

**Penerapan Metode Fuzzy C-Means Pada Pengelompokan Kabupaten/Kota  
di Pulau Kalimantan Berdasarkan Indikator Kesejahteraan Rakyat Tahun 2020**

*Application of the Fuzzy C-Means Method in the Grouping of Regencies/Cities in Kalimantan  
Island Based on People's Welfare Indicators in 2020*

Deviyana Nurmin<sup>1</sup>, Memi Nor Hayati<sup>2</sup>, Rito Goejantoro<sup>3</sup>

<sup>1,2</sup>Laboratorium Statistika Terapan, FMIPA, Universitas Mulawarman

<sup>3</sup>Laboratorium Statistika Komputasi, FMIPA, Universitas Mulawarman

E-Mail: [devi.dd591@gmail.com](mailto:devi.dd591@gmail.com)

**ABSTRACT**

*Clustering is a method of grouping data into several clusters or groups so that data in one cluster has a high level of similarity and data between clusters has a low level of similarity. The clustering method used in this research is Fuzzy C-Means (FCM). FCM is a data grouping technique in which the existence of each data point in a cluster is determined by the degree of membership. To optimize the grouping results, it is necessary to validate the number of clusters using Partition Coefficient (PC). The purpose of this study is to obtain optimal grouping results from the FCM method using the PC validity indices from the people's welfare indicator data in 56 regencies/cities on the island of Kalimantan in 2020. Based on the results of the analysis, the conclusion is that the optimal number of clusters is three clusters. The first cluster consists of 24 regencies/cities on the island of Kalimantan, the second cluster consists of 17 regencies/cities on the island of Kalimantan, and the third cluster consists of 15 regencies/cities on the island of Kalimantan.*

**Keywords:** Fuzzy C-Means, People's Welfare Indicator, Validity Index

**Pendahuluan**

*Clustering* atau klusterisasi adalah sebuah proses untuk mengelompokkan data ke dalam beberapa klaster atau kelompok sehingga data dalam satu klaster memiliki tingkat kemiripan yang tinggi dan data antar klaster memiliki tingkat kemiripan yang rendah (Tan, dkk., 2006). Dilakukan pengelompokan dengan menggunakan *fuzzy clustering* di mana dalam melakukan pengelompokan mempertimbangkan tingkat keanggotaan himpunan *fuzzy* sebagai dasar pembobotan (Jang, dkk., 1997).

Menurut Kusumadewi & Purnomo (2010), ada beberapa algoritma *fuzzy clustering*, salah satu di antaranya adalah *Fuzzy C-Means* (FCM). FCM adalah suatu teknik pengelompokan data yang mana keberadaan setiap titik data dalam suatu klaster ditentukan oleh derajat keanggotaannya. Konsep dasar FCM yaitu menentukan pusat klaster terlebih dahulu, di mana pusat klaster ini akan mengelompokkan setiap titik dalam suatu kelompok. Penentuan pusat klaster dan nilai keanggotaan akan dilakukan secara berulang sehingga didapat pusat klaster yang nantinya akan bergerak ke lokasi yang tepat dikarenakan kondisi pada pusat klaster dan titik setiap datanya berdasarkan derajat keanggotaan belum terlalu akurat (Setiawan, dkk., 2018).

Untuk mengoptimalkan hasil pengelompokan, maka perlu dilakukan validasi jumlah klaster menggunakan indeks validitas. Indeks validitas merupakan ukuran validitas untuk memperoleh

jumlah klaster optimal yang sepenuhnya dapat menjelaskan struktur data (Zhao & Franti, 2014). Klaster yang optimal adalah klaster yang memiliki varians yang kecil antar elemen/anggota di dalam setiap klaster (homogen) dan memiliki varians yang besar antar klasternya (heterogen). Metode untuk menghitung indeks validitas pada penelitian ini adalah *Partition Coefficient* (PC). Indeks PC mengukur jumlah *overlapping* antar kelompok (Wang & Zhang, 2007).

Pembangunan suatu negara dapat dinilai dari kesejahteraan rakyatnya. Kesejahteraan dapat diartikan sebagai suatu keadaan di mana setiap warga negara selalu berada dalam kondisi serba kecukupan segala kebutuhannya, baik material maupun spiritual (Basri & Munandar, 2009). Kesejahteraan rakyat adalah salah satu tujuan negara yang tertuang dalam Pembukaan Undang-Undang Dasar (UUD) 1945 alinea IV. Dalam melaksanakan program pembangunan perlu dilakukan pengidentifikasian karakteristik atau indikator berdasarkan tingkat kesejahteraan rakyat pada setiap daerah sehingga pemerintah dalam memutuskan kebijakan dan strategi pembangunan dapat sesuai dengan tujuan Negara (Alwi & Hasrul, 2018).

Suatu taraf kesejahteraan hanya dapat terlihat melalui aspek tertentu karena kesejahteraan rakyat memiliki dimensi yang sangat luas dan kompleks. Sehingga Badan Pusat Statistik (BPS) mengamati kesejahteraan rakyat dari berbagai aspek spesifik, yaitu kependudukan, pendidikan, kesehatan,

fertilitas dan keluarga berencana, perumahan, teknologi informasi dan komunikasi, dan lainnya.

