IMPLEMENTASI METODE K – MEANS DAN *FUZZY* C – MEANS DALAM KLASIFIKASI KELUARGA RISIKO *STUNTING* DI PROVINSI NUSA TENGGARA TIMUR

Sri Miftahul Rahmi[1*], Ria Faulina[2]
[1, 2]*Program Studi Statistika, Universitas Terbuka, Tangerang Selatan, Indonesia*

Penulis korespondensi: srimiftahulrahmi@gmail.com

ABSTRAK

Penurunan angka *stunting* menjadi prioritas utama dalam pembangunan nasional di Indonesia, terutama di Provinsi Nusa Tenggara Timur (NTT) yang memiliki prevalensi *stunting* tertinggi secara nasional. Penelitian ini bertujuan mengelompokkan 22 kabupsten/kota di Provinsi NTT berdasarkan delapan indikator Keluarga Risiko *Stunting* (KRS). Metode klasterisasi, yaitu K-Means dan *Fuzzy* C-Means (FCM) sebagai dua pendekatan klasterisasi non-hirarki yang memungkinkan penelusuran kedekaatan karakteristik antar wilayah. K-Means menghasilkan dua klaster utama yang memisahkan wilayah dengan nilai indikator KRS rendah dan tinggi secara tegas. Metode FCM juga membentuk dua klaster, namun memberikan gambaran yang lebih rinci melalui derajat keanggotaan setiap wilayah, sehingga pola tumpang tindih antar wilayah dapat terlihat lebih jelas. Perbandingan kedua metode menunjukkan bahwa K-Means memberikan pemisahan klaster yang sederhana dan langsung, sedangkan FCM lebih informatif untuk memahami perbedaan tingkat risiko antar kabupaten/kota. Secara keseluruhan, penelitian ini memberikan gambaran kompregensif mengenai penyebaran KRS di provinsi NTT dan dapat menjadi dasar bagi pemerintah daerah dalam menentukan prioritas intervensi *stunting* lebih adaptif dan tepat sasaran.

Kata Kunci: *Stunting*, KRS, *K-Means*, *Fuzzy C-Means*, Klasterisasi

1 PENDAHULUAN

Kesehatan merupakan aspek fundamental dalam peningkatan kualitas sumber daya manusia dan pembangunan nasional. Salah satu permasalahan kesehatan yang masih menjadi fokus utama di Indonesia adalah *stunting* yaitu kondisi gagal tumbuh pada anak akibat kekurangan gizi dalam jangka panjang, terutama pada periode 1000 Hari Pertama Kehidupan (HPK). Kondisi ini tidak hanya berdampak pada tinggi badan anak yang berada di bawah standar, tetapi juga mempengaruhi kemampuan kognitif, peningkatan risiko penyakit kronis, dan rendahnya produktivitas di masa mendatang (BPS, 2023). Pemerintah Indonesia melalui Rencana Pembangunan Jangka Menengah Nasional (RPJMD) 2020 – 2024 telah menetapkan penurunan *stunting* nasional hingga 14% pada tahun 2024. Namun, capaian di beberapa daerah masih menunjukkan ketimpangan sehingga membutuhkan intervensi yang lebih spesifik dan berbasis karakteristik wilayah, terutama kawasan timur Indonesia

Salah satu aspek penting dalam penanganan *stunting* adalah perhatian terhadap Keluarga Resiko *Stunting* (KRS). Badan Pendudukan dan Keluarga Berencana Nasional (BKKBN) mendefinisikan KRS sebagai keluarga yang memiliki satu atau lebih faktor resiko seperti keberadaan remaja putri, calon pengantin, ibu hamil, anak berusia 0 – 23 bulan, atau anak berusia 24 – 59 bulan dalam satu keluarga. Risiko tersebut semakin meningkat apabila keluarga berada pada kondisi ekonomi rendah, memiliki sanitasi yang buruk, serta memiliki akses terbatas terhadap air minum layak (Fatmaningrum, 2022). Oleh sebab itu, analisis terhadap karakteristik keluarga menjadi langkah strategis dalam memetakan potensi masalah dan menyusun kebijakan intervensi yang tepat sasaran.

