

---

---

## PENGELOMPOKAN KABUPATEN/KOTA DI JAWA TIMUR BERDASARKAN KASUS *STUNTING* BALITA MENGGUNAKAN ALGORITME *FUZZY PARTICLE SWARM OPTIMIZATION-FUZZY C-MEANS*

Sepnita Wulandari<sup>1</sup>, Robert Kurniawan<sup>2</sup>

<sup>1,2</sup>Jurusan Komputasi Statistik, Sekolah Tinggi Ilmu Statistik (STIS)

e-mail: <sup>1</sup>sepnitawulandari75@gmail.com, <sup>2</sup>robertk@stis.ac.id

### ABSTRACT

*Stunting is a condition that describe the presence of chronic malnutrition problem caused by various condition. East Java Province is a region that has the highest percentage of short toddler in Java Island. Moreover, there is high disparity in cross regency/city and the prevalence rate of stunting in the East Java Province is same as national prevalence rate. Meanwhile, Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2015-2019 sets the target of national prevalence rate of stunting toddler decreasing in 2019. Based on that problem, this research is clustering regency/city in East Java Province based on stunting toddler case. The clustering uses Fuzzy Particle Swarm Optimization-Fuzzy C-Means (FPSO-FCM). From the clustering result, this research obtains 2 cluster which are cluster of low stunting potential region (cluster 1) and high stunting potential region (cluster 2).*

**Keywords:** *stunting, FPSO-FCM, Jawa Timur, clustering.*

### PENDAHULUAN

*Stunting* merupakan suatu keadaan yang menggambarkan adanya masalah gizi kronis yang dipengaruhi oleh beberapa kondisi [14]. Data “Profil Kesehatan Provinsi Jawa Timur Tahun 2015” menunjukkan bahwa kecenderungan prevalensi status gizi balita di Jawa Timur mengalami penurunan (perbaikan) dari tahun 2010 hingga tahun 2015. Sementara itu, menurut data Kementerian Kesehatan Republik Indonesia, persentase balita pendek di Jawa Timur pada tahun 2015 adalah sebesar 27,1%. Persentase tersebut menempatkan Provinsi Jawa Timur menjadi wilayah dengan persentase balita pendek tertinggi di Pulau Jawa. Selain itu, disparitas yang sangat tinggi terjadi antar kabupaten/kota di Provinsi Jawa Timur dengan angka prevalensi *stunting* balita

pada tingkatan yang sama dengan angka nasional, yaitu prevalensi *stunting* balita tinggi sebesar 30-39% [9]. Di sisi lain, Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2015-2019 menetapkan target prevalensi *stunting* balita menurun dari angka 32,9% pada tahun 2013 menjadi sebesar 28% pada tahun 2019 [13]. Berdasarkan hal tersebut, penelitian ini melakukan pengelompokan di Provinsi Jawa Timur untuk melihat apakah kebijakan dan upaya yang dilakukan telah diterapkan pada daerah yang sesuai sehingga kebijakan untuk mengatasi kondisi *stunting* ini lebih tepat sasaran. Selain itu, hasil dari pengelompokan di Provinsi Jawa Timur dapat dijadikan rujukan bagi pemerintah untuk mengatasi prevalensi *stunting* balita pada tingkat nasional.

Pengelompokan kabupaten/kota di Provinsi Jawa Timur berdasarkan kasus

data *stunting* balita dalam penelitian ini menggunakan algoritme *Fuzzy Particle Swarm Optimization-Fuzzy C-Means* (FPSO-FCM). Algoritme ini digunakan sebagai solusi kelemahan algoritme *Fuzzy C-Means* (FCM) dan algoritme *Fuzzy Particle Swarm Optimization* (FPSO). Hal ini bertujuan untuk meningkatkan performa hasil pengelompokan (*clustering*). Selain itu, hasil pengelompokan akan dianalisis menggunakan tujuh indeks validitas (PC, MPC, CE, FS, XB, PI, dan Kwon) dan diuji secara statistik untuk menentukan variabel yang berperan sebagai variabel pembeda antar-*cluster*.

*Stunting* (balita pendek) adalah balita dengan status gizi yang berdasarkan panjang atau tinggi badan menurut umurnya bila dibandingkan dengan standar baku WHO-MGRS (*Multicentre Growth Reference Study*) tahun 2005 [14], nilai *z-score* balita tersebut kurang dari -2SD dan masuk kategori sangat pendek jika nilai *z-score* balita tersebut kurang dari -3SD. Seorang balita dapat dikatakan mengalami kondisi *stunting* (balita pendek) apabila hasil perbandingan antara tinggi badan dan nilai standar balita tersebut berada dibawah normal. Masalah *stunting* ini menggambarkan adanya masalah gizi kronis, dipengaruhi dari kondisi ibu/calon ibu, masa janin, dan masa bayi/balita, termasuk penyakit yang diderita selama masa balita. Selain itu, masalah *stunting* juga dipengaruhi kondisi-kondisi lain yang secara tidak langsung memengaruhi kesehatan.

*Clustering* merupakan teknik *data mining* yang bertujuan untuk menetapkan objek-objek menjadi sekumpulan kelompok terpisah yang disebut dengan *cluster* sehingga setiap objek yang menjadi anggota *cluster* yang sama akan memiliki karakteristik yang lebih mirip satu sama lain dibandingkan dengan objek dari *cluster* berbeda [7].

Terdapat dua jenis teknik *clustering*, yaitu *Hard Clustering* dan *Fuzzy Clustering* [8]. Pada *hard clustering*, setiap elemen data harus menjadi anggota atau tidak menjadi anggota dari suatu *cluster*. Sementara itu pada *fuzzy clustering*, setiap elemen data dapat menjadi anggota dari beberapa *cluster* dengan nilai keanggotaan yang berbeda-beda.

*Fuzzy Particle Swarm Optimization-Fuzzy C-Means* (FPSO-FCM) adalah algoritme yang dikembangkan oleh Hesam Izakian dan Ajith Abraham pada tahun 2011 [12]. Algoritme ini merupakan algoritme FCM yang diintegrasikan dengan algoritme FPSO untuk membentuk suatu algoritme *hybrid* di dalam teknik *clustering*. FPSO-FCM dikembangkan untuk menjadi solusi kelemahan yang ada pada FCM dan FPSO.

Algoritme FCM mudah terperangkap ke dalam *local optima*. Hal ini dapat diatasi menggunakan algoritme FPSO. Namun, algoritme FPSO memiliki performa yang lebih lambat karena FPSO membutuhkan evaluasi fungsi yang lebih banyak dibandingkan FCM. Algoritme FPSO-FCM menerapkan algoritme FCM ke dalam partikel di dalam kawanan pada setiap iterasi sehingga nilai *fitness* setiap partikel mengalami perbaikan dengan waktu yang lebih cepat.

