

Grouping Districts/Cities in Kalimantan Island Based on The People's Welfare Indicators Using Fuzzy C-Means and Subtractive Fuzzy C-Means Methods

Pengelompokan Kabupaten/Kota di Pulau Kalimantan Berdasarkan Indikator Kesejahteraan Rakyat Menggunakan Metode *Fuzzy C-Means* dan *Subtractive Fuzzy C-Means*

Nur Annisa Fitri^{1*}, Memi Nor Hayati^{2*}, Rito Goejantoro^{3*}

Abstract

Cluster analysis has the aim of grouping several objects of observation based on the data found in the information to describe the objects and their relationships. The grouping method used in this research is the Fuzzy C-Means (FCM) and Subtractive Fuzzy C-Means (SFCM) methods. The two grouping methods were applied to the people's welfare indicator data in 42 regencies/cities on the island of Kalimantan. The purpose of this study was to obtain the results of grouping districts/cities on the island of Kalimantan based on indicators of people's welfare and to obtain the results of a comparison of the FCM and SFCM methods. Based on the results of the analysis, the FCM and SFCM methods yield the same conclusions, so that in this study the FCM and SFCM methods are both good to use in classifying districts/cities on the island of Kalimantan based on people's welfare indicators and produce an optimal cluster of two clusters, namely the first cluster consisting of 10 Regencies/Cities on the island of Kalimantan, while the second cluster consists of 32 districts/cities on the island of Borneo.

Keywords: FCM, People's Welfare Indicators, SFCM, validity index

Abstrak

Analisis kluster mempunyai tujuan untuk mengelompokkan beberapa objek pengamatan berdasarkan data yang ditemukan dalam informasi untuk menggambarkan objek dan hubungannya. Metode pengelompokan yang digunakan dalam penelitian ini adalah metode *Fuzzy C-Means* (FCM) dan *Subtractive Fuzzy C-Means* (SFCM). Dua metode pengelompokan tersebut diterapkan pada data indikator kesejahteraan rakyat pada 42 Kabupaten/Kota di Pulau Kalimantan. Tujuan dari penelitian ini adalah untuk mendapatkan hasil pengelompokan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat serta untuk memperoleh hasil perbandingan metode FCM dan SFCM. Berdasarkan hasil analisis, metode FCM dan SFCM menghasilkan kesimpulan yang sama sehingga pada penelitian ini metode FCM dan SFCM

**Laboratorium Statistika Terapan dan Laboratorium Statistika Komputasi, Fakultas MIPA, Universitas Mulawarman*

Email Address: nurannisafitri200@gmail.com¹, meminorhayati@fmipa.unmul.ac.id², rito.goejantoro@fmipa.unmul.ac.id³



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

sama-sama baik untuk digunakan dalam mengelompokkan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat dan menghasilkan kluster optimal sebanyak dua kluster yaitu kluster pertama beranggotakan 10 Kabupaten/Kota di Pulau Kalimantan sedangkan kluster kedua beranggotakan 32 Kabupaten/Kota di Pulau Kalimantan.

Kata kunci: FCM, Indeks Validitas, Indikator Kesejahteraan Rakyat, SFCM.

1. PENDAHULUAN

1.1 Latar Belakang

Data mining adalah istilah yang digunakan untuk menggambarkan penemuan pengetahuan dalam database. Data mining adalah proses mengekstraksi dan mengidentifikasi informasi dan pengetahuan yang relevan dari sejumlah database dengan menggunakan teknik statistik, matematika, buatan dan pembelajaran mesin [10]. Teknik data mining dapat mengidentifikasi pola tersembunyi dan dapat digunakan secara efektif dalam klasifikasi, estimasi, prediksi, aturan asosiasi, *clustering* dan visualisasi [4].

Analisis statistika yang membagi objek pengamatan kedalam beberapa kelompok menurut variabel yang diamati disebut analisis kluster [13]. Dalam analisis kluster objek data yang dengan karakteristik yang sama akan dikelompokkan pada satu kelompok dan data dengan karakteristik yang berbeda akan dikelompokkan ke dalam kelompok lain [14]. Terkadang terdapat objek sebenarnya berada di antara dua kluster atau lebih, sehingga menggunakan metode *fuzzy clustering* [6]. Salah satu algoritma *fuzzy clustering* adalah *Fuzzy C-Means* [7].