Provinsi Nusa Tenggara Timur (NTT) merupakan salah satu wilayah dengan tingkat prevalensi *stunting* tertinggi di Indonesia. Berdasarkan hasil Survei Status Gizi Nasional (SSGI)

tahun 2024 prevalensi *stunting* di NTT masih mencapai 37% jauh di atas target nasional. NTT memiliki kondisi sosial, budaya, dan geografis yang beragam, serta terdiri dari 22 kabupaten/kota yang masing-masing memiliki tantangan berbeda terkait kesehatan, pendidikan, sanitasi, dan akses pangan. Kompleksitas kondisi tersebut menjadikan NTT sebagai wilayah yang membutuhkan pendekatan analitik berbasis data untuk mengidentifikasi faktor risiko dan merumuskan kebijakan penanggulangan *stunting* yang efektif, adaptif dan konseptual.

Dalam upaya mempercepat penurunan *stunting*, penggunaan analisis data menjadi semakin penting karena mendukung kebijakan berbasis bukti (*evidence – based policy*). Data *mining* berperan dalam menggali pola terstruktur dan hubungan tersembunyi dalam data sehingga dapat digunakan untuk memahami fenomena sosial dan kesehatan secara komprehensif (Presetyowati, 2017). Teknik yang umum digunakan dalam data *mining* adalah analisis klaster, yaitu metode statistik yang bertujuan mengelompokkan objek berdasarkan karakteristik sehingga objek dalam satu kelompok lebih homogen dibandingkan dengan kelompok lainnya (Kusuma, 2020).

Analisis klaster dapat dilakukan melalui dua pendekatan utama, yaitu metode hiraki dan nonhirarki. Metode nonhirarki seperti K-Means dan *Fuzzy* C-Means banyak digunakan dalam penelitian sosial kesehatan karena lebih fleksibel dan mampu menangani data dalam jumlah besar. Menurut Igual & Segui (2017) K-Means merupakan algoritma berbasis jarak dan mempartisi data ke dalam sejumlah klaster tertentu dan dikenal kesederhanaan serta kecepatannya. Namun, metode ini memiliki keterbatasan karena setiap objek hanya dapat menjadi anggota satu klaster secara tegas. Untuk mengatasi keterbatasan tersebut, FCM dikembangkan sebagai metode yang memungkinkan objek memiliki derajat keanggotaan pada lebih dari satu klaster, sehingga lebih sesuai untuk data dengan karakteristik yang kompleks dan tumpang tindih (Nurmin *et al*., 2022; Mahardika & Agus, 2024).

Beberapa penelitian sebelumnya menunjukkan bahwa metode klaster efektif digunakan dalam menganalisis *stunting*. Widiyanti dan Fitri (2023) mengelompokkan kabupaten/kota di Sumatera Barat berdasarkan indikator KRS menggunakan K-menas dan menemukan dua klaster utama dengan variasi signifikan pada sanitasi dan layanan kesehatan. Sedangkan penelitian Mahardika & Abadi (2024) berhasil memetakan risiko *stunting* pada tingkat puskesmas di Kabupaten Gunungkidul menggunakan K-Means dan FCM, dan hasilnya menunjukkan bahwa FCM memberikan pengelompokan yang lebih stabil dan konsisten berdasarkan metrik evaluasi.

Berbeda dari penelitian sebelumnya yang fokus pada provinsi lain atau unit analisis berbeda, penelitian ini berfokus pada pengelompokan kabupaten/kota berdasarkan delapan variabel KRS dengan penggunaan dua metode yaitu K-Means dan FCM yang memungkinkan dilakukannya perbandingan akurasi dan kualitas klaster sehingga dapat diketahui metode yang paling relevan digunakan dalam konteks NTT.

Dengan demikian, tujuan penelitian ini adalah untuk menganalisis dan mengelompokkan kabupaten/kota di Provinsi NTT berdasarkan indikator keluarga Risiko *Stunting* menggunakan metode K-Means dan FCM, serta mengevaluasi akurasi kedua metode tersebut untuk menentukan metode yang memberikan hasil pengelompokan stabil dan informatif. Hasil ini diharapkan dapat menjadi bahan pertimbangan daerah dalam merumuskan strategi penanganan *stunting* yang lebih efektif dan berkelanjutan di provinsi NTT.

2 METODE

Penelitian ini merupakan penelitian kuantitatif deskriptif dengan pendekatan data *mining* menggunakan teknik *clustering*. Pendekatan ini digunakan untuk mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan Keluarga Risiko *Stunting* (KRS). Teknik analisis yang digunakan adalah metode K-Means dan *Fuzzy* C-Means untuk menemukan pola kesamaan karakteristik wilayah.

