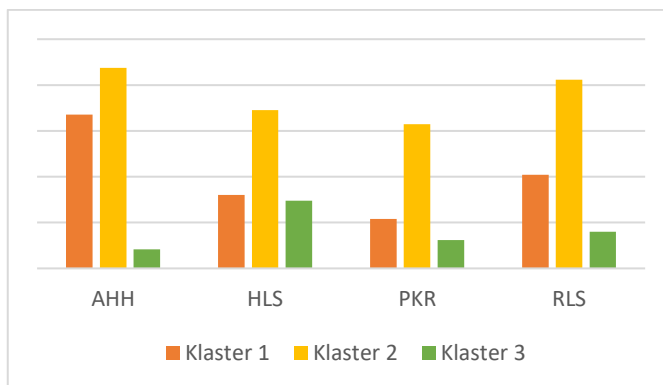
Pada Tabel 6 terlihat bahwa indeks terbesar CHI adalah 433,9344, yaitu metode SOM dengan jumlah kluster 3, *learning rate* 0,1, dan normalisasi *min-max* dan metode FCM dengan jumlah kluster 3, bobot pangkat 1,5, dan normalisasi *min-max*. Kedua metode analisis clustering tersebut memiliki hasil yang sama, sehingga dipilih salah satu yaitu metode FCM. Maka dari itu terpilih bahwa hasil *clustering* terbaik berdasarkan indeks terbesar CHI adalah metode FCM dengan jumlah kluster 3, bobot pangkat 1,5, dan normalisasi *min-max*.

5. Analisis Karakteristik

Analisis karakteristik dilakukan pada hasil clustering terbaik berdasarkan validasi kluster, yaitu pada DBI diperoleh metode SOM dengan jumlah kluster 3, *learning rate* 0,1 dan normalisasi *min-max* dan pada CHI diperoleh metode FCM dengan jumlah kluster 3, bobot pangkat 1,5, dan normalisasi *min-max*. Untuk mengetahui karakteristik dari setiap kluster digunakan nilai rata-rata dari setiap kluster. Nilai rata-rata dari setiap kluster metode SOM berdasarkan DBI terkecil pada Tabel 7 dan divisualisasikan dalam diagram batang pada Gambar 2.

Tabel 7. Nilai Rata-rata Variabel Hasil Clustering SOM

Kluster	AHH	HLS	PKR	RLS
1	71,777	13,017	10.660,292	7,522
2	73,264	14,508	14.450,4	10,217
3	67,498	12,92	9.813,25	5,905



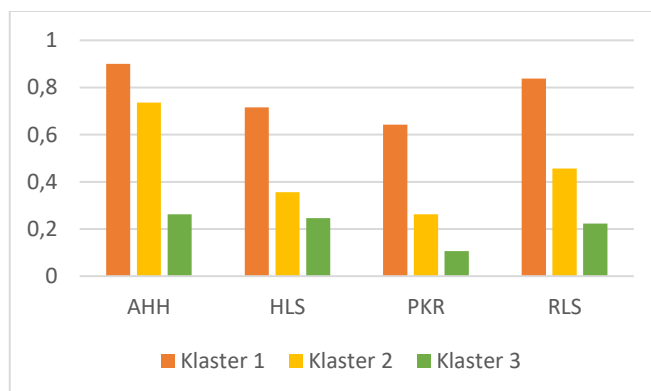
Gambar 2. Nilai Rata-rata Variabel Hasil *Clustering* SOM

Berdasarkan Tabel 7 dan Gambar 2, terlihat bahwa klaster 1 dikategorikan sedang, klaster 2 dikategorikan tinggi, dan klaster 3 dikategorikan rendah. Adapun klaster 1 sebanyak 24 anggota, klaster 2 sebanyak 10 anggota, dan klaster 3 sebanyak 4 anggota.

Nilai rata-rata dari setiap klaster metode FCM berdasarkan CHI terbesar pada Tabel 8 dan divisualisasikan dalam diagram batang pada Gambar 3.