Fuzzy C-Means merupakan metode pengelompokan yang banyak digunakan berdasarkan fungsi tujuan, yang telah digunakan dalam berbagai aplikasi [16]. FCM adalah sebuah metode dimana diasumsikan setiap titik data berasosiasi dengan semua kluster dengan nilai fungsi keanggotaan *fuzzy*. FCM dilakukan dengan meminimalkan fungsi objektif [11]. Metode FCM mempunyai kelemahan, diantaranya membutuhkan kelompok yang banyak dan nilai keanggotaan ditentukan sebelumnya [8].

Jika jumlah kluster tidak diketahui, metode *Subtractive Clustering* (SC) merupakan alternatif dari metode kluster lainnya. Metode SC mendapatkan hasil lebih konsisten daripada metode FCM [3]. Metode SC juga mempunyai kecepatan yang lebih baik dibanding metode FCM, akan tetapi akurasi pada metode SC lebih rendah dibandingkan metode FCM [5].

Subtractive Fuzzy C-Means (SFCM) merupakan metode *hybrid* yang menghubungkan kekurangan dan kelebihan dari metode FCM dan metode SC. Dalam penelitian [9] memperoleh kesimpulan bahwa secara umum metode SFCM dapat memberikan solusi lebih baik daripada metode FCM dan tingkat kecepatan yang diberikan lebih tinggi mengenai konvergensi fungsi objektif.

Sebagaimana tercantum dalam alinea IV Pembukaan UUD 1945, kesejahteraan rakyat merupakan salah satu tujuan negara. Dalam melaksanakan rencana pembangunan, pemerintah perlu menetapkan karakteristik yang sesuai dengan tingkat kesejahteraan rakyat dimasing-masing daerah, sehingga dapat sejalan dengan tujuan negara dalam merumuskan kebijakan dan strategi pembangunan [1].

Berdasarkan latar belakang tersebut, maka peneliti mengelompokkan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat menggunakan metode *fuzzy c-means* dan *subtractive fuzzy c-means* sehingga hasil pengelompokan tersebut dapat membantu

pemerintah untuk mempertimbangkan dalam membuat kebijakan dalam meningkatkan kesejahteraan rakyat di Kabupaten/Kota khususnya Pulau Kalimantan.

1.2 Fuzzy C-Means (FCM)

Konsep dasar dari metode FCM adalah menentukan pusat kluster sebagai rata-rata posisi setiap kluster, dimana pusat kluster masih belum akurat. Secara berulang kali memperbaiki pusat kluster serta nilai keanggotaan setiap titik data, sehingga pusat kluster bergerak mendekati lokasi yang tepat. Adapun algoritma dari metode FCM adalah sebagai berikut [7]:

1. Menentukan banyak kluster (c), pangkat (m), fungsi objektif awal (P_0), *error* terkecil yang diharapkan (ε), maksimum iterasi (*MaxIter*).
2. Membangkitkan bilangan acak m_{ik} , $i = 1, 2, \dots, n$; $k = 1, 2, \dots, c$ sebagai nilai keanggotaan awal U . Kemudian matriks partisi awal U dinormalisasikan menggunakan persamaan:

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (1.1)$$

Setelah itu hitung elemen matriks yang ternormalisasi dengan persamaan:

$$\mu_{iknorm} = \frac{\mu_{ik}}{Q_i} \quad (1.2)$$

3. Menghitung pusat kluster ke- k : v_{kj} , dengan $k = 1, 2, \dots, c$ dan $j = 1, 2, \dots, p$.

$$v_{kj} = \frac{\sum_{i=1}^n ((\mu_{iknorm})^m x_{ij})}{\sum_{i=1}^n (\mu_{iknorm})^m} \quad (1.3)$$