2. 1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari situs resmi BKKBN https://siga.bkkbn.go.id/. Data yang digunakan mencakup seluruh kabupaten/kota di Provinsi Nusa Tenggara Timur untuk periode tahun 2024. Data yang diolah terdiri atas delapan variabel yang merepresentasikan kondisi KRS. Adapun delapan variabel tersebut adalah dapat dilihat pada Tabel 1 berikut.

Tabel 1. Variabel penelitian

Variabel	Keterangan	Satuan Variabel	Skala Pengukuran
X_1	Keluarga yang memiliki istri dengan usia 15 s.d 49 tahun (PUS)	Jumlah Keluarga	Rasio
X_2	Keluarga yang memiliki istri dengan usia 15 s.d 49 tahun dan sedang hamil (PUS hamil)	Jumlah Keluarga	Rasio
X_3	Keluarga yang memiliki anak dengan usia 0 s.d 23 bulan (Baduta)	Jumlah Keluarga	Rasio
X_4	Keluarga yang memiliki anak dengan usia 24 s.d 59 bulan (Balita)	Jumlah Keluarga	Rasio
X_5	Keluarga tidak mempunyai sumber air minum layak	Jumlah Keluarga	Rasio
X_6	Keluarga tidak punya jamban layak	Jumlah Keluarga	Rasio
X_7	PUS 4 terlalu (terlalu muda < 20, terlalu tua 35 – 40, terlalu dekat jarak antar anak kurang 2 tahun, dan terlalu banyak > 2 anak)	Jumlah Keluarga	Rasio
X_8	PUS bukan KB modern	Jumlah Keluarga	Rasio

2. 2 Teknik Analisis Data

2.2.1 Pra – Pemrosesan Data

Tahap ini bertujuan untuk menyiapkan data sebelum dilakukan proses klasterisasi. Langkah yang dilakukan adalah normalisasi data, agar setiap variabel berada pada skala yang sebanding dan tidak saling mendominasi dalam pengelompokan. Normalisasi dilakukan dengan menggunakan metode *min-max normalization* dengan rumus (1) sebagai berikut:

$$X_{ij}^{*} = \frac{X_{ij} - X_{\min j}}{X_{\max j} - X_{\min j}} \quad (1)$$

Dimana:

$X'X_{ij}$: objek ke-i dan variable ke-j

$X_{\min j}$: Nilai minimum variabel ke-j

$X_{\max j}$: Nilai maksimum variabel ke-j

2.2.2 Analisis Klaster

Analisis klaster yang digunakan pada klasifikasi keluarga risiko *stunting* di Provinsi Nusa Tenggara Timur adalah K-Means dan FCM. K-Means merupakan salah satu metode analisis klaster nonhierarki, di mana algoritma pengelompokan datanya berdasarkan pada kemiripan atau kedekatan data dengan centroid atau pusat masing-masing kelompok. Algoritma K-Means juga secara tegas mengalokasikan data dari klaster tertentu, sehingga perpindahan data dari satu klaster ke klaster yang lainnya tidak akan dianggap (Budiharto, 2016). Berikut ini adalah algoritma K-Means.

1) Menentukan jumlah klaster (k) dengan metode *Silohutte Coefficient*.
2) Menetapkan cetroid awal secara acak
3) Hitung jarak antara objek dari setiap klaster dengan menggunakan persamaan Euclidean berikut

$$d(x,y) = \sqrt{\sum_{i=1}^{d}(x_i - y_i)^2} \quad (2)$$

dimana:
$d(x,y)$: jarak antara x dan y
x_i : pusat masa pada variabel ke
y_i : data dari i – t variabel
d : jumlah dimensi variabel data
i : indeks dari variabel

4) Mengelompokkan data ke klaster terdekat dan memperbaharui posisi centroid hingga hasil konvergen.