Tabel 8. Nilai Rata-rata Variabel Hasil *Clustering* FCM

Klaster	AHH	HLS	PKR	RLS
1	73,449	14,606	14.572,222	10,316
2	72,252	13,163	11.081,1	7,835
3	68,799	12,718	9.648	6,308



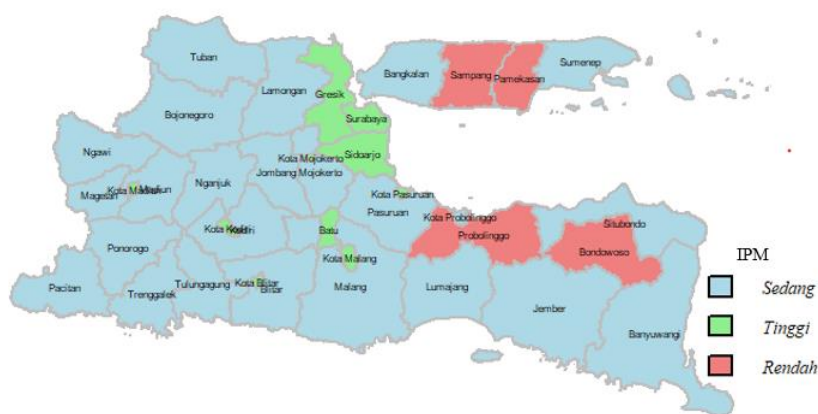
Gambar 3. Nilai Rata-rata Variabel Hasil *Clustering* FCM

Berdasarkan Tabel 8 dan Gambar 3, terlihat bahwa klaster 1 dikategorikan tinggi, klaster 2 dikategorikan sedang, dan klaster 3 dikategorikan rendah. Adapun klaster 1 sebanyak 9 anggota, klaster 2 sebanyak 20 anggota, dan klaster 3 sebanyak 9 anggota.

Pembahasan

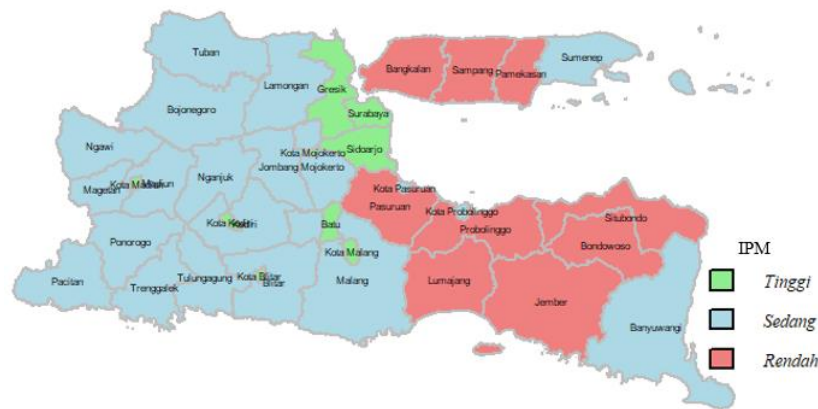
Hasil validasi klaster terbaik berdasarkan DBI diperoleh metode SOM dengan jumlah klaster 3, *learning rate* 0,1, dan normalisasi min-max. Pada klaster 1 dikategorikan indeks pembangunan manusia sedang karena klaster 1 merupakan klaster yang terdiri dari Angka Harapan Hidup (AHH) sedang, Harapan Lama Sekolah (HLS) sedang, Pendapatan Riil per Kapital (PKR) sedang, dan Rata-rata Lama Sekolah (RLS) sedang. Pada klaster 2 dikategorikan indeks pembangunan manusia tinggi karena klaster 2 merupakan klaster yang terdiri dari AHH tinggi, HLS tinggi, PKR tinggi, dan RLS tinggi. Pada Klaster 3 dikategorikan indeks pembangunan manusia rendah karena klaster 3 merupakan klaster yang terdiri dari AHH rendah, HLS rendah, PKR rendah, dan RLS rendah.

Anggota dari setiap kluster hasil *clustering* metode SOM divisualisasikan pada Gambar 4.