4. Menghitung nilai fungsi objektif pada iterasi ke- t .

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^p (x_{ij} - v_{kj})^2 \right] (\mu_{iknorm})^m \right) \quad (1.4)$$

5. Menghitung perubahan nilai keanggotaan:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^p (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{m-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^p (x_{ij} - v_{kj})^2 \right]^{\frac{-1}{m-1}}} \quad (1.5)$$

6. Mengulangi langkah ke-3 sampai dengan langkah ke-5 hingga kondisi ($|P_t - P_{t-1}| < \varepsilon$) atau $t > \text{MaxIter}$ terpenuhi

1.3 Subtractive Clustering (SC)

Metode SC memiliki konsep dasar yaitu untuk menentukan daerah dimana suatu variabel mempunyai potensi paling tinggi terhadap data sekitarnya. Data yang mempunyai tetangga paling tinggi akan menjadi pusat kluster lalu dikurangi potensinya, dan kemudian akan dipilih kembali titik lain dengan tetangga paling banyak sebagai pusat kluster berikutnya, langkah tersebut berulang kali akan sampai semua data teruji. Adapun algoritma dari metode SC adalah sebagai berikut [7]:

1. Menginput data yang diklusterkan: x_{ij} , dengan $i = 1, 2, \dots, n$; dan $j = 1, 2, \dots, p$.

2. Menentukan nilai jari-jari untuk setiap variabel data (r_j), *accept ratio* (ar), *squash factor* (q), *reject ratio* (rr), nilai Maksimum dan minimum data.
3. Menormalisasikan data dengan menggunakan persamaan:

$$x_{ijnorm} = \frac{x_{ij} - x_{jmin}}{x_{jmax} - x_{jmin}} \quad (1.6)$$

4. Menentukan nilai potensi awal untuk setiap data ($Y_f; i = 1, 2, \dots, n$).
 - a. Menghitung jarak T_j untuk setiap data dengan persamaan:

$$Dist_{ij} = \frac{T_j - x_{ijnorm}}{r}; i, f = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (1.7)$$

di mana $T_j = x_{jnorm}; j = 1, 2, \dots, p$

- b. Menentukan potensial awal tiap data dengan persamaan:

$$\text{Jika } m = 1, \text{ maka } Y_f = \sum_{i=1}^n e^{-4(Dist_{ij}^2)} \quad (1.8)$$

$$\text{Jika } m > 1, \text{ maka } Y_f = \sum_{i=1}^n e^{-4(\sum_{j=1}^p Dist_{ij}^2)} \quad (1.9)$$

5. Mencari data dengan potensi tertinggi:
 - a. $M = \max [Y_f | i = 1, 2, \dots, n]$; untuk iterasi ke-1 (1.10)
 - b. $Z = \max [Y_f | i = 1, 2, \dots, n]$; untuk iterasi ke-2,3, dan seterusnya (1.11)
6. Menentukan pusat kluster, kemudian potensinya akan dikurangi dengan titik-titik di sekitarnya dan menghitung nilai rasionya.

$$R = \frac{Z}{M} \quad (1.12)$$

Nilai $Z = M$ pada iterasi pertama

Setelah memperoleh nilai rasio, terdapat 3 kondisi:

- 1) Apabila nilai rasio $> \textit{accept ratio}$, maka pusat kluster (V_k) akan data diterima. Kemudian mengurangi potensi terhadap titik data yang lain menggunakan persamaan sebagai berikut:

$$\text{a) } S_{ij} = \frac{V_{kinorm} - x_{ijnorm}}{r \times q}; i = 1, 2, \dots, n; j = 1, 2, \dots, p \quad (1.13)$$

$$\text{b) } Y_{V_{kinorm}} = M \times e^{-4[\sum_{j=1}^p (S_{ij})^2]}; i = 1, 2, \dots, n \quad (1.14)$$

$$\text{c) } Y_i^t = Y_i^{t-1} - Y_{V_{kinorm}} \quad (1.15)$$

- 2) Apabila nilai rasio $< \textit{accept ratio}$ dan nilai rasio $> \textit{reject ratio}$, maka data dapat diterima jika jaraknya cukup jauh dari pusat kluster yang ada.