Metode FCM merupakan pengembangan dari K-Means yang memungkinkan setiap data memiliki derajat keanggotaan dilebih dari satu klaster. Konsep ini didasarkan pada teori himpunan *fuzzy*, dimana setiap data tidak secara mutlak tergabung dalam satu kelompok tertentu, melainkan memiliki peluang untuk menjadi bagian dari beberapa kelompok sekaligus. Nilai keanggotaan menunjukkan tingkat keketatan data terhadap masing – masing klaster, dimana semakin besar nilainya semakin tinggi kecenderungan data tersebut menjadi anggota klaster tertentu (Nurmin *et al*., 2022). Tahapan algoritma FCM adalah sebagai berikut(Ghezelbash *et al*., 2025; Mchara *et al*., 2025; Zhang & Huang, 2025):

1) Menentukan jumlah klaster dan parameter fuzzynes $(m > 1)$, yang mengatur tingkat ketidakpastian keanggotaan klaster.
2) Menginisiasikan matriks keanggotaan awal $U = [u_{i,j}]$ secara acak, dimana $u_{i,j}$ menunjukkan tingkat keanggotaan data ke -i pada klaster ke -j.
3) Menghitung pusat klaster (centroid) menggunakan rumus

$$V_{ij} = \frac{\sum_{k=1}^{n}((u_{ik})^m X_{ij})}{\sum_{k=1}^{n}(u_{ik})^m)} \quad (3)$$

dimana:
u_{ik} : nilai objek ke-i dengan pusat klaster ke-k
X_{ij} : objek data ke-i untuk distribusi ke-j
n : jumlah onjek penelitian
m : *fuzzier*

4) Menentukan Fungsi Objektif menggunakan persamaan:

$$P_t = \sum_{k=1}^{n}\sum_{i=1}^{c}\left(\left[\sum_{j=1}^{p}(X_{ij} - V_{ij})^2\right](u_{ik})^m\right) \quad (4)$$

dimana:
P : Jumlah atribut
c : jumlah klaster yang diinginkan
n : Jumlah objek penelitian
X_{ij} : Data ke-j untuk atribut ke-j
V_{ij} : Pusat klaster ke-i untuk atribut ke-j
u_{ik} : Derajat keanggotaan data ke-i pada klaster ke-k
m : *fuzzifier*

5) Memperbaharui derajat keanggotaan setiap data terhadap masing – masing klaster menggunakan persamaan:

$$u_{ik} = \frac{\left[\sum_{j=1}^{p}(X_{ij}-V_{ij})^2\right]^{\frac{-1}{m-1}}}{\sum_{i=1}^{c}\left[\sum_{j=1}^{p}(X_{ij}-V_{ij})^2\right]^{\frac{-1}{m-1}}} \quad (5)$$

dimana

P : Jumlah atribut
c : jumlah klaster yang diinginkan
X_{ij} : Data ke-j untuk atribut ke-j
V_{ij} : Pusat klaster ke-i untuk atribut ke-j
u_{ik} : Derajat keanggotaan data ke-i pada klaster ke-k
m : *fuzzifier*

6) Periksa Konvergensi merupakan cara untuk memastikan kondisi iterasi berhenti dapat menggunakan
   a) Apabila $|P_t - P_{t-1}| < \epsilon \; atau \; t > MaxIter$ maka berhenti;
   b) Apabila tidak: $t = t + 1$, makamennagulangi langkah ke 3
7) Menentukan anggota klaster dengan menggunakan nilai derajat keanggotaan pada iterasi terakhir untuk menentukan jumlah anggota dari satu klaster. Suatu data dapat dikatakan anggota dari satu klaster apabila derajat keanggotaan terbesar pada klaster tersebut (Nurmin et al., 2022)
8) Uji validasi menggunakan Uji dengan menggunakan:
   a) *Partition Coefficient Indeks* (PCI)
   Metode ini digunakan untuk menilai derajat keterbatasan karena hanya berfokus pada derajat keanggotaan tanpa memperhatikan informasi geometris pada vektor data. Jika nilai PCI mendekati 1 maka kualitas klaster yang di hasilkan juga semakin baik (Putri et al., 2022). Menggunakan persamaan sebagai berikut

$$PCI = \frac{1}{n}\left(\sum_{i=k}^{n}\sum_{k=1}^{c}u_{ik}^2\right) \quad (6)$$

   dimana

   n : Banyak objek Penelitian
   c : Banyak kelompok
   $u$ik : Nilai keanggotaan-k dengan pusat kelompok ke-i

   b) *Modified Pertition Coefficient* (MPC)
   MPC merupakan modifikasi dari PC dan mampu mengurangi perubahan yang monoton ketika jumlah klaster berubah, menurut Putri (2022) MPC dapat memvalidasi jumlah klaster yang tepat adapun persamaan metode MPC adalah sebagai berikut

