

Perbandingan Algoritma K-Means dan SFCM Pada Pengelompokan Rumah Tangga Miskin

Wiwi Widayani¹⁾, Harliana^{2*)}

¹⁾ Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta
Jl Ring Road Utara, Condongcatur Sleman, Yogyakarta

^{2*)} Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhadi Setiabudi
Jl Pangeran Diponegoro KM 2 Pesantunan Wanasari, Brebes

*Email corresponden author: harliana.hifzhiya@gmail.com

Abstrak

Secara definisi rumah tangga miskin dan penduduk miskin memiliki sudut pandang yang berbeda, dimana definisi rumah tangga miskin akan lebih ditekankan kepada individu yang akan dijadikan survey dalam menentukan penduduk miskin sedangkan penduduk miskin lebih kepada kumpulan dari beberapa rumah tangga miskin. Penelitian ini dilakukan untuk membandingkan antara algoritma *K-Means* dan *Fuzzy Subtractive Clustering* (SFCM) dalam mengelompokkan rumah tangga miskin. Kedua algoritma ini akan dibandingkan berdasarkan simpangan baku dan validitas hasil pengelompokan yang dihasilkan. Berdasarkan 6 pengujian yang telah dilakukan, maka didapatkan hasil bahwa dari sisi waktu algoritma *K-Means* mampu mengelompokkan lebih cepat bila dibandingkan dengan algoritma SFCM, namun dari sisi simpangan baku kelompok, simpangan baku antar kelompok, maupun akurasi maka algoritma SFCM memiliki performa yang lebih baik bila dibandingkan dengan algoritma *K-Means*.

Kata kunci: K-Means, SFCM, Pengelompokan

Abstract

Based on the definition, poor households and poor people have different perspective, where the definition of poor households will be emphasized to individuals will be used as a survey to determine the poor population while the poor population is a collection of several poor households. This research was conducted to compare between the K-Means algorithm and Subtractive Fuzzy C-Means (SFCM) in classifying poor households. These two algorithms will be compared based on the standard deviation and the result validity of the resulting group. Based on 6 tests that have been done, the results show that in terms of time the K-Means algorithm is able to group faster when compared to the SFCM algorithm, but from the side of standard group deviations, intergroup standard deviations and accuracy, so the SFCM algorithm has better performance compared to the K-Means algorithm.

Keywords: K-Means, SFCM, clustering

1. PENDAHULUAN

Rumah tangga miskin merupakan salah satu sektor penting yang digunakan pemerintah dalam proses perhitungan kemiskinan. Secara definisi rumah tangga miskin dan penduduk miskin memiliki definisi yang berbeda, dimana rumah tangga miskin adalah suatu individu / keluarga / rumah tangga yang dijadikan survey dalam menentukan penduduk miskin. Sedangkan penduduk miskin adalah kumpulan dari beberapa individu / keluarga / rumah tangga yang memiliki rata-rata pengeluaran perbulan dibawah garis kemiskinan[1]. Berbagai macam pendekatan yang dapat digunakan untuk mengukur kemiskinan diantaranya Data Dasar Rumah Tangga (DDRT) dan Survey Rumah Tangga (SRT) yang dilakukan oleh Badan Ketahanan Pangan Kabupaten Cirebon,

Survey Sosial Ekonomi Nasional (Susenas) yang dilakukan oleh BPS, serta Basis Data Terpadu yang dilakukan oleh Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K). Pada penelitian ini penulis menggunakan DDRT karena kriteria yang terdapat dalam DDRT mampu mewakili kondisi rumah tangga yang disurvei mulai dari pendidikan anggota rumah tangga, kondisi rumah, konsumsi sehari-hari, kemampuan daya beli, aset kepemilikan kendaraan, serta pengeluaran rumah tangga selama satu bulan terakhir [2].

Berdasarkan DDRT tersebut, maka penelitian ini bertujuan untuk membandingkan hasil pengelompokan antara algoritma *K-Means* dan SFCM dalam mengelompokkan rumah tangga miskin. Penelitian mengenai *K-Means* dalam mengelompokkan kemiskinan telah banyak dilakukan sebelumnya, diantaranya yaitu perbandingan FCM dan *K-Means* dalam mengelompokkan kemiskinan di Desa Girijati Yogyakarta, menurut hasil penelitian menunjukkan bahwa *K-Means* memiliki akurasi yang lebih besar yaitu 83,33% bila dibandingkan dengan FCM yang hanya 50% [3]. Sedangkan jika dilihat dari simpangan baku terhadap hasil pengelompokan, maka algoritma *K-Means* memiliki nilai simpangan baku S_b dan S_w terkecil bila dibandingkan dengan algoritma FCM[4]. Namun apabila dilihat dari sisi lamanya waktu perhitungan, maka algoritma SFCM memiliki waktu pemrosesan lebih lama tetapi memiliki sebaran lebih baik bila dibandingkan dengan algoritma FCM [5].