Berdasarkan uraian tersebut, maka penulis tertarik untuk melakukan penelitian dengan judul “Penerapan Metode *Fuzzy C-Means* Pada Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Kesejahteraan Rakyat Tahun 2020” yang bertujuan untuk memperoleh hasil pengelompokan yang optimal menggunakan indeks validitas PC dari kabupaten/kota di Pulau Kalimantan menggunakan metode FCM berdasarkan indikator kesejahteraan rakyat. Sehingga informasi yang dihasilkan diharapkan dapat menjadi bahan referensi dan pertimbangan pemerintah setempat dalam membuat kebijakan atau keputusan terkait kesejahteraan rakyat.

### **Data Mining**

*Data mining* merupakan proses pencarian ataupun penambangan informasi baru dengan mencari pola-pola yang menarik dan tersembunyi (*hidden pattern*) dari suatu kumpulan data yang berukuran besar yang tersimpan dalam suatu basis data, data *warehouse*, atau tempat penyimpanan data lainnya (Tan, dkk., 2006). Ada beberapa teknik yang dapat digunakan dalam *data mining* untuk mendapatkan pola-pola dan informasi tersembunyi, yaitu *classification*, *neural network*, *decision tree*, *genetic algorithm*, *clustering*, *Online Analytical Processing (OLAP)*, dan *association rules* (Prasetyo, dkk., 2020).

### **Analisis Multivariat**

Menurut Widarjono (2015), analisis multivariat berasal dari kata *multy* dan *variate* yang berarti alat analisis lebih dari dua variabel. Analisis multivariat berhubungan dengan semua teknik statistik yang secara simultan menganalisis sejumlah pengukuran pada objek. Analisis ini dikelompokkan menjadi dua kelompok besar yaitu, metode dependensi merupakan analisis ketergantungan yang digunakan untuk menjelaskan atau meramalkan nilai variabel terikat berdasarkan lebih dari satu variabel bebas yang memengaruhinya. Metode ini dibagi menurut jumlah variabel terikatnya. Apabila terdapat satu variabel terikat, maka analisis yang digunakan adalah analisis varians, analisis kovarians, regresi berganda, dan analisis konjoin. Apabila terdapat lebih dari satu variabel terikat, maka analisis yang digunakan adalah analisis multivariat, analisis diskriminan, dan korelasi kanonik. Sedangkan metode interdependensi merupakan analisis yang digunakan untuk memberikan arti kepada suatu *set* variabel (kelompok variabel) atau pengelompokan suatu *set* variabel menjadi kelompok yang lebih sedikit jumlahnya dan masing-masing kelompok membentuk variabel baru yang disebut faktor.

Beberapa contoh analisis interdependensi adalah analisis faktor, analisis kluster, dan penskalaan multidimensional.

### **Analisis Kluster**

Analisis kluster merupakan analisis yang bertujuan untuk mengelompokkan objek pengamatan yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan diantaranya. Tujuannya adalah agar objek-objek yang bergabung ke dalam sebuah kelompok merupakan objek-objek yang mirip (atau berhubungan) satu sama lain dan berbeda (atau tidak berhubungan) dengan objek yang berada pada kelompok lain. Lebih besar kemiripannya (homogenitas) dalam suatu kelompok maka lebih besar perbedaannya di antara kelompok lainnya (Prasetyo, 2012).

Menurut Prasetyo (2012), metode dalam analisis kluster dibagi menjadi dua yaitu, metode hirarki dan non hirarki. Metode hirarki dimulai dengan pengelompokan dua atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Contoh metode hirarki adalah *single linkage*, *complete linkage*, dan *average linkage*. Sedangkan metode non hirarki dimulai dengan menentukan terlebih dahulu banyak kluster yang diinginkan. Contoh metode non hirarki adalah *c-means* dan *fuzzy c-means*.

### **Teori Himpunan Fuzzy**

Himpunan *fuzzy* didasarkan pada gagasan untuk memperluas jangkauan fungsi keanggotaan pada himpunan *crisp* sedemikian sehingga fungsi tersebut mencakup bilangan *real* pada interval  $[0,1]$ . Derajat keanggotaannya menunjukkan bahwa suatu elemen dalam semesta pembicaraan tidak hanya berada pada 0 dan 1, namun juga nilai yang terletak di antaranya. Dengan kata lain, nilai kebenaran suatu pernyataan tidak hanya bernilai benar atau salah, nilai 1 menunjukkan benar, nilai 0 menunjukkan salah dan masih ada nilai-nilai yang terletak antara benar dan salah (Kusumadewi & Purnomo, 2010).

### **Fuzzy Clustering**

*Fuzzy clustering* adalah metode pengelompokan berdasarkan derajat keanggotaan yang mencakup himpunan *fuzzy* sebagai dasar pembobotan bagi pengelompokan. Masing-masing data diberikan nilai kemungkinan untuk bisa bergabung ke setiap kelompok yang ada, artinya data tidak mutlak menjadi anggota satu kelompok saja, tetapi juga mempunyai nilai kemungkinan untuk menjadi anggota kelompok lain dengan derajat keanggotaan terbesar menunjukkan kecenderungan yang tinggi suatu data untuk

menjadi anggota kelompok tertentu (Jang, dkk., 1997).