$$MPC(c) = 1 - \frac{c}{c-1}(1 - PC(c)) \quad (7)$$

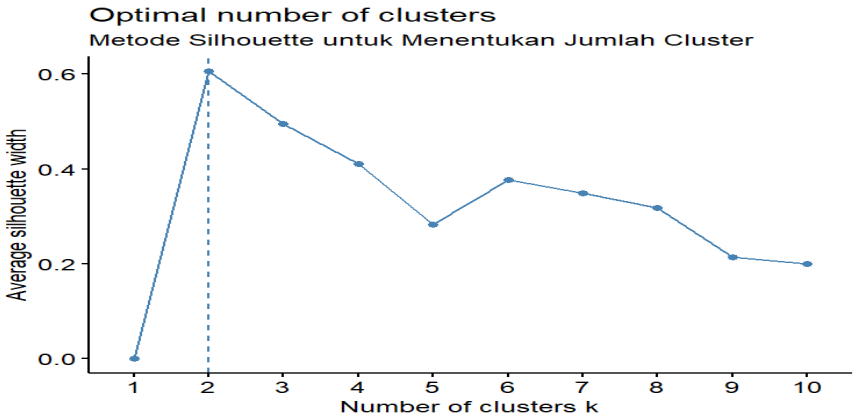
   dimana:

   $MPC(c)$: *Modification Peertition Coefficient* untuk jumlah klaster tertentu
   c : Jumlah klaster
   $PC(c)$: *Partition Coefficient* (PC) untuk jumlah klaster yang sama

3 HASIL DAN PEMBAHASAN

3.1 Hasil klasterisasi K-Means

Analisis K-Means diawali dengan menentukan jumlah klaster optimal menggunakan plot *Average Silhoutte Width* seperti pada Gambar 1. *Average Silhoutte Width* adalah rata-rata nilai *silhouette* dari seluruh data pada satu jumlah klaster tertentu (*k*). Jumlah klaster optimal ditentukan berdasarkan nilai rata-rata *Silhoutte* tertinggi.



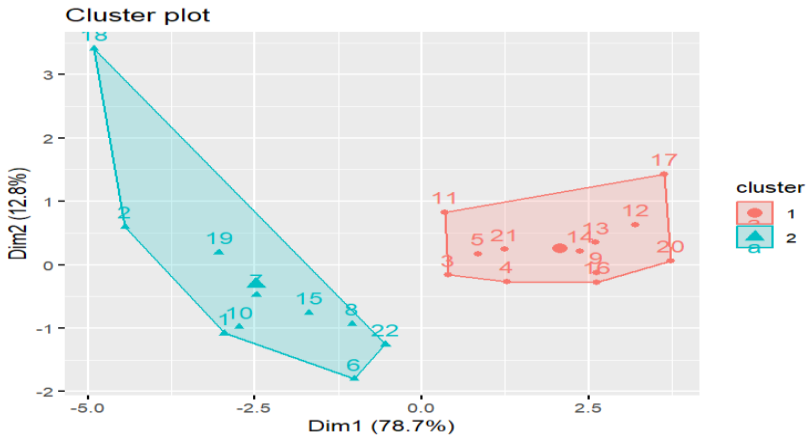
Gambar 1. Nilai *Silhoutte Coefficient* setiap K-Klaster

Hasil perhitungan menunjukkan nilai *Silhoutte Coefficient* tertingggi pada k=2, sehingga jumlah klaster yang digunakan adalah dua kelompok. Hal ini dikarenakan pada klaster ketiga dan selanjutnya terjadi penurunan nilai *average silhoutte width.* Nilai *silhoutte* sebesar 0,6 menunjukkan kualitas pengelompokan yang cukup baik, dengan jarak antar klaster yang jelas dan pengelompokan internal yang stabil.

Tabel 2. Pusat klaster K-Means

Klaster	PUS	PUS Hamil	Baduta	Balita	SAM	JAMBAN	PUS 4T	BPKBM
1	0.2565	0.1984	0.3458	0.2544	0.1039	0.1373	0.2588	0.2360
2	0.7777	0.6913	0.8313	0.7264	0.2805	0.3593	0.7678	0.7442

Berdasarkan hasil perhitungan pusat klaster pada Tabel 2, klaster 1 ditandai dengan nilai rata-rata delapan variabel KRS yang masih rendah pada PUS, PUS Hamil, Baduta, Balita, PUS 4T dan BPKBM. Hal ini menunjukkan bahwa kabupaten/kota di klaster ini memiliki kondisi risiko *stunting* yang relatif lebih rendah. Sementara itu, klaster 2 memiliki nilai rata-rata lebih tinggi pada hampir semua variabel, menggambarkan wilayah dengan risiko *stunting* lebih besar. Adapun visualisasi pengelompokan dapat dilihat pada Gambar 2.