Berdasarkan kelebihan dan kekurangan antara algoritma *K-Means* dan SFCM tersebut, maka penelitian ini akan membandingkan keduanya berdasarkan simpangan baku kelompok (S_w), simpangan baku antar kelompok (S_b), maupun validitas hasil pengelompokan (*accuracy*, *error rate*, *precision*, *recall*, dan *sensitivity*). Sehingga didapatkan algoritma terbaik dalam mengelompokkan rumah tangga miskin

2. TINJAUAN PUSTAKA

2.1 Rumah Tangga Miskin

Secara definisi kemiskinan memiliki berbagai macam pengertian, konsep dan sudut pandang berbeda, diantaranya yaitu: kemiskinan alami, kemiskinan struktural, kemiskinan kultural, kemiskinan absolute ataupun kemiskinan relative[6]. Apabila dilihat dari karakteristik responden pada DDRT yang digunakan, maka perhitungan rumah tangga miskin pada tingkat desa yang dilakukan lebih cenderung pada konsep kemiskinan relative yang membandingkan tingkat kesejahteraan penduduk dengan lingkungannya[7]. DDRT akan melakukan pengelompokan variable menjadi 5 indikator yaitu[8]:

- a. Keterangan umum rumah tangga yang terdiri atas:
 - 1) Pendidikan kepala keluarga (B4K3)
 - 2) Banyaknya anggota keluarga (B4K4)
 - 3) Banyaknya anggota rumah tangga balita (B4K5)
 - 4) Banyaknya anak usia sekolah 7-15 tahun (B4K6)
 - 5) Banyaknya anak usia 7-15 tahun yang masih sekolah (B4K7)
- b. Keterangan kondisi rumah tempat tinggal yang terdiri atas:
 - 1) Luas lantai (m^2) (B4K8)
 - 2) Jenis lantai (B4K9)
 - 3) Sumber air minum (B4K10)
 - 4) Sumber penerangan (B4K11)
- c. Keterangan rumah tangga mengkonsumsi daging/ayam/ikan/telur selama satu minggu yang lalu (B4K12)
- d. Keterangan tentang ketersediaan bahan makanan pokok (B4K13)
- e. Keterangan lapangan usaha dari pekerjaan utama rumah tangga (B4K14)
- f. Keterangan tentang kemampuan daya beli rumah tangga yang diukur melalui kemampuan membeli pakaian dalam satu tahun terakhir (B4K15)
- g. Keterangan tentang kepemilikan asset yang meliputi:
 - 1) Luas lahan sawah / lading / kebun (B4K16)
 - 2) Kepemilikan kendaraan bermotor (B4K17)

- 3) Kepemilikan sepeda / sampan / kendaraan bermotor lainnya (B4K18)
- 4) Kepemilikan tempat tidur dengan busa / kasur (B4K19)
- 5) Kepemilikan hewan ternak besar, seperti sapi, kerbau, kuda dan kambing (B4K20)
- h. Keterangan tentang pengeluaran rumah tangga yang meliputi:
 - 1) Pengeluaran untuk makan sebulan (B4K21)
 - 2) Total pengeluaran sebulan (B4K22)

2.2 Algoritma K-Means

K-Means merupakan salah satu algoritma *clustering* yang cocok digunakan ketika atribut yang dipakai bersifat numerik, prinsip kerja algoritma ini yaitu meminimalisasi jarak antar data ke pusat *cluster*[9]. Adapun langkah-langkah algoritma K-Means adalah sebagai berikut[10]:

- a. Inisialisasi:
 - 1) Jumlah k sebagai jumlah *cluster* yang ingin dibentuk
 - 2) *Centroid* (pusat *cluster*) awal. Biasanya *centroid* awal yang terbentuk adalah rerata dari data yang digunakan
- b. Hitung jarak antara data ke *centroid* melalui persamaan 1
$$d(P, Q) = \sqrt{\sum_{j=1}^p (x_j(P) - X_j(Q))^2} \quad (1)$$
- c. Kelompokkan data berdasarkan jarak terdekat data ke *centroid*
- d. Tentukan *centroid* baru
- e. Ulangi langkah b dan seterusnya jika *centroid* baru dan *centroid* lama tidak berubah.