Langkah pengerjaan dalam kondisi ini adalah sebagai berikut:

- a) $Md = - 1$;
- b) Mengerjakan untuk $k = 1, 2, \dots, c$; dimana c =banyaknya kluster

$$Sd_k = \sum_{j=1}^p \left(\frac{I_j - V_{kjnorm}}{r} \right)^2; j=1,2,\dots,p \quad (1.16)$$

Jika ($Md < 0$) atau ($Sd < Md$), maka $Md = Sd_k$

Jika ($Sd_k > Md$), maka Md tidak berubah

$$Mds = \sqrt{Md} \quad (1.17)$$

Di mana Mds merupakan jarak terdekat data calon pusat kluster dengan pusat kluster.

Jika ($\text{rasio} + Mds \geq 1$), maka data diterima sebagai pusat kluster.

Jika ($\text{rasio} + Mds < 1$), maka data tidak diterima kemudian tidak dipertimbangkan kembali sebagai pusat kluster, dan data potensi ditetapkan menjadi 0.

3) Apabila nilai rasio $<$ *accept ratio* dan nilai rasio $<$ *reject ratio*, maka iterasi berhenti.

7. Mengembalikan pusat kluster ke bentuk aslinya (denormalisasi) dengan menggunakan persamaan:

$$v_{kj} = V_{kjnorm} \times (x_{jmax} - x_{jmin}) + x_{jmin} \quad (1.18)$$

1.4 Subtractive Fuzzy C-Means (SFCM)

Metode *subtractive Fuzzy C-Means* (SFCM) merupakan kombinasi dari metode SC dan metode FCM. Banyaknya kluster akan ditentukan dengan metode SC sedangkan metode FCM digunakan untuk nilai keanggotaan awal, sehingga tidak perlu membangkitkan bilangan secara acak. Metode FCM digunakan untuk memperoleh pusat kluster dan nilai keanggotaan untuk menentukan penempatan data ke dalam kluster. Algoritma metode SFCM adalah sebagai berikut [9]:

1. Menentukan jari-jari (r), *accept ratio* (ar), *squash factor* (q), *reject ratio* (rr), pangkat (m), fungsi objektif awal (P_0), maksimum iterasi ($MaxIter$), *error* terkecil yang diharapkan (ϵ).
2. Menormalisasi setiap data dengan menggunakan persamaan (1.6).
3. Menghitung potensi awal untuk setiap data yang telah ternormalisasi dengan persamaan (1.9).
4. Mencari titik dengan potensi tertinggi dengan menggunakan persamaan (1.10) dan (1.11).
5. Menentukan pusat kluster, kemudian potensinya akan dikurangi dengan titik-titik di sekitarnya dan menghitung nilai rasionya menggunakan persamaan (1.12) kemudian rasio dibandingkan dengan ar dan rr . Apabila nilai *ratio* $>$ ar , maka pusat kluster (V_k) akan data diterima. Setelah mendapatkan kluster baru selanjutnya adalah mengurangi potensi titik data yang lain. Apabila nilai *ratio* $<$ ar dan nilai *ratio* $>$ rr , maka data dapat diterima jika jaraknya cukup jauh dari pusat kluster yang ada. Apabila nilai *ratio* $<$ ar dan nilai *ratio* $<$ rr , maka iterasi berhenti.
6. Mengembalikan pusat kluster ke bentuk aslinya (denormalisasi) dengan persamaan (1.18).
7. Menghitung nilai keanggotaan awal FCM berdasarkan pusat kluster yang diperoleh dengan metode SC menggunakan persamaan (1.5).
8. Memperbarui pusat kluster dengan persamaan (1.3)
9. Menghitung nilai fungsi objektif pada iterasi ke- t dengan persamaan (1.4)
10. Menghitung perubahan nilai keanggotaan dengan persamaan (1.5)

11. Mengulangi langkah ke-8 sampai dengan langkah ke-10 hingga $|P_t - P_{t-1}| < \varepsilon$ atau $t > MaxIter$ terpenuhi