Gambar 2. Plot klaster K-Means

Visualisasi klaster melalui Gambar 2 menunjukkan pemisahan dua klaster yang tegas, hasil pengelompokan wilayah sebagaimana pada klaster 1 berisi kabupaten/ kota seperti Timor

Tengah Utara, Belu, Alor, Ngada, Sumba Timur, Sumba Barat, Lembata, Rote Ndao, Nageko, Sumba Tengah, Sabu Raijua dan Malaka. Kelompok ini memiliki karakteristik demografis dan sanitasi yang cenderung lebih baik. Sebaliknya, klaster 2 terdiri dari Kupang, Timor Tengah Selatan, Flores Timur, Sikka, Ende, Manggarai, Manggarai Barat, Sumba Barat Daya, Manggarai Timur, Kota Kupang. Daerah-daerah ini memiliki jumlah PUS dan PUS Hamil tinggi, beban Baduta dan Balita lebih besar, serta tantangan pada sanitasi dan partisipasi KB modern.

Secara umum, K-Means mampu memberikan gambaran risiko *stunting* di Provinsi NTT, namun sifat pengelompokan yang tegas membuat variasi internal antara daerah tidak sepenuhnya tergambar. Hal ini menjadi alasan bahwa hasil K-Menas perlu d bandingkan dengan FCM yang lebih fleksibel.

3.2 Hasil Klasterisasi *Fuzzy* C-Means

Proses klasterisasi dilakukan dengan kluster (k) = 2, pembobotan (m) = 2, iterasi Maksimum (*Maxiter*) = 1000, $\varepsilon = 0{,}001$, fungsi objektif awal = 0 dan iterasi awal =1. Setelah itu diperoleh nilai pusat klaster pada iterasi ke-10 seperti yang ada pada Tabel 3.

Tabel 3. Pusat klaster pada iterasi pertama (t=10)

Klaster	PUS	PUS Hamil	Baduta	Balita	SAM	JAMBAN	PUS 4T	BPKBM
1	0.2449	0.1880	0.3342	0.2422	0.0984	0.1210	0.2451	0.2326
2	0.7834	0.7154	0.8485	0.7412	0.2833	0.3737	0.7725	0.7174

Pada Tabel 3 menunjukkan dua pusat klaster yang memiliki pola serupa dengan K-Means, namun dengan nilai yang lebih halus karena mempertimbangkan derajat keanggotaan. Pada sajian data tersebut klaster 2 tetap muncul sebagai kelompok dengan nilai paling tinggi pada variabel PUS, PUS Hamil, Baduta, Balita, Jamban, PUS 4T dan BPKBM. Sebaliknya klaster 1 menjadi representasi daerah dengan risiko *stunting* lebih rendah. Pola ini konsisten dengan karakteristik demografis di NTT, di mana perbedaan akses kesehatan, sanitasi dan konsenterasi penduduk sangat mencolok antar wilayah.

Tabel 4. Derajat keanggotaan pada iterasi ke-10

No.	Kabupaten/Kota	u_{i1}	u_{i2}	Klaster
1	Kupang	0,0369	0,9631	2
2	Timor Tengah Selatan	0,0916	0,9084	2
3	Timor Tengah Utara	0,0586	0,4132	2
4	Belu	0,8911	0,1089	1
5	Alor	0,8055	0,1945	1
6	Flores Timur	0,2608	0,7392	2
7	Sikka	0,0416	0,9584	2
8	Ende	0,2539	0,7461	2
9	Ngada	0,9866	0,0134	1
10	Manggarai	0,0409	0,9591	2
11	Sumba Timur	0,6568	0,3452	1
12	Sumba Barat	0,947	0,053	1
13	Lembata	0,9788	0,0212	1
14	Rote Ndao	0,9933	0,0067	1
15	Manggarai Barat	0,1181	0,8819	2
16	Nagekeo	0,9833	0,0167	1
17	Sumba Tengah	0,9003	0,0997	1

No.	Kabupaten/Kota	u_{i1}	u_{i2}	Klaster
18	Sumba Barat Daya	0,2105	0,7895	2
19	Manggarai Timur	0,0798	0,9202	2
20	Sabu Raijua	0,9267	0,0733	1
21	Malaka	0,9118	0,0882	1
22	Kota Kupang	0,4442	0,5558	2