Indeks validitas digunakan untuk mengevaluasi hasil *clustering* dan performa suatu metode *clustering*. Penelitian ini menggunakan tujuh indeks validitas, yaitu *Partition Coefficient* (PC), *Classification Entropy* (CE), *Modified Partition Coefficient* (MPC), *Partition Index* (PI), *Fukuyama Sugeno* (FS), *Xie-Beni* (XB), dan *Kwon Index*. Indeks validitas tersebut sama dengan yang digunakan pada penelitian yang dilakukan oleh [11], kecuali untuk *Kwon*

*Index*. Ketujuh indeks tersebut adalah sebagai berikut :

1. *Partition Coefficient (PC)*

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

dengan  $k = 1, 2, \dots, c$ ,  $i = 1, 2, \dots, n$ ,  $c$  merupakan jumlah *cluster*,  $n$  merupakan jumlah objek,  $\mu_{ik}$  merupakan derajat keanggotaan objek  $x_i$  pada *cluster* ke- $k$  yang dapat dihitung menggunakan persamaan (16), dan  $x_i$  merupakan baris data ke- $i$ . Rentang nilai indeks PC adalah  $\frac{1}{c} \leq PC \leq 1$ . Nilai PC yang paling besar menunjukkan hasil clustering yang paling optimal.

2. *Classification Entropy (CE)*

$$CE = -\frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{ik} \ln(\mu_{ik}) \quad (2)$$

Rentang nilai indeks CE adalah  $0 \leq 1 - PC \leq CE \leq \ln(c)$ . Nilai CE yang semakin kecil menunjukkan hasil clustering semakin baik.

3. *Modified Partition Coefficient (MPC)*

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

Rentang nilai indeks MPC adalah  $0 \leq MPC \leq 1$ . Nilai MPC paling besar menunjukkan hasil clustering yang paling optimal.

4. *Partition Index (PI)*

$$PI = \sum_{k=1}^c \frac{\sum_{i=1}^n (\mu_{ik})^w \|x_i - a_k\|^2}{n_k \sum_{l=1}^c \|a_l - a_k\|^2} \quad (4)$$

dengan  $w$  merupakan tingkat kekaburan (*fuzzifier*),  $\|x_i - a_k\|$  merupakan jarak Euclidean titik data  $x_i$  dengan pusat *cluster*  $a_k$  dan  $\|a_l - a_k\|$  merupakan jarak Euclidean antar pusat *cluster*. Nilai *PI* yang semakin kecil menunjukkan partisi *cluster* yang lebih baik.

5. *Fukuyama Sugeno (FS)*

$$FS = J_m(\mu, a) - K_m(\mu, a) \\ = \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^w \|x_i -$$

$$a_k\|^2 - \sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^w \|a_k - \bar{a}\|^2 \quad (5)$$

dengan  $\|a_k - \bar{a}\|$  merupakan jarak Euclidean antara pusat *cluster*  $a_k$  dengan rata-rata pusat *cluster* dan  $\bar{a} = \sum_{k=1}^c \frac{a_k}{c}$ .

Nilai *FS* yang semakin kecil menunjukkan partisi *cluster* yang lebih baik.

6. *Xie-Beni (XB)*

$$XB = \frac{J_m(\mu, a) / n}{sep(a)} \\ = \frac{\sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^w \|x_i - a_k\|^2}{n(\min_{k \neq l} (\|a_k - a_l\|^2))} \quad (6)$$

dengan  $\min_{k \neq l} (\|a_k - a_l\|^2)$  merupakan jarak minimum antar pusat *cluster*. Semakin kecil nilai *XB* menunjukkan semakin optimum jumlah *cluster* yang digunakan. (2)

7. *Kwon Index*

$$Kwon = \frac{\sum_{k=1}^c \sum_{i=1}^n \mu_{ik}^2 \|x_k - a_i\|^2 + \frac{1}{c} \sum_{i=1}^c \|a_i - \bar{a}\|^2}{\min_{k \neq l} (\|a_k - a_l\|^2)} \quad (7)$$

**METODE PENELITIAN**

**Sumber Data dan Variabel Penelitian**

Data studi kasus *stunting* balita dalam penelitian ini dibatasi pada wilayah provinsi Jawa Timur menggunakan data *cross section* tahun 2013. Data tersebut merupakan data sekunder yang diperoleh dari publikasi “Indeks Pembangunan Kesehatan Masyarakat 2013” dan “Profil Kesehatan Provinsi Jawa Timur 2013” dari Kementerian Kesehatan. Selain itu, data juga diperoleh dari data mentah Survei Sosial Ekonomi Nasional (Susenas) tahun 2013 yang dilakukan oleh BPS. Data yang dipilih adalah data tahun 2013 dikarenakan variabel-variabel *stunting* yang paling lengkap merupakan data *stunting* pada tahun 2013.

Variabel-variabel penelitian yang digunakan dalam penelitian ini sama dengan yang digunakan oleh Bambang S

Hadi dalam penelitian berjudul “Pendekatan *Modified Particle Swarm Optimization* dan *Artificial Bee Colony* pada *Fuzzy Geographically Weighted Clustering*” (2017). Terdapat 20 variabel yang digunakan dalam data studi kasus *stunting* balita, yaitu rata-rata jumlah ART (Anggota Rumah Tangga) pada ruta balita ( $X_1$ ), persentase rumah tangga balita yang tidak memiliki sanitasi layak ( $X_2$ ), persentase rumah tangga balita yang tidak memiliki sumber air minum layak ( $X_3$ ), rata-rata lama pemberian ASI balita ( $X_4$ ), persentase ibu balita berpendidikan < SMA ( $X_5$ ), persentase ibu balita bekerja ( $X_6$ ), pengeluaran rata-rata rumah tangga balita ( $X_7$ ), rata-rata konsumsi kalori ( $X_8$ ), prevalensi ISPA pada balita ( $X_9$ ), prevalensi diare pada balita ( $X_{10}$ ), persentase balita yang telah imunisasi lengkap ( $X_{11}$ ), persentase balita gizi buruk dan kurang ( $X_{12}$ ), kecukupan puskesmas ( $X_{13}$ ), kecukupan dokter ( $X_{14}$ ), kecukupan tenaga ahli gizi ( $X_{15}$ ), kecukupan bidan ( $X_{16}$ ), kecukupan posyandu ( $X_{17}$ ), persentase Berat Badan Lahir Rendah atau BBLR ( $X_{18}$ ), cakupan pelayanan kesehatan bayi (umur 29 hari-11 bulan) ( $X_{19}$ ), dan cakupan pelayanan kesehatan balita (umur 12-59 bulan) ( $X_{20}$ ).