**Fuzzy C-Means**

Metode *Fuzzy C-Means* (FCM) merupakan salah satu teknik pengelompokan data yang mana keberadaan setiap datanya dalam suatu kluster ditentukan oleh derajat keanggotaan. Konsep dasar FCM, pertama menentukan pusat kluster yang akan menandai lokasi rata-rata untuk setiap kluster, pusat kluster ini masih belum akurat. Dengan cara memperbaiki pusat kluster dan nilai keanggotaan setiap titik data secara berulang, maka akan dapat dilihat bahwa pusat kluster akan bergerak menuju lokasi yang tepat (Kusumadewi & Purnomo, 2010).

**Indeks Validitas**

Indeks validitas merupakan suatu ukuran yang digunakan untuk menentukan jumlah kelompok yang optimal yang dapat menjelaskan struktur data (Wang & Zhang, 2007). Bezdek & Dunn (1975) menyarankan indeks validitas kluster untuk *fuzzy clustering*, satu diantaranya yaitu *Partition Coefficient* (PC). Indeks PC digunakan untuk mengukur jumlah *overlapping* antar kelompok. Semakin besar nilai PC, maka semakin baik kluster yang dihasilkan.

**Kesejahteraan Rakyat**

Menurut Perserikatan Bangsa-Bangsa (PBB), kesejahteraan adalah suatu kondisi atau keadaan sejahtera baik fisik, mental maupun sosial dan tidak hanya perbaikan-perbaikan penyakit sosial tertentu saja. Menurut Badan Pusat Statistik (2020), kualitas penduduk sangat ditentukan oleh berbagai faktor baik faktor internal, maupun faktor eksternal mulai dari suasana lingkungan keluarga, lingkungan masyarakat dan kebijakan yang diambil oleh pemerintah. Guna menilai kualitas penduduk diperlukan indikator ataupun ukuran yang dapat menunjukkan kondisi penduduk. Adapun indikator untuk mengamati kesejahteraan rakyat adalah kependudukan, pendidikan, kesehatan, fertilitas dan keluarga berencana, perumahan, teknologi informasi dan komunikasi, dan lainnya.

**Variabel Penelitian**

Adapun variabel yang digunakan dalam penelitian ini adalah sebagai berikut:

- X<sub>1</sub> : Angka Beban Ketergantungan
- X<sub>2</sub> : Kepemilikan Akta Kelahiran
- X<sub>3</sub> : APS 7-12 Tahun
- X<sub>4</sub> : APS 13-15 Tahun
- X<sub>5</sub> : APS 16-18 Tahun
- X<sub>6</sub> : APM SD
- X<sub>7</sub> : APM SMP
- X<sub>8</sub> : APM SMA

- X<sub>9</sub> : Keluhan Kesehatan
- X<sub>10</sub> : Wanita Menurut Umur Perkawinan Pertama
- X<sub>11</sub> : Penggunaan Alat/Cara KB
- X<sub>12</sub> : Fasilitas Tempat Buang Air Besar
- X<sub>13</sub> : Sumber Air Minum Bersih
- X<sub>14</sub> : Sumber Penerangan Listrik
- X<sub>15</sub> : Penggunaan Internet
- X<sub>16</sub> : Korban Kejahatan
- X<sub>17</sub> : PIP
- X<sub>18</sub> : BPNT

**Metode Penelitian**

Adapun langkah-langkah yang dilakukan dalam analisis data adalah sebagai berikut:

1. Melakukan pengelompokan data dengan metode *fuzzy c-means*. Berikut adalah tahapan dalam algoritma *fuzzy c-means*:
  - a. Menentukan:
    - i. Banyak kluster (*c*) merupakan banyaknya kluster yang akan dibentuk.
    - ii. Pangkat (*m*) umumnya bernilai  $1,25 \leq m \leq 2$  yang di mana semakin besar nilai *m* yang digunakan maka menunjukkan bahwa pusat data semakin berada di tengah data.
    - iii. Maksimum iterasi (*MaxIter*) merupakan perulangan yang akan berhenti jika nilai maksimal iterasi sudah tercapai.
    - iv. *Error* terkecil yang diharapkan ( $\epsilon$ ) merupakan batasan nilai yang membuat perulangan akan berakhir setelah didapatkan nilai *error* yang diharapkan.
    - v. Fungsi objektif awal (*P*<sub>0</sub>) merupakan suatu fungsi yang akan dioptimumkan (maksimum atau minimum), nilai 0 berarti untuk mendapatkan nilai minimum.
  - b. Membangkitkan bilangan *random*  $\mu_{ik}, i = 1, 2, \dots, n; k = 1, 2, \dots, c$  sebagai elemen-elemen matriks partisi awal *U* berukuran  $i \times k$ .
  - c. Menghitung pusat kluster ke-*k*:  $v_{kj}$ , di mana  $k = 1, 2, \dots, c$  dan  $j = 1, 2, \dots, p$  dengan menggunakan persamaan sebagai berikut:
 
$$v_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^m x_{ij})}{\sum_{i=1}^n (\mu_{ik})^m} \tag{1}$$
 di mana,
    - $\mu_{ik}$  : Nilai keanggotaan pada data sampel ke-*i* dan kluster ke-*k*
    - $x_{ij}$  : Data sampel ke-*i* dan variabel ke-*j*
    - m* : Pangkat
  - d. Menghitung fungsi objektif pada iterasi ke-*t*, *P*<sub>*t*</sub> dengan menggunakan persamaan sebagai berikut:

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left( \left[ \sum_{j=1}^p (x_{ij} - v_{kj})^2 \right] (\mu_{ik})^m \right) \quad (2)$$

di mana,

$x_{ij}$  : Data sampel ke- $i$  dan variabel ke- $j$

$v_{kj}$  : Pusat kluster ke- $k$  dan variabel ke- $j$

$\mu_{ik}$  : Nilai keanggotaan pada data sampel ke- $i$  dan kluster ke- $k$

$m$  : Pangkat

- e. Menghitung perubahan matriks keanggotaan dengan menggunakan persamaan sebagai berikut:

$$\mu_{ik} = \frac{\left[ \sum_{j=1}^p (x_{ij} - v_{kj})^2 \right]^{-\frac{1}{m-1}}}{\sum_{k=1}^c \left[ \sum_{j=1}^p (x_{ij} - v_{kj})^2 \right]^{-\frac{1}{m-1}}} \quad (3)$$

di mana,

$x_{ij}$  : Data sampel ke- $i$  dan variabel ke- $j$

$v_{kj}$  : Pusat kluster ke- $k$  dan variabel ke- $j$

$m$  : Pangkat

- f. Mengulang langkah c sampai dengan langkah e hingga  $(|P_t - P_{t-1}| < \epsilon)$  atau  $t > MaxIter$  terpenuhi.  
g. Mengulangi langkah b sampai dengan langkah f untuk kluster yang berbeda.  
h. Menghitung indeks validitas *Partition Coefficient* (PC) untuk setiap kluster dengan menggunakan persamaan sebagai berikut:

$$PC(c) = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^c (\mu_{ik})^2 \quad (4)$$

di mana,  $\mu_{ik}$  adalah nilai keanggotaan akhir pada data sampel ke- $i$  dan kluster ke- $k$ .

2. Menginterpretasikan hasil kluster yang optimal.

### Hasil dan Pembahasan

1. Metode Pengelompokan *Fuzzy C-Means* (FCM)

- a. Penentuan Nilai-Nilai Parameter Metode FCM

Sebelum melakukan proses pengelompokan menggunakan metode *Fuzzy C-Means* (FCM), langkah awal yang perlu dilakukan adalah menentukan banyak kluster yang akan digunakan, pangkat ( $m$ ), maksimum iterasi ( $MaxIter$ ), dan *error* terkecil yang diharapkan ( $\epsilon$ ). Pada penelitian ini menggunakan banyak kluster 3 sampai dengan 6,  $m = 2$ ,  $MaxIter = 1000$ ,  $\epsilon = 10^{-5}$ . Pada penelitian ini contoh perhitungan dilakukan dengan menggunakan banyak kluster  $c = 3$ .

- b. Membangkitkan Bilangan *Random*

Setelah menentukan nilai-nilai parameter, langkah selanjutnya dalam proses pengelompokan menggunakan metode FCM adalah melakukan pembangkitan bilangan

*random*  $\mu_{ik}$ , sebagai elemen-elemen matriks keanggotaan awal  $U$  berukuran  $i \times k$ . Nilai keanggotaan awal dengan  $c = 3$  dapat dilihat pada Tabel 1.

**Tabel 1.** Nilai Keanggotaan Awal dengan  $c = 3$  Pada Metode FCM

Kabupaten/ Kota	Nilai Keanggotaan Awal		
	Kluster 1	Kluster 2	Kluster 3
Malinau	0,2832	0,0708	0,6460
Bulungan	0,4500	0,1200	0,4300
Tana Tidung	0,3409	0,1250	0,5341
⋮	⋮	⋮	⋮
Bontang	0,4528	0,3585	0,1887

- c. Menghitung Pusat Kluster

Setelah membangkitkan bilangan *random*  $\mu_{ik}$ , langkah selanjutnya adalah menghitung pusat kluster awal dengan menggunakan persamaan (1). Adapun contoh perhitungan pusat kluster awal dengan  $c = 3$  pada pusat kluster pertama adalah sebagai berikut:

$$\begin{aligned} v_{1,1} &= \frac{\sum_{i=1}^{56} ((\mu_{i1})^2 x_{i1})}{\sum_{i=1}^{56} (\mu_{i1})^2} \\ &= \frac{((0,2832)^2 \times 52,61) + \dots + ((0,4528)^2 \times 45,99)}{(0,2832)^2 + \dots + (0,4528)^2} \\ &= \frac{397,6264}{8,2877} \\ &= 47,9777 \end{aligned}$$

⋮

$$\begin{aligned} v_{1,18} &= \frac{\sum_{i=1}^{56} ((\mu_{i1})^2 x_{i18})}{\sum_{i=1}^{56} (\mu_{i1})^2} \\ &= \frac{((0,2832)^2 \times 6,80) + \dots + ((0,4528)^2 \times 9,79)}{(0,2832)^2 + \dots + (0,4528)^2} \\ &= \frac{75,8760}{8,2877} \\ &= 9,1552 \end{aligned}$$

Dilakukan perhitungan yang sama pada pusat kluster kedua dan ketiga untuk mendapatkan pusat kluster selanjutnya. Sehingga pusat kluster yang telah diperbarui dapat dilihat pada Tabel 2.