FCM tidak langsung menempatkan data secara tegas, tetapi memberikan nilai keanggotaan u_{i1} dan u_{i1}pada masing-masing klaster. Pada Tabel 4 beberapa daerah menunjukkan keanggotaan yang sangat tegas misalnya Rote Ndao $u_{i1} = 0{,}9933$ dan Ngada $u_{i1} = 0{,}9866$ menandakan karakteristik wilayah yang sanggat mirip dengan klaster 1. Sebaliknya Kupang $u_{i2} = 0{,}931$ dan Sikka ($u_{i2} = 0{,}984$ sangat identik dengan klaster 2. Ciri khas metode FCM terlihat pada daerah *boderline seperti* Kota Kupang $u_{i1} = 0{,}4442$ dan nilai $u_{i2} = 0{,}5558$ nilai ini menunjukkan bahwa karakteristik wilayahnya tidak sepenuhnya tegas masuk satu kalster, sehingga pendekatan *fuzzy* lebih mampu menangkap transisi kondisi sosial demografis di wilayah tersebut.

Klaster 2 memuat daerah Kupang, Timor Tengah Selatan, Timor Tengah Utara, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Sumba Barat Daya, Manggarai Timur, dan Kota Kupang. Daerah-daerah dalam klaster ini dicirkan oleh tingginya jumlah PUS, PUS Hamil, Baduta, serta Balita sehingga beban pelayanan kesehatan ibu dan anak relatif besar. Selain itu tingginya proporsi keluarga tanpa akses jamban layak, meningkatnya kasus PUS 4T serta rendahnya penggunaan KB modern menandakan ada masalah reproduksi dan sanitasi yang berpotensi memperburuk risiko *stunting*. Kondisi tersebut mengindikasikan perlunya intervensi lebih sensitif seperti peningkatan sanitasi, penguatan layanan kesehatan ibu dan anak, edukasi gizi serta penguatan KB modern. Disisi lain klaster 1 berisi Kabupaten/kota Belu, Alor, Ngada, Sumba Timur, Sumba Barat, Lembata, Rite Ndao, Nagekeo, Sumba Tengah, Sabu Raijua, dan Malaka. Menggambarkan wilayah dengan risiko lebih rendah karena sebagian besar variabel KRS berada di bawah rata-rata. Hal ini mencerminkan daerah tersebut telah melakukan upaya mitigasi KRS dengan lebih baik sebagaimana sejalan dengan pendekatan intervensi terpadu yang di rekomendasikan BKKBN (2021).

Selain itu nilai dari validasi klaster FCM dengan menggunakan *Partition Coefficient Indeks* (PCI) sebesar 0,7965 menunjukkan kejelasan klaster cukup baik dimana, sebagian besar data memiliki afilasi dominan pada satu klaster sedangkan nilai *Modified Partition Coefficient* (MPC) sebesar 3,593 menunjukkan kualitas partisi pada kategori sedang, namun masih dalam batas valid. Nilai PCI dan MPC memperlihatkan bahwa FCM tidak hanya mampu membentuk dua klaster dengan stabil tetapi juga mengakomodasi variasi internal antar wilayah yang tidak tertangkap oleh K-Means.

3.3 Perbandingan hasil K-Means dan FCM

K-Means dan FCM menghasilkan jumlah klaster sama, tetapi anggota berbeda. Baik K-Means maupun FCM menghasilkan dua klaster, namun dengan distribusi daerah yang tidak identik. FCM menggeser beberapa wilayah yang memiliki karakteristik di antara dua kecenderungan seperti Kota Kupang dan Timur Tengah Utara menjadi anggota klaster risiko tinggi, mengikuti pola keanggotaan yang lebih konseptual. FCM lebih fleksibel dalam menangkap variasi antar wilayah melalui nilai derajat keanggotaan yang mungkinkan lebih relevan untuk data sosial kesehatan, sedangkan K-Means memasukkan setiap daerah secara tegas pada satu klaster, sehingga transisi gradatif antar wilayah tidak terlihat. Interpretasi kebijakan lebih kuat jika digambarkan dengan FCM hal ini dikarenakan FCM membantu

mengidentifikasi wilayah dengan karakteristik menengah yang membutuhkan perhatian khusus meski tidak ekstrim, misalnya kota Kupang, Flores Timur dan Ende.