### Metode Analisis

Metode analisis yang digunakan adalah algoritme *Fuzzy Particle Swarm Optimization-Fuzzy C-Means* (FPSO-FCM). Algoritme FPSO-FCM pada penelitian ini menggunakan ukuran jarak *Euclidean*. Adapun tahapan analisis dari algoritme FPSO-FCM adalah sebagai berikut:

1. Menentukan parameter awal:
  - a. Data yang akan dikelompokkan dan syarat penghentian algoritme berupa  $it_m$  dan  $cons_{it}$  untuk FPSO,  $it_m$  untuk FCM, serta  $it_m$  dan  $cons_{it}$  untuk FPSO-FCM.

Pada penelitian ini, syarat penghentian algoritme untuk FPSO adalah  $it_m = 100$  kali dan  $cons_{it} = 50$  kali; syarat penghentian algoritme untuk FCM adalah  $it_m = 5$  kali; dan syarat penghentian algoritme untuk FPSO-FCM adalah  $it_m = 10$  kali dan  $cons_{it} = 2$  kali.

- b. Parameter FPSO-FCM yaitu jumlah partikel ( $p$ ), laju belajar ( $c_1, c_2$ ), momen inersia ( $iw$ ), jumlah *cluster* ( $c$ ), dan *fuzzifier* ( $w$ ). Pada penelitian ini, nilai masing-masing partikel yang digunakan adalah  $p = 10$  (jumlah partikel bisa diubah tergantung pada peneliti), laju belajar = 2 ( $c_1$ ) dan 2 ( $c_2$ ),  $iw = 0,9$  (nilai minimum = 0,1), dalam uji coba penelitian nilai  $c$  yang digunakan adalah 2 hingga 5 *cluster*, dan nilai  $w$  yang digunakan dalam uji coba penelitian adalah 1,5 hingga 4 dengan peningkatan sebanyak 0,1 disetiap percobaan.
2. Membangkitkan suatu kawanan sejumlah  $p$  partikel terdiri dari matriks posisi ( $X$ ), matriks kecepatan ( $V$ ), matriks  $Pbest$ , dan matriks  $Gbest$ .  $Pbest$  merupakan posisi terbaik yang pernah dikunjungi oleh suatu partikel. Matriks  $Pbest$  diperoleh dengan membandingkan nilai *fitness* seluruh matriks  $X$  dan mengambil salah satu matriks  $X$  dengan nilai *fitness* paling kecil sebagai matriks  $Pbest$ . Pada iterasi pertama, matriks  $Pbest$  sama dengan matriks  $X$  untuk seluruh partikel.  $Gbest$  merupakan posisi terbaik yang pernah dikunjungi oleh seluruh partikel sejak iterasi pertama.  $Gbest$  merupakan matriks  $Pbest$  dengan nilai *fitness* paling kecil. Seluruh matriks tersebut berukuran  $n \times c$ , dengan  $n$  adalah jumlah data dan  $c$  adalah jumlah cluster.

3. Menginisialisasi matriks kawanan :

- a. Matriks posisi ( $\mathbf{X}$ ) sejumlah  $p$  partikel secara acak menggunakan persamaan (8) dan (9) (Izaki dan Abraham, 2011). Sejumlah partikel tersebut merupakan solusi potensial yang bisa dipilih berdasarkan solusi yang menunjukkan hasil terbaik. Jika jumlah partikel sebanyak 10 ( $p=10$ ), maka matriks posisi ( $\mathbf{X}$ ) yang diinisialisasi ada sebanyak 10 matriks  $\mathbf{X}$ . Adapun setiap matriks  $\mathbf{X}$  ditunjukkan sebagai berikut:

$$\mathbf{X}_{n \times c} = \begin{bmatrix} \mu_{1,1} & \mu_{1,2} & \dots & \mu_{1,c} \\ \mu_{2,1} & \mu_{2,2} & \dots & \mu_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n,1} & \mu_{n,2} & \dots & \mu_{n,c} \end{bmatrix}$$

dengan  $n$  merupakan jumlah objek pada data dan  $c$  merupakan jumlah *cluster*.

$$\mu_{ik} \in [0,1] \quad (8)$$

$$\sum_{i=1}^c \mu_{ik} = 1 \quad (9)$$

dengan ketentuan :

$$i = 1, 2, \dots, n ; k = 1, 2, \dots, c$$

- b. Matriks kecepatan ( $\mathbf{V}$ ) sejumlah  $p$  partikel secara acak dengan ketentuan :

$$\mathbf{V}_{n \times c} = \begin{bmatrix} v_{1,1} & v_{1,2} & \dots & v_{1,c} \\ v_{2,1} & v_{2,2} & \dots & v_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n,1} & v_{n,2} & \dots & v_{n,c} \end{bmatrix}$$

dengan  $n$  merupakan jumlah objek pada data dan  $c$  merupakan jumlah *cluster*. Jika jumlah partikel sebanyak 10 ( $p=10$ ), maka matriks kecepatan ( $\mathbf{V}$ ) yang diinisialisasi ada sebanyak 10 matriks  $\mathbf{V}$ .

Matriks  $\mathbf{V}$  tersebut diinisialisasi secara random dengan menggunakan range  $[-1,1]$ . Pada

penelitian ini, penghitungan angka acak menggunakan fungsi `runif()` pada aplikasi R dengan batas nilai minimum -1 dan batas nilai maksimum 1.

- c. Matriks  $\mathbf{Pbest}$  sejumlah  $p$  partikel. Penjelasan mengenai matriks  $\mathbf{Pbest}$  ada pada langkah 2 dan 4.3.
- d. Matriks  $\mathbf{Gbest}$  untuk kawanan partikel. Penjelasan mengenai matriks  $\mathbf{Gbest}$  ada pada langkah 2 dan 4.4.

4. Menjalankan algoritme FPSO :

- 4.1 Menghitung pusat *cluster* untuk setiap partikel menggunakan persamaan (10) [6].

$$a_k(t) = \frac{\sum_{i=1}^n (\mu_{ik}(t-1))^w x_i}{\sum_{i=1}^n (\mu_{ik}(t-1))^w} \quad (10)$$

dengan  $a_k$  merupakan pusat *cluster* ke- $k$ , dan  $t$  merupakan iterasi.