1.5 Indeks Validitas

Indeks validitas merupakan digunakan untuk menentukan mengukur banyak kluster yang optimal [15]. Indeks validitas yang umum digunakan adalah:

1. *Partition Coefficient (PC)*

Indeks PC menggunakan nilai keanggotaan sebagai evaluasi. Semakin tinggi nilai PC, maka semakin baik kluster yang dihasilkan dan dapat dihitung dengan menggunakan persamaan [12]:

$$PC(c) = \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^c (\mu_{ik})^2 \quad (1.19)$$

2. *Classification Entropy (CE)*

Menurut [2], Indeks CE digunakan untuk mengukur kekaburan dari kluster. Indeks CE mirip dengan indeks PC karena menggunakan nilai keanggotaan sebagai evaluasi. Persamaan indeks CE sebagai berikut:

$$CE(c) = -\frac{1}{n} \sum_{i=1}^n \sum_{k=1}^c \mu_{ik} \ln(\mu_{ik}) \quad (1.20)$$

Menurut [15], indeks CE memiliki rentang antara 0 hingga $\ln(c)$. Semakin rendah nilai CE, maka semakin baik pengelompokannya.

3. *Modified Partition Coefficient (MPC)*

Seiring perubahan nilai c , nilai PC dan nilai CE cenderung berubah secara monoton. Dave (1996) melakukan modifikasi nilai PC (MPC) terhadap kecenderungan monotonik tersebut. Persamaan indeks MPC adalah sebagai berikut:

$$MPC(c) = 1 - \frac{c}{c-1} (1 - PC) \quad (1.21)$$

Indeks MPC memiliki rentang antara 0 hingga 1, semakin besar nilai MPC (mendekati 1) akan menghasilkan kluster yang lebih baik [12].

2. HASIL DAN PEMBAHASAN

2.1 Metode Pengelompokan *Fuzzy C-Means (FCM)*

Pada penelitian ini menggunakan banyak kelompok 2 sampai dengan 8, $m = 2$, $MaxIter = 1000$, $\varepsilon = 10^{-5}$. Berikut hasil analisis dengan menggunakan persamaan (1.19), (1.20) dan (1.21) yang dapat dilihat pada Tabel 2.1 sebagai berikut :

Tabel 2.1 Nilai Indeks Validitas untuk Keseluruhan Parameter c Pada Metode FCM

Banyak Kluster (c)	2	3	4	5	6	7	8
PC	0,8887	0,8041	0,7229	0,7235	0,6435	0,6545	0,6340
CE	0,1918	0,3715	0,5246	0,5484	0,7126	0,7178	0,7967
MPC	0,7775	0,7062	0,6305	0,6544	0,5722	0,5969	0,5817

Berdasarkan Tabel 2.1 diperoleh nilai PC, CE, dan MPC dari banyak kluster 2 sampai dengan 8, nilai PC dan MPC terbesar berturut-turut adalah 0,8887 dan 0,7775 pada banyak kluster 2 sedangkan nilai CE terkecil adalah 0,1918 pada banyak kluster 2. Sehingga diperoleh kluster yang optimal adalah 2 kluster. Adapun anggota kluster dengan banyak kluster 2 pada metode FCM adalah sebagai berikut:

Kluster 1 : Bengkayang, Sanggau, Mempawah, Sekadau, Kapuas Hulu, Kayong Utara, Melawi, Tanah Laut, Singkawang, Banjar, Kota Baru, Tapin, Hulu Sungai Selatan, Barito Kuala, Tanah Bumbu, Tabalong, Banjarbaru, Balangan, Hulu Sungai Utara, Paser, Kutai Barat, Penajam Paser Utara, Berau, Mahakam Ulu, Bontang, Balikpapan, Malinau, Tana Tidung, Bulungan, Tarakan dan Nunukan.

Kluster 2 : Sambas, Landak, Ketapang, Sintang, Kubu Raya, Pontianak, Banjarmasin, Kutai Kartanegara, Kutai Timur dan Samarinda.