4 KESIMPULAN

Berdasarkan hasil analisis menggunakan metode K-Means dan *Fuzzy* C-Means terhadap delapan indikator Keluarga Risiko *Stunting* (KRS) di Provinsi Nusa Tenggara Timur (NTT), diperoleh beberapa kesimpulan penting yakni secara keseluruhan K-Means memberikan gambaran awal pola risiko *stunting* yang cukup baik dan mudah diinterpretasikan. Namun, FCM menghasilkan pemetaan yang lebih mendalam dan realistis terhadap kondisi sosial demografis antar wilayah di Provinsi NTT. Dengan mempertimbangkan nilai pusat klaster, anggota klaster, serta validasinya, metode FCM dinilai lebih unggul dan lebih tepat digunakan untuk analisis Keluarga Risiko *Stunting* karena mampu menangkap tingkat variasi risiko yang tidak bersifat tegas antar wilayah.

DAFTAR PUSTAKA

Badan Pusat Statistik. (2023). *Laporan Indeks Khusus Penanganan Stunting 2021–2022*. *Sustainability (Switzerland)*, 11(1), 1–14.

Budiharto, W. (2016). *Machine Learning dan Computational Intelligence*. Yogyakarta: Andi.

Badan Kependudukan dan Keluarga Berencana Nasional. (2023). *Konsep Pemutahiran, Verifikasi dan Validasi Keluarga Berisiko Stunting.*

Badan Kependudukan dan Keluarga Berencana Nasional. (2021). Kebijakan dan Strategi Percepatan Penurunan *Stunting* di indonesia. *Sustainability* (*Switzerland*), 11(1), 1–110

Fatmaningrum, W., Nadhiroh, S. R., Raikhani, A., Utomo, B., Masluchah, L., & Patmawati. (2022). Analisis situasi upaya percepatan penurunan *stunting* dengan pendekatan keluarga berisiko *stunting*. *Media Gizi Indonesia*, 17(1SP), 139–144.

Ghezelbash, R., Maghsoudi, A., & Daviran, M. (2025). A New Framework For Exploration Targeting: Integrating Multifractal Geochemical Analysis, Structural Controls And Fuzzy C-Means Unsupervised Clustering. *Geochemistry*, 85(4), 126340. https://doi.org/10.1016/J.CHEMER.2025.126340

Igual, L., & Segui, S. (2017). *Introduction to Data Science*. Springer.

Kusuma, P. D. (2020). *Machine Learning: Teori, Program dan Studi Kasus*. Yogyakarta: Deepublish Publisher.

Mahardika, B. W., & Abadi, A. M. (2024). Implementation of K-Means and *Fuzzy* C-Means clustering for mapping toddler *stunting* cases in Gunungkidul District. *Journal of Mathematics and Its Applications*, 18(4), 2231–2246.

Mchara, W., Manai, L., Khalfa, M. A., Raissi, M., Dimassi, W., & Hannachi, S. (2025). A hybrid deep learning framework for global irradiance prediction using fuzzy C-Means, CNN-WNN, and Informer models. *Cleaner Engineering and Technology*, *28*, 101061. https://doi.org/10.1016/J.CLET.2025.101061

Nurmin, D., Hayati, M. N., & Goejantoro, R. (2022). Penerapan metode *Fuzzy* C-Means pada pengelompokan kabupaten/kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat tahun 2020. *Eksponensial*, 13(2), 189–198.

Prasetyowati, E. (2017). *Data Mining: Pengelompokan Data untuk Informasi dan Evaluasi*. Pamekasan: Duta Media Publishing.

Putri, G. N. S., Ispriyanti, D., & Widiharih, T. (2022). Implementasi algoritma *Fuzzy* C-Means dan *Fuzzy* Possibilistic C-Means untuk klasterisasi data tweets pada akun Twitter Tokopedia. *Jurnal Gaussian*, 11(1), 86–98.

Widiyanti, F., & Fitri, F. (2024). Application of the K-Means clustering algorithm to the case of *stunting* risk families in districts/cities of West Sumatra Province in 2023. *Mathematical Journal of Modelling and Forecasting*, 2(2), 1–8.

Zhang, H., & Huang, S. L. (2025). Improved fuzzy C-means clustering algorithm based on fuzzy particle swarm optimization for solving data clustering problems. *Mathematics and Computers in Simulation*, *233*, 311–329. https://doi.org/10.1016/J.MATCOM.2025.02.012