- 4.2 Menghitung nilai *fitness* untuk setiap partikel menggunakan persamaan (11) [17]. Jika jumlah partikel sebanyak 10, persamaan (11) akan dihitung sebanyak 10 kali berdasarkan nilai parameter masing-masing partikel.

$$J = \sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^w \text{dist}^2(x_i, a_k) \quad (11)$$

$$\text{dist}^2(x_i, a_k) = (x_i - a_k)(x_i - a_k)' \quad (12)$$

dengan ketentuan :

$$i = 1, 2, \dots, n ; k = 1, 2, \dots, c$$

dengan  $J$  merupakan fungsi objektif FCM,  $c$  merupakan jumlah *cluster* pada data,  $n$  merupakan jumlah objek pada data,  $a_k$  merupakan pusat *cluster* ke- $k$ ,  $x_i$  merupakan baris data ke- $i$ ,  $\text{dist}^2(x_i, a_k)$  merupakan jarak Euclidean antara objek  $x_i$  dengan  $a_k$ .

4.3 Mengevaluasi **Pbest** untuk setiap partikel. Matriks **Pbest** diperoleh dengan ketentuan sebagai berikut:

- a. Jika nilai *fitness* partikel lebih kecil daripada nilai *fitness Pbest* partikel iterasi sebelumnya, **Pbest** partikel merupakan matriks **X** partikel saat ini.
- b. Jika nilai *fitness* partikel lebih besar daripada nilai *fitness Pbest* partikel iterasi sebelumnya, **Pbest** partikel merupakan **Pbest** partikel sebelumnya.

4.4 Mengevaluasi **Gbest** untuk kawanan. **Gbest** merupakan matriks posisi partikel (**X**) dengan nilai *fitness* terkecil dari seluruh iterasi. Seluruh matriks **X** pada setiap iterasi dihitung nilai *fitness* masing-masing. Dari penghitungan tersebut, salah satu matriks **X** yang memiliki nilai paling kecil akan menjadi matriks **Gbest**.

4.5 Membarui matriks kecepatan (**V**) untuk setiap partikel menggunakan persamaan (13) [12].

$$\begin{aligned} V_{(t)} = & iw \otimes V_{(t-1)} \oplus \\ & (c_1 r_1) \otimes (Pbest_{(t-1)} - \\ & X_{(t-1)}) \oplus (c_2 r_2) \otimes \\ & (Gbest_{(t-1)} - X_{(t-1)}) \end{aligned} \quad (13)$$

dengan *iw* merupakan momen inersia yang mengalami penurunan sebesar 5% di setiap iterasi,  $c_1$  dan  $c_2$  merupakan konstanta (*learning factor*), serta  $r_1$  dan  $r_2$  yang merupakan nilai random dengan range [0,1] (kedua nilai tersebut diperoleh dari angka acak yang berada pada range 0 sampai 1. Pada penelitian ini, penghitungan

angka acak menggunakan fungsi  $runif\{\}$  pada aplikasi R).

4.6 Membarui matriks posisi (**X**) untuk setiap partikel menggunakan persamaan (14) [12].

$$X_{(t)} = X_{(t-1)} \oplus V_{(t)} \quad (14)$$

Setelah matriks **X** setiap partikel diperbarui, matriks **X** tersebut dinormalisasi sehingga ketentuan (8) dan (9) terpenuhi. Adapun tahap normalisasi adalah sebagai berikut:

1. Mengubah seluruh elemen matriks posisi yang bernilai negatif menjadi nol (0).
2. Jika ditemukan baris dengan seluruh elemen bernilai nol (0), evaluasi baris matriks posisi tersebut menggunakan angka random dengan interval [0,1].
3. Lakukan transformasi matriks menggunakan persamaan berikut [12]:

$$X_{normal} = \begin{bmatrix} \mu_{1,1}/\sum_{k=1}^c \mu_{1,k} & \dots & \mu_{1,c}/\sum_{k=1}^c \mu_{1,k} \\ \vdots & \ddots & \vdots \\ \mu_{n,1}/\sum_{k=1}^c \mu_{n,k} & \dots & \mu_{n,c}/\sum_{k=1}^c \mu_{n,k} \end{bmatrix} \quad (15)$$

4.7 Cek kondisi syarat penghentian algoritme FPSO. Jika belum terpenuhi, ulangi langkah 4.1-4.6.

5. Menjalankan algoritme FCM : (13)

5.1 Menghitung pusat *cluster* untuk setiap partikel menggunakan persamaan (10).

5.2 Menghitung jarak antara data dengan seluruh pusat *cluster* menggunakan persamaan (12).

5.3 Menghitung derajat keanggotaan seluruh objek menggunakan persamaan (16) [12].

$$\mu_{ik} = \frac{1}{\sum_{l=1}^c \left( \frac{\text{dist}^2(x_i, a_k)}{\text{dist}^2(x_i, a_l)} \right)^{\frac{2}{w-1}}} \quad (16)$$

- 5.4 Mengevaluasi **Pbest** untuk setiap partikel. Penjelasan mengenai matriks **Pbest** ada pada langkah 2 dan 4.3.
- 5.5 Mengevaluasi **Gbest** untuk kawan partikel. Penjelasan mengenai matriks **Gbest** ada pada langkah 2 dan 4.4.
- 5.6 Cek kondisi syarat penghentian algoritme FCM. Jika belum terpenuhi, ulangi langkah 5.1-5.5.
6. Cek kondisi syarat penghentian algoritme FPSO-FCM. Jika belum terpenuhi, ulangi langkah 4-5.

Hasil akhir pengelompokan diperoleh dari matriks **Gbest** iterasi terakhir. Penentuan kelompok setiap objek dilihat pada derajat keanggotaan atau  $\mu$  objek yang bersangkutan pada *cluster* tertentu. Misalnya jika objek A memiliki nilai  $\mu$  sebesar 0,3 pada *cluster* 1 dan nilai  $\mu$  sebesar 0,7 pada *cluster* 2. Dengan demikian objek A berada pada *cluster* 2. Begitu juga untuk penentuan kelompok objek yang lainnya.

Penentuan ukuran jarak terbaik (jumlah *cluster* dan *fuzzifier*) dilakukan dengan menghitung dan membandingkan setiap nilai indeks validitas pada setiap hasil uji coba. Seluruh penghitungan nilai indeks validitas tersebut dilakukan seperti pada penjelasan Bab 2.5.