2.2 Metode Pengelompokan *Subtractive Fuzzy C-Means* (SFCM)

Pada penelitian ini menggunakan $r = 0,9400$; $r = 0,9000$; $r = 0,7500$; $r = 0,7100$; $r = 0,7000$; $r = 0,6900$ dan $r = 0,6880$; $q = 1,5000$; $ar = 0,7500$; $rr = 0,5000$; $MaxIter = 1000$, $m = 2$ dan $\varepsilon = 10^{-5}$. Berikut hasil analisis dengan menggunakan persamaan (1.19), (1.20) dan (1.21), yang dapat dilihat pada Tabel 2.2 sebagai berikut:

Tabel 2.2 Nilai Indeks Validitas untuk Keseluruhan Parameter c Pada Metode SFCM

Banyak Kluster (c)	2	3	4	5	6	7	8
r	0,9400	0,9000	0,7500	0,7100	0,7000	0,6900	0,6880
PC	0,889	0,804	0,723	0,724	0,643	0,654	0,633
CE	0,192	0,372	0,525	0,548	0,713	0,718	0,783
MPC	0,777	0,706	0,631	0,654	0,572	0,597	0,580

Berdasarkan Tabel 2.2 diperoleh nilai PC, CE, dan MPC dari banyak kluster 2 sampai dengan 8, nilai PC dan MPC terbesar berturut-turut adalah 0,8890 dan 0,7770 pada banyak kluster 2 sedangkan nilai CE terkecil adalah 0,1920 pada banyak kluster 2. Sehingga diperoleh kluster yang optimal adalah 2 kluster. Adapun anggota kluster dengan $r = 0,9400$ pada metode SFCM adalah sebagai berikut:

Kluster 1 : Bengkayang, Sanggau, Mempawah, Sekadau, Kapuas Hulu, Kayong Utara, Melawi, Tanah Laut, Singkawang, Banjar, Kota Baru, Tapin, Hulu Sungai Selatan, Barito Kuala, Tanah Bumbu, Tabalong, Banjarbaru, Balangan, Hulu Sungai Utara, Paser, Kutai Barat, Penajam Paser Utara, Berau, Mahakam Ulu, Bontang, Balikpapan, Malinau, Tana Tidung, Bulungan, Tarakan dan Nunukan.

Kluster 2 : Sambas, Landak, Ketapang, Sintang, Kubu Raya, Pontianak, Banjarmasin, Kutai Kartanegara, Kutai Timur dan Samarinda.

2.3 Interpretasi Hasil Kluster

Setelah dilakukan pengelompokan menggunakan dua metode *Fuzzy C-Means* (FCM) dan *Subtractive Fuzzy C-Means* (SFCM), kemudian dilakukan validasi kluster dengan melihat nilai

PC, CE dan MPC pada masing-masing hasil pengelompokan. Adapun hasil pengelompokan menggunakan metode FCM dan SFCM dapat dilihat pada Tabel 2.3 sebagai berikut:

Tabel 2.3 Nilai Indeks Validitas Metode FCM dan SFCM

Metode	Nilai PC	Nilai CE	Nilai MPC	Banyak Klaster
FCM	0,8887	0,1918	0,7775	2
SFCM	0,8887	0,1918	0,7775	2

Berdasarkan Tabel 2.3 dapat disimpulkan bahwa pengelompokan menggunakan metode FCM dan metode SFCM menghasilkan kesimpulan yang sama sehingga metode FCM dan metode SFCM pada penelitian ini sama-sama baik untuk digunakan.