Berdasarkan Tabel 2, dapat diketahui bahwa nilai 47,9777 menunjukkan nilai pusat kluster pada variabel pertama di kluster pertama, nilai 76,3065 menunjukkan nilai pusat kluster pada variabel kedua di kluster pertama, nilai 99,2036 menunjukkan nilai pusat kluster pada variabel ketiga di kluster pertama dan demikian seterusnya hingga pada variabel kedelapan belas di kluster pertama. Nilai 46,6698 menunjukkan nilai pusat kluster pada variabel pertama di kluster kedua, nilai 80,3315 menunjukkan nilai pusat kluster pada variabel kedua di kluster kedua, nilai 99,3027 menunjukkan nilai pusat kluster pada variabel ketiga di kluster kedua dan demikian seterusnya hingga pada variabel kedelapan belas di kluster kedua. Nilai 48,0806 menunjukkan nilai pusat kluster pada variabel pertama di kluster ketiga, nilai 75,0502 menunjukkan nilai pusat kluster pada variabel kedua di kluster ketiga, nilai

99,1425 menunjukkan nilai pusat kluster pada variabel ketiga di kluster ketiga dan demikian seterusnya hingga pada variabel kedelapan belas di kluster ketiga.

**Tabel 2.** Elemen Pusat Kluster Awal dengan  $c = 3$  Pada Metode FCM

Variabel	Pusat Kluster		
	1	2	3
$X_1$ (Angka Beban Ketergantungan)	47,9777	46,6698	48,0806
$X_2$ (Kepemilikan Akta Kelahiran)	76,3065	80,3315	75,0502
$X_3$ (APS 7-12 Tahun)	99,2036	99,3027	99,1425
$X_4$ (APS 13-15 Tahun)	94,8812	94,3332	95,5477
$X_5$ (APS 15-18 Tahun)	73,1282	70,6137	71,3773
$X_6$ (APM SD)	97,8927	98,3259	97,2434
$X_7$ (APM SMP)	75,8461	75,0362	76,8114
$X_8$ (APM SMA)	60,1163	55,9807	57,5401
$X_9$ (Keluhan Kesehatan)	28,9691	26,9891	26,4134
$X_{10}$ (Wanita Menurut Umur Perkawinan Pertama)	16,9162	18,3405	16,9312
$X_{11}$ (Penggunaan Alat/Cara KB)	58,4461	60,1076	59,0451
$X_{12}$ (Fasilitas Tempat Buang Air Besar)	86,3095	84,0470	85,4499
$X_{13}$ (Air Minum Bersih)	69,8684	66,0559	64,3689
$X_{14}$ (Sumber Penerangan Listrik)	98,0827	98,3240	97,7434
$X_{15}$ (Penggunaan Internet)	54,3029	52,2382	51,5337
$X_{16}$ (Korban Kejahatan)	0,8425	0,8633	0,8220
$X_{17}$ (PIP)	7,3690	5,7847	6,1362
$X_{18}$ (BPNT)	9,1552	8,9897	8,3464

**d. Menghitung Fungsi Objektif**

Setelah menghitung pusat kluster awal, selanjutnya nilai keanggotaan pada Tabel 1 dan pusat kluster awal pada Tabel 2 digunakan untuk menghitung nilai fungsi objektif pada  $t = 1$  dengan  $c = 3$  dengan menggunakan persamaan (2), hasil perhitungannya adalah sebagai berikut:

$$\begin{aligned} & \sum_{j=1}^{18} (x_{1j} - v_{1j})^2 \\ &= (52,61 - 47,9777)^2 \\ &+ (71,21 - 76,3065)^2 + \dots \\ &+ (6,80 - 9,1552)^2 \\ &= 752,7081 \\ &\vdots \\ & \sum_{j=1}^{18} (x_{56j} - v_{3j})^2 \\ &= (45,99 - 48,0806)^2 \\ &+ (84,03 - 75,0502)^2 + \dots \\ &+ (9,79 - 8,3464)^2 \\ &= 2.416,0125 \end{aligned}$$

$$\begin{aligned} P_t &= \sum_{i=1}^{56} \sum_{k=1}^3 \left( \left[ \sum_{j=1}^{18} (x_{ij} - v_{kj})^2 \right] \mu_{ik}^2 \right) \\ &= (752,7081 \times (0,2832)^2) \\ &+ (1.157,7544 \times (0,0708)^2) \\ &+ (967,2972 \times (0,6460)^2) + \dots \\ &+ (2.416,0125 \times (0,1887)^2) \\ &= 29.863,6750 \end{aligned}$$

Nilai fungsi objektif pada  $t = 1$  sebesar 29.863,6750. Nilai fungsi objektif awal ( $P_0$ ) adalah 0 sehingga  $|P_1 - P_0| = 29.863,6750 > \epsilon = 10^{-5}$  karena perubahan fungsi objektif masih lebih besar dari nilai  $\epsilon$ , maka proses dilanjutkan ke iterasi berikutnya. Iterasi berhenti ketika  $|P_t - P_{t-1}| < 0,00001$  atau  $t > 1000$ .