Setelah proses pengelompokan selesai, selanjutnya dilakukan analisis variabel yang menjadi pembeda antar *cluster*. Variabel pembeda antar-*cluster* diperoleh dengan cara menguji asumsi normalitas data setiap atribut menggunakan uji *Shapiro-Wilk*. Jika asumsi normalitas terpenuhi, selanjutnya dilakukan uji kesamaan varians menggunakan uji *Levene*. Apabila

variens antar *cluster* sama, digunakan uji ( $t$  varians sama dan tidak diketahui. (16) Apabila varians antar *cluster* tidak sama, digunakan uji  $t$  varians beda dan tidak diketahui. Sedangkan jika asumsi normalitas tidak terpenuhi, digunakan uji *U Mann-Whitney* untuk sampel independen. Selanjutnya, jika hasil akhir pengujian variabel pembeda (baik menggunakan uji *Levene*, uji  $t$ , maupun uji *U Mann-Whitney*) suatu variabel tolak hipotesis nol, cukup bukti untuk menyatakan bahwa variabel tersebut merupakan variabel yang berbeda antar *cluster*.

## HASIL DAN PEMBAHASAN

### Perbandingan Pada Penelitian Terkait

Untuk menganalisis apakah algoritme FPSO-FCM yang digunakan pada penelitian ini telah berhasil meningkatkan performa hasil *clustering*, maka dilakukan perbandingan antara hasil penelitian ini dengan penelitian terkait yang pernah dilakukan sebelumnya. Penelitian terkait tersebut adalah penelitian yang dilakukan oleh [12].

Data uji coba yang digunakan adalah data Iris. Kemudian hasil nilai *fitness* matriks **Gbest** pengelompokan pada data Iris menggunakan algoritme FPSO-FCM pada penelitian ini dibandingkan dengan hasil yang diperoleh pada penelitian [12] menggunakan algoritme FCM dan FPSO. Hasil perbandingan dapat dilihat pada Tabel 1.

**Tabel 1.** Hasil Nilai *Fitness* Alogritme FCM, FPSO, dan FPSO-FCM Pada Hasil Pengelompokan Data Iris

Hasil penelitian Izakian dan Abraham		Hasil penelitian FPSO-FCM
FCM	FPSO	
(1)	(2)	(3)
67,92	66,26	<b>60,5057</b>

Pada Tabel 1, parameter yang digunakan adalah jumlah *cluster* sebanyak 3, *fuzzifier* sebanyak 2, jumlah partikel sebanyak 10, dan laju belajar ( $c_1, c_2$ ) sebesar 2. Sedangkan algoritme FPSO-FCM yang digunakan pada penelitian ini diimplementasikan menggunakan *software R*.

Berdasarkan perbandingan pada Tabel 1, dapat dilihat bahwa nilai *fitness* yang dihasilkan oleh algoritme FPSO-FCM menunjukkan nilai yang lebih kecil dibandingkan hasil pengelompokan menggunakan algoritme FCM dan FPSO. Dengan demikian, dapat disimpulkan bahwa algoritme FPSO-FCM menghasilkan pengelompokan yang lebih baik pada data Iris dibandingkan algoritme FCM dan FPSO.

### Hasil Pengelompokan Data Studi Kasus

Pengelompokan data studi kasus menggunakan algoritme FPSO-FCM

menghasilkan solusi terbaik berupa jumlah *cluster* ( $c$ ) terbaik sebanyak 2 dan *fuzzifier* ( $w$ ) terbaik sebesar 1,5. Hasil tersebut diperoleh dari pengelompokan uji coba menggunakan FPSO-FCM sebanyak lima kali untuk setiap nilai *fuzzifier* (dengan penambahan nilai *fuzzifier* sebesar 0,1, nilai minimum *fuzzifier* sebesar 1,5 dan nilai maksimum *fuzzifier* sebesar 4) dan setiap jumlah *cluster* (dengan minimum jumlah *cluster* sebanyak 2 dan maksimum jumlah *cluster* sebanyak 5). Solusi terbaik yang dipilih dari ketujuh indeks validitas dapat dilihat pada Tabel 2. Sedangkan untuk parameter  $p$  yang digunakan adalah sebesar 10,  $iw$  sebesar 0,9 (nilai minimum = 0,1), serta  $c_1$  sebesar 2 dan  $c_2$  sebesar 2. Plot, diagram radar, dan pemetaan hasil pengelompokan dapat dilihat pada Lampiran 1 dan 2. Adapun hasil pengelompokan data studi kasus dapat dilihat pada Tabel 3.

**Tabel 2** Jumlah *cluster* dan *fuzzifier* terbaik FPSO-FCM beserta indeks validitas data studi kasus *stunting* balita di Provinsi Jawa Timur pada Tahun 2013 dengan jarak Euclidean

$c$	$w$	PC	MPC	CE	FS	XB	PI	Kwon
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<b>2</b>	<b>1,5</b>	<b>0,920</b>	<b>0,839</b>	<b>0,155</b>	20370,85	<b>0,199</b>	0,21204	7,392
2	1,6	0,884	0,767	0,217	35630,17	0,201	0,14029	<b>7,382</b>
2	3,7	0,506	0,012	0,687	21380,65	0,768	<b>0,00006</b>	95,616
2	4	0,502	0,004	0,691	<b>17382,77</b>	1,755	0,00009	267,819

Berdasarkan Tabel 2, dapat dilihat bahwa nilai indeks PC, MPC, CE, dan XB menunjukkan indeks terbaik diperoleh saat jumlah *cluster* sebanyak 2 dan *fuzzifier* sebesar 1,5. Sedangkan indeks Kwon menunjukkan hasil terbaik saat jumlah *cluster* sebanyak 2 dan *fuzzifier* sebesar 1,6. Indeks PI menunjukkan hasil terbaik saat jumlah *cluster* sebanyak 2 dan *fuzzifier* sebesar 3,7. Indeks FS menunjukkan hasil terbaik saat jumlah *cluster* sebanyak 2

dan *fuzzifier* sebesar 4. Dari tabel tersebut dapat disimpulkan bahwa jumlah *cluster* terbaik sebanyak 2 dan *fuzzifier* terbaik sebesar 1,5.