3. KESIMPULAN

Berdasarkan hasil analisis data dan pembahasan, maka diambil kesimpulan sebagai berikut:

1. Berdasarkan hasil indeks validitas pada metode FCM jumlah klaster yang optimal untuk mengelompokkan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat adalah sebanyak 2 klaster. Adapun anggota pada klaster 1 berjumlah 32 Kabupaten/Kota sedangkan anggota pada klaster 2 berjumlah 10 Kabupaten/Kota.
2. Berdasarkan hasil indeks validitas pada metode SFCM jumlah klaster yang optimal untuk mengelompokkan Kabupaten/Kota di Pulau Kalimantan berdasarkan indikator kesejahteraan rakyat adalah sebanyak 2 klaster dengan $r=0,9400$. Adapun anggota pada klaster 1 berjumlah 32 Kabupaten/Kota sedangkan anggota pada klaster 2 berjumlah 10 Kabupaten/Kota.
3. pengelompokan menggunakan metode FCM dan metode SFCM menghasilkan kesimpulan yang sama sehingga pada penelitian ini metode FCM dan metode SFCM sama-sama baik untuk digunakan.

DAFTAR PUSTAKA

- [1] Alwi, W., Hasrul, Muh., 2018. Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat. *Jurnal MSA*, Vol.6, No.1, pp. 35-42.
- [2] Azar, A. T, El-Said, S. A., Hassanien, A. E., 2013. Fuzzy and Hard Clustering Analysis for Thyroid Disease. *Computer Methods and Programs in Biomedicine*, Vol.111, No.1, pp. 1-16.
- [3] Bataineh, K. M, Naji, M., Saqer, M., 2011. A Comparison Study Between Various Fuzzy Clustering Algorithms. *Jordan Journal of Mechanical and Industrial Engineerin*, Vol.5, No.4, pp. 335-343.
- [4] Dubey, A. K., Grupta, U., Jain, S., 2018. Comparative Study of K-Means and Fuzzy C-Means Algorithms on The Breast Cancer Data. *International Journal on Advanced Science Engineering Information Technology*, Vol. 8, No.1, pp. 18-29.

- [5] Hossen J., Rahman, A., Sayeed, S., Samsuddin, K., Rokhani, F., 2011. A Modified Hybrid Fuzzy Clustering Algorithm for Data Partitions. *Australian Journal of Basic and Applied Sciences*, Vol.5, No.8, pp. 674-681.
- [6] 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.
- [7] Kusumadewi, S, Purnomo, H, 2010. *Aplikasi Logika Fuzzy untuk Pendukung Keputusan Edisi 2*. Yogyakarta: Graha Ilmu.
- [8] Le, T., Altman, T., 2011. A new initialization method for the Fuzzy C-Means Algorithm using Fuzzy Subtractive Clustering. *Proc. International Conference on Information and Knowledge Engineering*, Vol.1, pp. 144-150.
- [9] Liu, W. Y., Xiao, C. J., Wang, B. W., Shi, Y., Fang, S. F., 2003. Study on Combining Subtractive Clustering with Fuzzy C-Means Clustering. *International Conference on Machine Learning and Cybernetics*, Vol.5, pp. 2659-2662.
- [10] Maselena., A., Sugiyarti, E., Jasmi, K. A., Basiron, B., Huda, M., Shankar K., 2018. Decision Support System of Scholarship Grantee Selection Using Data Mining. *Internasional Journal of Pure and Applied Mathematics*, Vol. 119, No. 15, pp. 2239-2249.
- [11] Pantula, P. D., Miriyala, S. S., Mitra, K., 2020. An Evolutionary Neuro-Fuzzy C-Means Clustering Technique. *Engineering Applications of Artificial Intelligence*, Vol. 89, pp. 103435-103450.
- [12] Prasetyo, E., 2014. *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: ANDI.
- [13] Santoso, S., 2015. *Menguasai Statistik Multivariat*. Jakarta: PT Elex Media Komputindo.
- [14] Sardar, T. H, Ansari, Z., 2018. An Analysis of Map Reduce Efficiency in Document Clustering Using Parallel K-Means Algorithm. *Future Computing and Infoematics Journal*, Vol. 3, Issue 2, pp. 200-209.
- [15] Wang, W., Zhang, Y., 2007. On Fuzzy Cluster Validity Indices. *Fuzzy Sets System*, Vol.158, No.19, pp. 2095-2117.
- [16] Zhu, X., Pedrycz, W., Li, Z., 2017. Fuzzy Clustering with Nonlinearly Transformed Data. *Applied Soft Computing*, Vol. 61, pp. 364-376.