Gambar 4. Peta Daerah Hasil *Clustering* Metode SOM

Hasil validasi kluster terbaik berdasarkan CHI diperoleh metode FCM dengan jumlah kluster 3, bobot pangkat 1,5, dan normalisasi min-max. Pada kluster 1 dikategorikan indeks pembangunan manusia tinggi karena kluster 1 merupakan kluster yang terdiri dari Angka Harapan Hidup (AHH) tinggi, Harapan Lama Sekolah (HLS) tinggi, Pendapatan Riil per Kapital (PKR) tinggi, dan Rata-rata Lama Sekolah (RLS) tinggi. Pada kluster 2 dikategorikan indeks pembangunan manusia sedang karena kluster 2 merupakan kluster yang terdiri dari AHH sedang, HLS sedang, PKR sedang, dan RLS sedang. Pada Kluster 3 dikategorikan indeks pembangunan manusia rendah karena kluster 3 merupakan kluster yang terdiri dari AHH rendah, HLS rendah, PKR rendah, dan RLS rendah. Anggota dari setiap kluster hasil *clustering* metode FCM divisualisasikan pada Gambar 5.



Gambar 5. Peta Daerah Hasil *Clustering* Metode FCM

Berdasarkan hasil *clustering* metode SOM dan FCM, terdapat perbedaan dan persamaan hasil *clustering* kedua metode tersebut. Perbedaan dari kedua hasil *clustering* tersebut terletak pada anggota kluster, yang mana beberapa anggota kluster berada pada kluster yang berbeda. Terlihat pada Gambar 4 dan Gambar 5, terdapat 6 Kabupaten/Kota yang berada pada kluster yang berbeda, yaitu terdiri dari Lumajang, Jember, Situbondo, Pasuruan, bangkalan, dan Kota Pasuruan. Persamaan dari kedua hasil *clustering* tersebut adalah data Indikator IPM Kabupaten/Kota Jawa Timur dapat diklasterisasi sebanyak 3 kluster dengan normalisasi *min-max*.

Pada penelitian tersebut terdapat kluster dengan kategori indeks pembangunan manusia rendah. indeks pembangunan manusia rendah menggambarkan masyarakat yang bermukim di daerah dengan indeks pembangunan yang rendah. Oleh karena itu, wilayah pada kluster dengan kategori rendah tersebut membutuhkan perhatian lebih dari pemerintah dalam perencanaan sasaran program agar sesuai dengan kebutuhan setiap daerah, sehingga hal ini dapat meningkatkan indeks pembangunan manusia secara merata.

SIMPULAN

Metode SOM dan FCM dapat diterapkan untuk pengelompokan Kabupaten/kota Jawa Timur berdasarkan indikator IPM dengan 38 Kabupaten/Kota dan 4 indikator IPM. Hasil clustering terbaik ditentukan menggunakan dua metode validasi kluster, yaitu DBI dan CHI. Berdasarkan DBI terbaik diperoleh metode SOM dengan jumlah kluster 3, learning rate 0,1, dan normalisasi min-max. berdasarkan CHI terbaik diperoleh metode FCM dengan jumlah kluster 3, bobot pangkat 1,5, dan normalisasi min-max.

Berdasarkan kedua hasil clustering tersebut, terdapat perbedaan dan persamaan hasil clustering. Perbedaan dari kedua hasil *clustering* tersebut terletak pada anggota kluster, yang mana beberapa anggota kluster berada pada kluster yang berbeda. Persamaan dari kedua hasil *clustering* tersebut adalah data Indikator IPM Kabupaten/Kota Jawa Timur dapat diklusterisasi sebanyak 3 kluster dengan normalisasi *min-max*.

penelitian ini hanya menganalisis satu wilayah provinsi untuk analisis *clustering*. Oleh karena itu, bagi peneliti yang nantinya akan melakukan analisis *clustering* menggunakan indikator indeks pembangunan manusia dapat menambahkan wilayah provinsi lainnya yang ada di Indonesia. Diharapkan hasil penelitian ini dapat menjadi sumber informasi dan referensi bagi pemerintah dalam membuat kebijakan dan perencanaan program yang tepat sasaran.