**Tabel 3** Hasil pengelompokan data studi kasus menggunakan FPSO-FCM

<i>Cluster 1</i>	<i>Cluster 2</i>		
(1)	(2)		
Kab. Sidoarjo	Kab. Pacitan	Kab. Jember	Kab. Nganjuk
Kota Kediri	Kab. Ponorogo	Kab. Banyuwangi	Kab. Madiun
Kota Blitar	Kab. Trenggalek	Kab. Bondowoso	Kab. Magetan
Kota Pasuruan	Kab. Tulungagung	Kab. Situbondo	Kab. Ngawi
Kota Mojokerto	Kab. Blitar	Kab. Probolinggo	Kab. Bojonegoro
Kota Madiun	Kab. Kediri	Kab. Pasuruan	Kab. Tuban
Kota Surabaya	Kab. Malang	Kab. Mojokerto	Kab. Lamongan
Kota Batu	Kab. Lumajang	Kab. Jombang	Kab. Gresik
	Kab. Bangkalan	Kab. Sampang	Kab. Pamekasan
	Kab. Sumenep	Kota Malang	Kota Probolinggo

**Variabel Pembeda Antar-Cluster**

Dari uji statistik yang dilakukan, variabel-variabel yang membedakan kedua *cluster* tersebut adalah variabel persentase rumah tangga balita yang tidak memiliki sanitasi layak ( $X_2$ ), rata-rata lama pemberian ASI balita ( $X_4$ ), persentase ibu balita berpendidikan dibawah SMA ( $X_5$ ), persentase ibu balita bekerja ( $X_6$ ), pengeluaran rata-rata rumah tangga balita ( $X_7$ ), persentase balita gizi buruk dan kurang ( $X_{12}$ ), kecukupan dokter ( $X_{14}$ ), kecukupan tenaga ahli gizi ( $X_{15}$ ), kecukupan posyandu ( $X_{17}$ ), dan cakupan pelayanan kesehatan bayi ( $X_{19}$ ). Adapun diagram venn variabel pembeda antar-*cluster* dapat dilihat pada Lampiran 2.

**Tabel 4** Variabel pembeda antar-*cluster*

Variabel	Rata-Rata	
	<i>Cluster 1</i>	<i>Cluster 2</i>
(1)	(2)	(3)
$X_2$	6,82	39,05
$X_4$	14,83	16,97
$X_5$	33,45	63,43
$X_6$	49,15	39,24
$X_7$	3,393	2,145
$X_{12}$	14,38	18,85
$X_{14}$	102,05	18,29
$X_{15}$	13,681	4,375
$X_{17}$	91,78	72,95
$X_{19}$	83,64	90,62

**Interpretasi Hasil Pengelompokan**

Berdasarkan hasil variabel pembeda antar-*cluster*, variabel-variabel yang membedakan adalah  $X_2, X_4, X_5, X_6, X_7, X_{12}, X_{14}, X_{15}, X_{17}$ , dan  $X_{19}$ . Pada variabel  $X_2, X_5, X_6$ , dan  $X_{12}$ , semakin tinggi nilai variabel maka semakin tinggi angka prevalensi *stunting* balita. Sedangkan pada variabel  $X_4, X_7, X_{14}, X_{15}, X_{17}$ , dan  $X_{19}$ , semakin rendah nilai variabel maka semakin tinggi angka prevalensi *stunting* balita.

Dari 10 variabel pembeda antar-*cluster*, hanya terdapat 3 variabel pada *cluster 1* yang mengakibatkan peningkatan angka prevalensi *stunting* balita. Pada variabel persentase ibu balita bekerja ( $X_6$ ), *cluster 1* memiliki nilai rata-rata yang lebih tinggi dibandingkan *cluster 2*. Sedangkan pada variabel rata-rata lama pemberian ASI balita ( $X_4$ ) dan variabel cakupan pelayanan kesehatan balita ( $X_{19}$ ), *cluster 1* memiliki nilai rata-rata yang lebih rendah dibandingkan *cluster 2*. Hal ini sesuai dengan angka prevalensi status gizi balita tinggi badan menurut umur (TB/U) salah satu anggota *cluster 1*, yaitu Kota Surabaya. Menurut Kementerian Kesehatan (), pada tahun 2013 Kota Surabaya merupakan kabupaten/kota dengan persentase status balita pendek dan sangat pendek terkecil,

dengan balita pendek sebesar 16,8% dan balita sangat pendek sebesar 5,9%.

Pada *cluster 2*, terdapat 7 variabel yang mengakibatkan peningkatan angka prevalensi *stunting* balita. Pada variabel persentase rumah tangga balita yang tidak memiliki sanitasi layak ( $X_2$ ), persentase ibu balita berpendidikan dibawah SMA ( $X_5$ ), dan persentase balita gizi buruk dan kurang ( $X_{12}$ ), *cluster 2* memiliki nilai rata-rata yang lebih tinggi dibandingkan *cluster 1*. Sedangkan pada variabel pengeluaran rata-rata rumah tangga balita ( $X_7$ ), kecukupan dokter ( $X_{14}$ ), kecukupan tenaga ahli gizi ( $X_{15}$ ), dan kecukupan posyandu ( $X_{17}$ ), *cluster 2* memiliki nilai rata-rata yang lebih rendah dibandingkan *cluster 1*. Hal ini sesuai dengan angka prevalensi status gizi balita tinggi badan menurut umur (TB/U) salah satu anggota *cluster 2*, yaitu Kabupaten Bondowoso. Menurut Kementerian Kesehatan (Riskesdas dalam Angka Provinsi Jawa Timur 2013), pada tahun 2013 Kabupaten Bondowoso merupakan kabupaten/kota dengan persentase status balita pendek dan sangat pendek tertinggi, dengan balita pendek sebesar 35,8% dan balita sangat pendek sebesar 20,6%.

Dapat disimpulkan bahwa *cluster 2* memiliki keadaan kesehatan yang lebih buruk dibandingkan *cluster 1*. Selain itu, fasilitas dan tenaga kesehatan pada *cluster 1* lebih baik daripada *cluster 2*. Dengan demikian, *cluster 1* dapat dikategorikan sebagai *cluster* dengan potensi *stunting* balita yang rendah sedangkan *cluster 2* dikategorikan sebagai *cluster* dengan potensi *stunting* balita yang tinggi.

## KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, tujuan penelitian telah dicapai sehingga dapat disimpulkan beberapa hal berikut ini :