**e. Menghitung Perubahan Matriks Keanggotaan**

Setelah menghitung nilai fungsi objektif pada  $t = 1$ , langkah selanjutnya adalah menghitung perubahan nilai keanggotaan dengan  $c = 3$  dengan menggunakan persamaan (3). Contoh perhitungan menggunakan data nilai keanggotaan pertama (Kabupaten Malinau) sebagai berikut:

$$\begin{aligned} \mu_{1,1} &= \frac{[\sum_{j=1}^{18} (x_{1j} - v_{1j})^2]^{-\frac{1}{2-1}}}{\sum_{k=1}^3 [\sum_{j=1}^{18} (x_{1j} - v_{kj})^2]^{-\frac{1}{2-1}}} \\ &= \frac{(752,7081)^{-1}}{(752,7081)^{-1} + (1.157,7544)^{-1} + (967,2972)^{-1}} \\ &= 0,4118 \\ \mu_{1,2} &= \frac{[\sum_{j=1}^{18} (x_{1j} - v_{2j})^2]^{-\frac{1}{2-1}}}{\sum_{k=1}^3 [\sum_{j=1}^{18} (x_{1j} - v_{kj})^2]^{-\frac{1}{2-1}}} \\ &= \frac{(1.157,7544)^{-1}}{(752,7081)^{-1} + (1.157,7544)^{-1} + (967,2972)^{-1}} \\ &= 0,2677 \\ \mu_{1,3} &= \frac{[\sum_{j=1}^{18} (x_{1j} - v_{3j})^2]^{-\frac{1}{2-1}}}{\sum_{k=1}^3 [\sum_{j=1}^{18} (x_{1j} - v_{kj})^2]^{-\frac{1}{2-1}}} \\ &= \frac{(967,2972)^{-1}}{(752,7081)^{-1} + (1.157,7544)^{-1} + (967,2972)^{-1}} \\ &= 0,3205 \end{aligned}$$

Perhitungan perubahan nilai keanggotaan dengan  $c = 3$  dilakukan pada data nilai keanggotaan kedua (Kabupaten Bulungan) hingga data nilai keanggotaan ke-56 (Kota Bontang) untuk setiap kluster. Hasil nilai keanggotaan yang telah diperbarui dapat dilihat pada Tabel 3.

Setelah menghitung perubahan matriks keanggotaan, langkah selanjutnya adalah menghitung kembali pusat kluster, fungsi objektif dengan pusat kluster yang telah diperbarui dan menghitung perubahan matriks kembali dengan  $c = 3$ .

**Tabel 3.** Nilai Keanggotaan yang Telah Diperbarui dengan  $c = 3$  Pada Metode FCM

Kabupaten/ Kota	Nilai Keanggotaan		
	Klaster 1	Klaster 2	Klaster 3
Malinau	0,4118	0,2677	0,3205
Bulungan	0,4408	0,2442	0,3150
Tana Tidung	0,3196	0,3288	0,3516
⋮	⋮	⋮	⋮
Bontang	0,3917	0,3103	0,2980

Iterasi berhenti ketika  $|P_t - P_{t-1}| < 0,00001$  atau  $t > 1000$ . Pada penelitian ini langkah berhenti pada iterasi ke-32. Diperoleh hasil akhir nilai keanggotaan 56 kabupaten/kota yang dapat dilihat pada Tabel 4.

**Tabel 4.** Nilai Keanggotaan Akhir dengan  $c = 3$  Pada Metode FCM

Kab/ Kota	Nilai Keanggotaan			Klaster yang Diikuti
	1	2	3	
Malinau	0,1719	0,7543	0,0738	2
Bulungan	0,2654	0,6390	0,0956	2
Tana Tidung	0,4029	0,2774	0,3196	1
⋮	⋮	⋮	⋮	⋮
Bontang	0,1645	0,7654	0,0701	2

Berdasarkan nilai keanggotaan akhir pada Tabel 4, pada data pertama (Kabupaten Malinau) mempunyai nilai keanggotaan sebesar 0,1719 pada klaster pertama, 0,7543 pada klaster kedua, dan 0,0738 pada klaster ketiga. Diperoleh nilai keanggotaan terbesar yaitu 0,7543 sehingga data pertama paling tepat menjadi anggota klaster kedua. Data kedua (Kabupaten Bulungan) mempunyai nilai keanggotaan sebesar 0,2654 pada klaster pertama, 0,6390 pada klaster kedua, dan 0,0956 pada klaster ketiga. Diperoleh nilai keanggotaan terbesar yaitu 0,6390 sehingga data kedua paling tepat menjadi anggota klaster kedua. Data ketiga (Kabupaten Tana Tidung) mempunyai nilai keanggotaan sebesar 0,4029 pada klaster pertama, 0,2774 pada klaster kedua, dan 0,3196 pada klaster ketiga. Diperoleh nilai keanggotaan terbesar yaitu 0,4029 sehingga data ketiga paling tepat menjadi anggota klaster pertama. Penentuan klaster yang diikuti dilakukan dengan cara yang sama pada data selanjutnya. Selain perhitungan pada klaster  $c = 3$ , dilakukan juga perhitungan dengan langkah yang sama terhadap banyak klaster 4 sampai dengan 6.

f. Menghitung Indeks Validitas

Pada penelitian ini dilakukan validitas menggunakan *Partition Coefficient* (PC). Nilai keanggotaan akhir pada Tabel 4 digunakan untuk menghitung nilai PC.