UCAPAN TERIMA KASIH

Terimakasih kepada Koordinator Prodi Matematika dan seluruh jajaran Dosen Prodi Matematika yang telah membagikan ilmu serta memberikan bimbingan hingga artikel ini dapat diselesaikan.

DAFTAR PUSTAKA

- Aliyatussya'ni, & Saputro, D. R. S. (2022). ALGORITME PARTITIONING AROUND MEDOID (PAM) DENGAN CALINSKI-HARABASZ INDEX UNTUK CLUSTERING DATA OUTLIER. *UNEJ e-Proceeding*, 22-29.
- Badan Pusat Statistik (BPS). (2021). Indeks Pembangunan Manusia 2021. Jakarta, Indonesia: Badan Pusat Statistik.
- Badan Pusat Statistik (BPS) Provinsi Jawa Timur. (2021). Tabel Dinamis Indeks Pembangunan Manusia. Jawa Timur, Indonesia: Badan Pusat Statistik Provinsi Jawa Timur.
- Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press.
- Cahyanti, D. N., Muchtolifah, & Sishadiyati. (2021). Faktor-faktor Indeks Pembangunan Manusia Di Provinsi Jawa Timur. *Jambura Economic Education Journal*, 3(2), 93-101.
- Deana, R., Suhaedi, D., & Harahap, E. (2022). Konstruksi Sistem Inferensi Fuzzy Menggunakan Subtractive Fuzzy C-Means pada Data Parkinson. *In Bandung Conference Series: Mathematics*, 2(1), 51-58.
- Deni, Midyanti, D. M., & Hidayati, R. (2022). Penentuan Pencemaran Air Menggunakan Metode Self Organizing Maps (SOM). *Indonesian Journal of Computer Science*, 11(1), 128-137.
- Dhanachandra, N., & Chanu, Y. J. (2020). An image segmentation approach based on fuzzy c-means and dynamic particle swarm optimization algorithm. *Multimedia tools and applications*, 79(25-26), 18839-18858.
- Faradonbeh, R. S., Haghshenas, S. S., Taheri, A., & Mikaeil, R. (2020). Application of self-organizing map and fuzzy c-mean techniques for rockburst clustering in deep underground projects. *Neural Computing and Applications*, 32, 8545-8559.

- Herlinda, V., Darwis, D., & Dartono. (2021). Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means. *Jurnal Teknologi dan Sistem Informasi*, 2(2), 94-99.
- Khairati, A. F., Adlina, A. A., Hertono, G. F., & Handari, B. D. (2019, February). Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan K-Means MMCA. In *PRISMA, Prosiding Seminar Nasional Matematika* (Vol. 2, pp. 161-170).
- Nawaratne, R., Alahakoon, D., De Silva, D., & Yu, X. (2019, July). HT-GSOM: Dynamic self-organizing map with transience for human activity recognition. In *2019 IEEE 17th international conference on industrial informatics (INDIN)*, Vol. 1, 270-273.
- Nugroho, C. A., Hendrawan, R. A., & Hafidz, I. (2012). Clustering kelompok swadaya masyarakat (KSM) dalam menentukan kebijakan bantuan badan pemberdayaan masyarakat di kota surabaya dengan menggunakan metode self-organizing map (SOM) dan K-means. *Jurnal Teknik ITS*, 1(1), A368-A373.
- Poerwanto, B., & Ali, B. (2019). Implementasi Algoritma Fuzzy C-means dalam Mengelompokkan Kecamatan di Tana Luwu Berdasarkan Produktifitas Hasil Perkebunan. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 19(1), 163-172.
- Sitompul, B. J. D., Sitompul, O. S., & Sihombing, P. (2019). Enhancement clustering evaluation result of davies-bouldin index with determining initial centroid of k-means algorithm. *Journal of Physics: Conference Series*, 1235(1).
- United Nations Development Programme (UNDP). (2022). *Human Development Report 2021-22: Uncertain Times, Unsettled Lives: Shapping our Future in a Transforming World*. New York.
- Wang, X., & Xu, Y. (2019, July). An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index. In *IOP Conference Series: Materials Science and Engineering* (Vol. 569, No. 5, p. 052024). IOP Publishing.