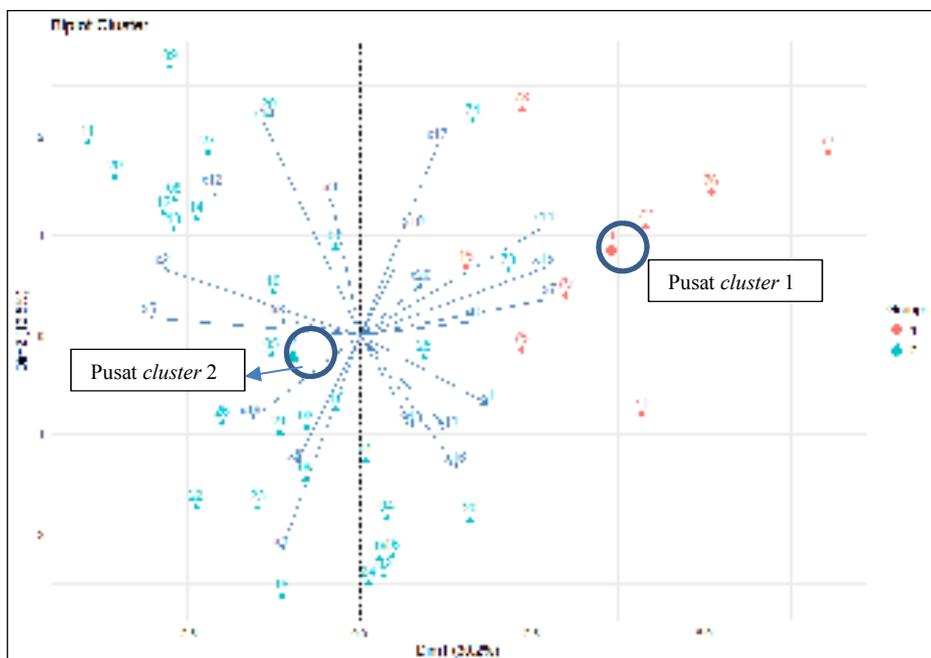
1. Algoritme *Fuzzy Particle Swarm Optimization-Fuzzy C-Means* (FPSO-FCM) mampu meningkatkan performa hasil *clustering*.
2. Pembagian anggota *cluster* pada data studi kasus dapat dilihat pada Tabel 3.
3. Variabel pembeda antar-*cluster* dapat dilihat pada Tabel 4.
4. Daerah yang menjadi anggota *cluster 1* dikategorikan sebagai daerah dengan potensi *stunting* balita yang rendah sedangkan daerah yang menjadi anggota *cluster 2* dikategorikan sebagai daerah dengan potensi *stunting* balita yang tinggi

## DAFTAR PUSTAKA

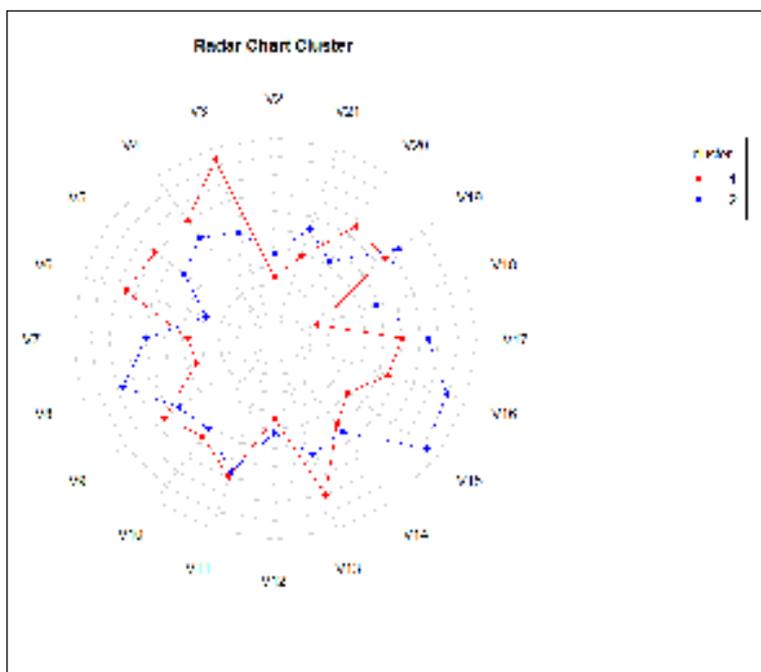
- [1] Abonyi, J. dan Feil, B., 2007. *Cluster Analysis for Data Mining and System Identification*. Berlin. Birkhäuser Verlag AG.
- [2] Bensaid, A. M., 1996. Validity-Guided (Re)Clustering with Applications to Image Segmentation. *IEEE Transactions on Fuzzy Systems*, Vol. 4, No. 2, pp 112-123.
- [3] Chen, M. dan Ludwig, S. A., 2014. Particle Swarm Optimization Based Fuzzy Clustering Approach to Identity Optimal Number of Clusters. *Journal of Artificial Intelligence and Soft Computing Research*, Vol. 4, No. 1, pp 43-56.
- [4] Dave, R. N., 1996. Validating Fuzzy Partitions Obtained Through C-Shells Clustering. *Pattern Recognition Letters*, Vol. 17, No. 6, pp 613-623.
- [5] Flores-Sintas, A., Cadenas, J. M. dan Martin, F., 1999. Membership Functions in The Fuzzy C-Means Algorithm. *Fuzzy Sets and Systems*, Vol. 101, No.1, pp 49-58.

- 
- 
- [6] Ghosh, S. dan Dubey, S. K., 2013. Comparative Analysis of K-Means and Fuzzy C-Means Algorithms. *International Journal of Advanced Computer Science and Application*, Vol. 4, No. 4, pp 35-39.
- [7] Gorunescu, F., 2011. *Data Mining Concepts, Models and Techniques*. Heidelberg. Springer Science dan Business Media.
- [8] Gosain, A. dan Dahiya, S., 2016. Performance Analysis of Various Fuzzy Clustering Algorithms: A Review. *Procedia Computer Science* 79, pp 100-111.
- [9] Hadi, B. S., 2017. *Pendekatan Modified Particle Swarm Optimization dan Artificial Bee Colony pada Fuzzy Geographically Weighted Clustering* (Tesis, Institut Teknologi Sepuluh Nopember).
- [10] Halkidi, M., Baristakis, Y. dan Vazirgiannis, M., 2001. On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, Vol. 17, No. 2-3, pp 107-145.
- [11] Haqiqi, B.N. dan Kurniawan, R., 2015. Analisis Perbandingan Metode Fuzzy C-Means Dan Subtractive Fuzzy C-Means. *Media Statistika*, Vol. 8, No. 2, pp 59-67.
- [12] Izakian, H. dan Abraham, A., 2011. Fuzzy C-Means and Fuzzy Swarm for Fuzzy Clustering Problem. *Expert Systems with Applications*, Vol. 38, No. 3, pp 1835-1838.
- [13] Kementerian Kesehatan Republik Indonesia, 2015. *Rencana Strategis Kementerian Kesehatan Tahun 2015-2019*. Jakarta. Kementerian Kesehatan RI.
- [14] Kementerian Kesehatan Republik Indonesia, 2016. *Situasi Balita Pendek*. Jakarta Selatan. Kementerian Kesehatan RI Pusat Data dan Informasi.
- [15] Kementerian Kesehatan Republik Indonesia, 2013. *Risikedas dalam Angka Provinsi Jawa Timur 2013*. Jakarta. Lembaga Penerbitan Badan Litbangkes.
- [16] Kwon, S. H., 1998. Cluster Validity Index for Fuzzy Clustering. *Electronic Letters*, Vol. 34, No. 22, pp 2176-2177.
- [17] Liu, H.-C., Jeng, B.-C., Yih, J.-M. dan Yu, Y.-K., 2009. Fuzzy C-Means Algorithm Based on Standard Mahalanobis Distance. *Proceedings of the 2009 International Symposium on Information Processing*, pp 422-427.
- [18] Millenium Challenge Account, 2016. *Millenium Challenge Account*. URL: <http://mca-indonesia.go.id/wp-content/uploads/2015/01/Backgrounder-Stunting-ID.pdf>
- [19] Pal, N. R. dan Bezdek, J. C., 1995. On Cluster Validity for Fuzzy C-Means Model. *IEEE Transaction on Fuzzy System*, Vol. 3, No. 3, pp 370-379.
- [20] Wang, W. dan Zhang, Y., 2007. On Fuzzy Cluster Validity Indices. *Fuzzy Sets and Systems*, Vol. 158, No. 19, pp 2095 – 2117.
- [21] Xie, X. L. dan Gerardo, B., 1991. A Validity Measure for Fuzzy Clustering. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 8, pp 841-847.