Nilai PC dihitung dengan menggunakan persamaan (4) di mana nilai PC yang semakin besar menandakan kualitas klaster yang didapatkan semakin baik. Contoh perhitungan

nilai PC menggunakan  $c = 3$  adalah sebagai berikut:

$$\begin{aligned}
 PC(3) &= \frac{1}{56} \sum_{i=1}^{56} \sum_{k=1}^3 (\mu_{ik})^2 \\
 &= \frac{1}{56} ((0,1719)^2 + (0,7543)^2 + \dots + (0,0701)^2) \\
 &= 0,5096
 \end{aligned}$$

Dari hasil perhitungan didapatkan nilai PC dengan  $c = 3$  sebesar 0,5096.

Adapun hasil perhitungan nilai indeks validitas untuk  $c = 3$  sampai dengan 6 dapat dilihat pada Tabel 5.

**Tabel 5.** Nilai Indeks Validitas untuk Keseluruhan Parameter  $c$  Pada Metode FCM

Banyak Klaster ( $c$ )	PC
3	<b>0,5096</b>
4	0,3910
5	0,3260
6	0,3051

Berdasarkan Tabel 5, diperoleh nilai PC dari banyak klaster 3 sampai dengan 6. Nilai PC terbesar adalah 0,5096 pada banyak klaster 3. Sehingga diperoleh klaster yang optimal adalah 3 klaster.

Adapun anggota klaster berdasarkan klaster yang optimal adalah sebagai berikut:

Klaster 1: Tana Tidung, Nunukan, Kotawaringin Barat, Kotawaringin Timur, Barito Selatan, Barito Utara, Sukamara, Lamandau, Seruyan, Katingan, Pulang Pisau, Gunung Mas, Barito Timur, Tanah Laut, Kotabaru, Banjar, Barito Kuala, Tapin, Hulu Sungai Selatan, Hulu Sungai Tengah, Hulu Sungai Utara, Balangan, Kutai Barat, dan Mahakam Ulu.

Klaster 2: Malinau, Bulungan, Tarakan, Singkawang, Palangka Raya, Tabalong, Tanah Bumbu, Banjarmasin, Banjarbaru, Paser, Kutai Kartanegara, Kutai Timur, Berau, Penajam Paser Utara, Balikpapan, Samarinda, dan Bontang.

Klaster 3: Sambas, Bengkayang, Landak, Mempawah, Sanggau, Ketapang, Sintang, Kapuas Hulu, Sekadau, Melawi, Kayong Utara, Kubu Raya, Pontianak, Kapuas, dan Murung Raya.

2. Interpretasi Hasil Klaster

Setelah menentukan klaster yang optimal dan anggota yang terbentuk, maka selanjutnya adalah menginterpretasi hasil klaster. Secara deskriptif, ukuran yang biasa dipakai untuk proses interpretasi adalah klaster *centroid* yaitu rata-rata nilai objek yang terdapat dalam klaster pada setiap variabel.

Berdasarkan hasil metode FCM diperoleh nilai rata-rata variabel untuk masing-masing klaster yang dapat dilihat pada Tabel 6.

**Tabel 6.** Nilai Rata-Rata Variabel Berdasarkan Metode FCM untuk Masing-Masing Klaster

Variabel	Klaster ke-k		
	1	2	3
$X_1$ (Angka Beban Ketergantungan)	47,2921	46,3638	49,6742
$X_2$ (Kepemilikan Akta Kelahiran)	83,9984	73,5050	72,5433
$X_3$ (APS 7-12 Tahun)	99,4376	99,4424	98,6983
$X_4$ (APS 13-15 Tahun)	94,9102	97,1089	92,2116
$X_5$ (APS 15-18 Tahun)	69,1847	78,7685	67,2741
$X_6$ (APM SD)	98,4664	97,7550	97,6680
$X_7$ (APM SMP)	77,4537	79,3619	68,7562
$X_8$ (APM SMA)	57,2431	65,7429	50,4092
$X_9$ (Keluhan Kesehatan)	29,0017	25,9661	26,8256
$X_{10}$ (Wanita Menurut Umur Perkawinan Pertama)	20,6562	14,2408	16,9032
$X_{11}$ (Penggunaan Alat/Cara KB)	63,5140	53,1845	62,0974
$X_{12}$ (Fasilitas Tempat Buang Air Besar)	83,6202	92,0828	79,2398
$X_{13}$ (Air Minum Bersih)	68,7031	86,8761	43,0711
$X_{14}$ (Sumber Penerangan Listrik)	98,5926	99,4911	94,8149
$X_{15}$ (Penggunaan Internet)	49,7222	64,1533	40,7320
$X_{16}$ (Korban Kejahatan)	0,7390	1,0637	0,6159
$X_{17}$ (PIP)	6,4075	5,2239	7,6785
$X_{18}$ (BPNT)	8,8510	8,2771	9,3613

Berdasarkan Tabel 6, dapat dilihat rata-rata setiap variabel berbeda dari ketiga klaster yang terbentuk. Dari variabel angka beban ketergantungan, keluhan kesehatan, wanita menurut umur perkawinan pertama, penggunaan alat/cara KB, PIP, dan BPNT pada klaster kedua memiliki rata-rata variabel yang lebih rendah dibandingkan pada klaster pertama dan ketiga. Sedangkan pada variabel kepemilikan akta kelahiran, APS, APM, fasilitas tempat buang air besar, air minum bersih, sumber penerangan listrik, penggunaan internet, dan korban kejahatan pada klaster ketiga memiliki rata-rata variabel yang lebih rendah dibandingkan pada klaster pertama dan kedua.

**Kesimpulan**

Berdasarkan hasil indeks validitas PC pada metode FCM, jumlah klaster yang optimal untuk pengelompokan kabupaten/kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat adalah sebanyak 3 klaster. Adapun anggota pada klaster 1 berjumlah 24 kabupaten/kota yaitu Tana Tidung, Nunukan, Kotawaringin Barat, Kotawaringin Timur, Barito Selatan, Barito Utara, Sukamara, Lamandau, Seruyan, Katingan, Pulang Pisau, Gunung Mas, Barito Timur, Tanah Laut, Kotabaru, Banjar, Barito Kuala, Tapin, Hulu Sungai Selatan, Hulu Sungai Tengah, Hulu Sungai Utara, Balangan, Kutai Barat, dan Mahakam Ulu.

Anggota pada klaster 2 berjumlah 17 kabupaten/kota yaitu Malinau, Bulungan, Tarakan, Singkawang, Palangka Raya, Tabalong, Tanah Bumbu, Banjarmasin, Banjarbaru, Paser, Kutai Kartanegara, Kutai Timur, Berau, Penajam Paser Utara, Balikpapan, Samarinda, dan Bontang. Sedangkan anggota pada klaster 3 berjumlah 15 kabupaten/kota yaitu Sambas, Bengkayang, Landak, Mempawah, Sanggau, Ketapang, Sintang, Kapuas Hulu, Sekadau, Melawi, Kayong Utara, Kubu Raya, Pontianak, Kapuas, dan Murung Raya.

Adapun pada anggota klaster 2, pemerintah dapat memperhatikan variabel angka beban ketergantungan, keluhan kesehatan, wanita menurut umur perkawinan pertama, penggunaan alat/cara KB, PIP, dan BPNT dalam menentukan tindakan lebih lanjut karena pada klaster kedua variabel-variabel tersebut memiliki rata-rata yang lebih rendah dibandingkan klaster pertama dan ketiga. Pada anggota klaster 3, pemerintah dapat memperhatikan variabel kepemilikan akta kelahiran, APS, APM, fasilitas tempat buang air besar, air minum bersih, sumber penerangan listrik, penggunaan internet, dan korban kejahatan dalam menentukan tindakan lebih lanjut karena pada klaster ketiga variabel-variabel tersebut memiliki rata-rata yang lebih rendah dibandingkan klaster pertama dan kedua. Sedangkan pada anggota klaster 1, pemerintah dapat memperhatikan variabel angka beban ketergantungan, APS, APM SMP, APM SMA, fasilitas tempat buang air besar, air minum bersih, sumber penerangan listrik, penggunaan internet, korban kejahatan, PIP, dan BPNT dalam menentukan tindakan lebih lanjut karena pada klaster pertama variabel-variabel tersebut memiliki rata-rata yang masih terbilang rendah dibandingkan klaster lain yang memiliki rata-rata yang lebih tinggi.

**Daftar Pustaka**

Alwi, W., & Hasrul, M. (2018). Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat. *Jurnal MSA*, 6(1), 35-42.

Basri, F., & Munandar, H. (2009). *Lanskap Ekonomi Indonesia*. Jakarta: Kencana.

Bezdek, J., & Dunn, J. (1975). Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions. *IEEE Transactions on Computers*, 835-838.

Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. New Jersey: Prentice-Hall Inc.

- Kusumadewi, S., & Purnomo, H. (2010). *Aplikasi Logika Fuzzy untuk Pendukung Keputusan Edisi 2*. Yogyakarta: Graha Ilmu.
- Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi menggunakan Matlab*. Yogyakarta: Andi Offset.
- Prasetyo, S., Mustafid, & Hakim, A. (2020). Penerapan Fuzzy C-Means Kluster untuk Segmentasi Pelanggan E-Commerce dengan Metode Recency Frequency Monetary (RFM). *Jurnal Gaussian*, 421-433.
- Setiawan, R., Ernawati, & Efendi, R. (2018). Klasifikasi Kawasan Permukiman Tingkat Kelurahan Untuk Pembangunan Sistem Berbasis Data Kualitas Permukiman (Studi Kasus: 67 Kelurahan di Kota Bengkulu). *Jurnal Pseudocode*, 5(1), 45-55.
- Tan, P., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Boston: Pearson Education.
- Wang, W., & Zhang, Y. (2007). On Fuzzy Cluster Validity Indices. *Fuzzy Sets System*, 158, 2095-2117.
- Widarjono, A. (2015). *Analisis Multivariat Terapan dengan Program SPSS, AMOS, dan SMARTPLS Edisi Kedua*. Yogyakarta: UPM STIM YKPN.
- Zhao, Q., & Franti, P. (2014). WB-index : A Sum-of-Squares Based Index for Cluster Validity. *Data & Knowledge Engineering Elsevier B.V*, 77-89.