LAMPIRAN 1

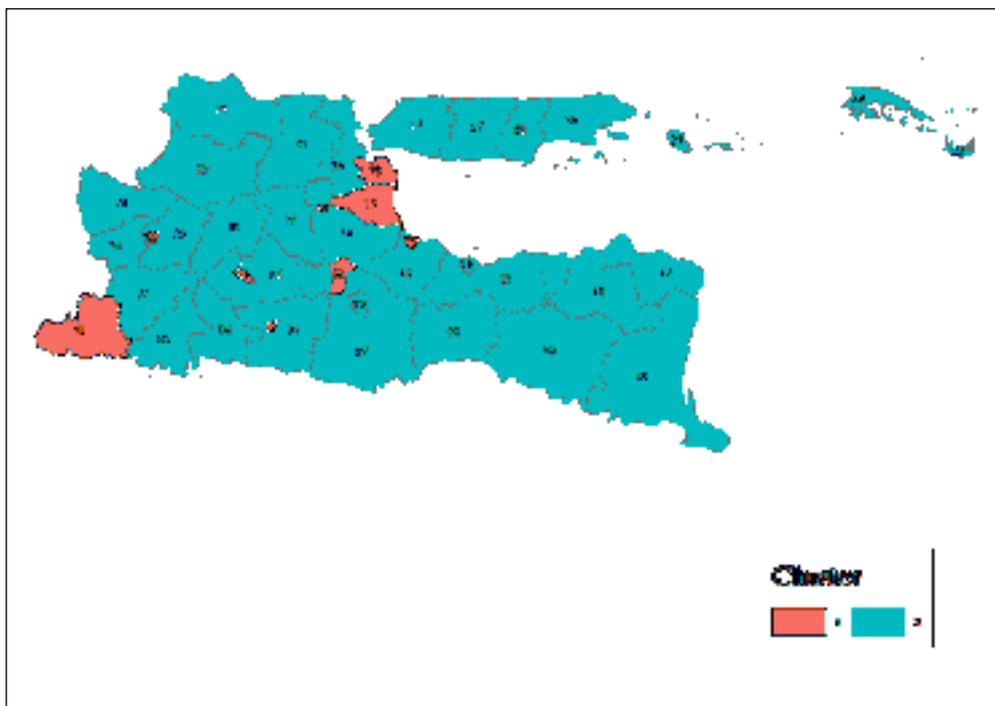


**Gambar 1** Plot hasil pengelompokan FPSO-FCM data studi kasus *stunting* balita di Provinsi Jawa Timur pada Tahun 2013.

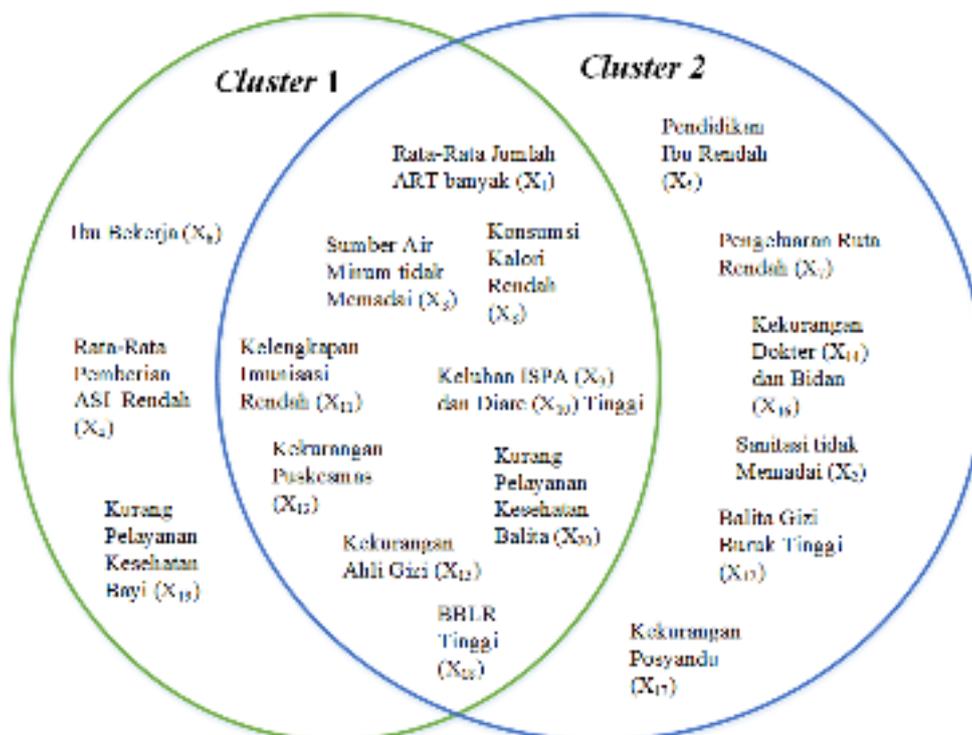


**Gambar 2** Diagram radar hasil pengelompokan FPSO-FCM data studi kasus *stunting* balita di Provinsi Jawa Timur pada Tahun 2013.

LAMPIRAN 2



Gambar 3 Pemetaan hasil pengelompokan kabupaten/kota di Jawa Timur Tahun 2013.



Gambar 4 Diagram venn permasalahan stunting balita di Provinsi Jawa Timur