2.3 Algoritma SFCM

Subtractive Fuzzy C Means (SFCM) atau *Fuzzy Subtractive Clustering* merupakan salah satu algoritma perluasan dari FCM, dimana jumlah *cluster* yang akan dibentuk tidak dapat ditentukan sebelumnya, selain itu perbedaan mendasar antara algoritma SFCM dengan FCM adalah pada derajat keanggotaan awal yang terbentuk belum tentu berjumlah satu[11].

Adapun langkah-langkah dari algoritma SFCM adalah sebagai berikut:

- a. Input data yang akan di *cluster* : X_{ij} dengan $i = 1, 2, \dots, n$ dan $j = 1, 2, \dots, m$
- b. Tetapkan nilai dari:
 - 1) r_j (jari-jari setiap atribut data), $j = 1, 2, \dots, m$
 - 2) q (*squash factor*)
 - 3) *Accept ratio*
 - 4) *Reject ratio*
 - 5) X_{min} (minimum data diperbolehkan)
 - 6) X_{max} (maksimum data diperbolehkan)
- c. Normalisasi dengan persamaan:

$$X_{ij} = \frac{X_{ij} - X_{min_j}}{X_{max_j} - X_{min_j}}, i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (2)$$

- d. Tentukan potensi awal tiap titik data:
 - 1) $i = 1$
 - 2) kerjakan hingga $i = n$, jika
 - a) $T_j = X_{ij}; j = 1, 2, \dots, m$ (3)
 - b) Hitung : $Dist_{kj} = \frac{T_j - X_{kj}}{r}$, $j = 1, 2, \dots, m; k = 1, 2, \dots, n$ (4)
 - c) Potensi awal:
 - Jika $m = 1$, maka $D_i = \sum_{k=1}^n e^{-4(Dist^2_{kj})}$ (5)
 - Jika $m > 1$, maka $D_i = \sum_{k=1}^n e^{-4(\sum_{j=1}^m Dist^2_{kj})}$ (6)
 - d) $i = i + 1$
- e. Cari titik dengan potensi tertinggi, dengan:

- 1) $M = \max[D_i | i = 1, 2, \dots, n]$
- 2) Pilih $h = i$, sedemikian sehingga $D_i = M$
- f. Tentukan pusat *cluster* dan kurangi potensinya terhadap titik-titik disekitarnya
 - 1) $Center = []$
 - 2) $V_j = X_{ij}; j = 1, 2, \dots, m$
 - 3) $C = 0$ (jumlah *cluster*)
 - 4) Kondisi = 1
 - 5) $Z = M$
 - 6) Kerjakan jika (kondisi $\neq 0$) & ($Z \neq 0$)
 - a) Kondisi = 0 (sudah tidak ada calon pusat baru lagi)
 - b) $Ratio = Z/M$
 - c) Jika $ratio > accept\ ratio$, maka kondisi = 1 (ada calon pusat baru)
 - d) Jika tidak,
 - i. $Ratio > reject\ ratio$ (calon baru akan diterima sebagai pusat jika keberadaannya akan memberikan keseimbangan terhadap data-data yang letaknya cukup jauh dengan pusat *cluster* yang telah ada), maka lanjutkan
 - ii. $Md = -1$
 - iii. Kerjakan untuk $i=1$ sampai $i=C$

$$G_{ij} = \frac{V_j - Center_{ij}}{r}; j = 1, 2, \dots, m \quad (7)$$

$$Sd_i = \sum_{j=1}^m (G_{ij})^2 \quad (8)$$

Jika ($Md < 0$) atau ($Sd < Md$), maka $Md = Sd$
 - iv. $Smd = \sqrt{Md}$
 - v. Jika $(ratio + Smd) \geq 1$, maka kondisi = 1 (data diterima sebagai pusat *cluster*)
 - vi. Jika $(ratio + Smd) < 1$, maka kondisi = 2 (data tidak dipertimbangkan sebagai pusat *cluster*)
 - e) Jika kondisi = 1 (calon pusat baru diterima sebagai pusat baru), lanjutkan:
 - i. $C = C + 1;$
 - ii. $Center_c = V$
 - iii. Kurangi potensi dari titik-titik di dekat pusat t:

$$S_{ij} = \frac{V_j - X_{ij}}{r_j * q}; i = 1, 2, \dots, m \quad (9)$$

$$DC_i = M e^{-4 [\sum_{j=1}^4 (S_{ij})^2]}; i=1, 2, \dots, n \quad (10)$$

$$D = D - D_c$$

Jika $D_i \leq 0$, maka $D_i = 0; i=1, 2, \dots, n$

$$Z = \max[D_i | i = 1, 2, \dots, n];$$

Pilih $h = i$, sedemikian hingga $D_i = Z$
 - f) Jika kondisi = 2 (calon pusat baru tidak diterima sebagai pusat baru), maka:
 - i. $D_h = 0;$
 - ii. $Z = \max[D_i | i = 1, 2, \dots, n]$
 - iii. Pilih $h = i$, sedemikian sehingga $D_i = Z$
- g. Kembalikan pusat *cluster* dari bentuk ternormalisasi ke bentuk semula

$$center_{ij} = Center_{ij} * (XMax_j - XMin_j) + XMin_j \quad (11)$$
- h. Hitung nilai sigma *cluster*

$$\sigma = r_j * \frac{XMax_j - XMin_j}{\sqrt{8}} \quad (12)$$

3. METODE PENELITIAN

Tahapan penelitian yang dilakukan dalam penelitian ini, yaitu:

- a) Pengumpulan data rumah tangga miskin
Pada tahapan pengumpulan data, peneliti akan menganalisis data hasil DDRT yang telah dimiliki oleh Dinas Ketahanan Pangan Kabupaten Cirebon. Penelitian ini menggunakan 753 responden dengan 5 indikator pada 19 kriteria.
- b) *Pre-processing* data rumah tangga miskin
Pada tahapan *pre-processing*, data yang sudah didapatkan selanjutnya akan dianalisa kembali apakah sudah sesuai dengan kebutuhan baik dari sisi kelengkapan data, penyesuaian pengisian data, serta tidak mengandung data yang *redundancy*.
- c) Implementasi data kedalam algoritma K-Means dan SFCM
Pada tahapan implementasi, data pada poin (c) tersebut selanjutnya akan diterapkan pada algoritma K-Means dan SFCM untuk mencari pengelompokan yang dihasilkan.
- d) Analisa metode terbaik
Pada tahapan pencarian metode terbaik antara K-Means dan SFCM, peneliti akan menggunakan nilai simpangan baku yang dihasilkan melalui nilai simpangan baku kelompok (S_w) dan simpangan baku antar kelompok (S_b). Jika suatu algoritma memiliki nilai S_w lebih kecil dibandingkan dengan nilai S_b maka algoritma tersebut memiliki persamaan yang tinggi dan menjadi algoritma terbaik dalam *clustering*, persamaan untuk mencari nilai S_w dan S_b tersebut terdapat pada persamaan (13) dan (14)[12]

$$S_{wj} = \frac{1}{c} \sum_{k=1}^c S_{kj} \quad (13)$$

$$S_{bj} = \left[\frac{1}{c-1} \sum_{k=1}^c (X_{kj} - X)^2 \right]^{1/2} \quad (14)$$

- e) Pengujian
Pada tahap pengujian, peneliti akan melakukan uji validasi kedua algoritma K-Means dan SFCM melalui *silhouette coefficient* dan *confusion matrix*. *Silhouette coefficient* digunakan karena mampu menguji kekonsistensian titik objek suatu *cluster* berdasarkan kemiripan dirinya sendiri dengan ketidakmiripan dirinya dengan *cluster* lain, persamaan *silhouette coefficient* terdapat pada persamaan (15)[13].

$$S(i) = \frac{x(i)-y(i)}{\max\{y(i),x(i)\}} \quad (15)$$

Sedangkan *confusion matrix* digunakan karena mampu mengukur kinerja hasil yang diperoleh sistem dengan hasil yang sesungguhnya, persamaan *confusion matrix* terdapat pada persamaan (16), (17) dan (18)[14]

$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (16)$$

$$presisi = \frac{TP}{FP+TP} \times 100\% \quad (17)$$

$$recall = \frac{TP}{FN+TP} \times 100\% \quad (18)$$

- f) Penarikan kesimpulan
Pada tahapan penarikan kesimpulan, peneliti akan menganalisa berdasarkan poin (d) dan (e) yang telah dihasilkan sebelumnya.

Adapun tahapan metode penelitian yang dilakukan, terangkum pada Gambar 1.



Gambar 1. Alur penelitian yang dilakukan

4. PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data yang berasal dari DDRT rumah tangga miskin tingkat desa, peta administrasi pangan serta peta rawan pangan Kabupaten Cirebon Tahun 2007. Berdasarkan 3 data primer tersebut, maka peneliti mendapatkan data awal berjumlah 764 data dengan 5 indikator dan 19 kriteria, namun setelah dilakukan *cleaning* data pada tahap *pre-processing* dengan menghilangkan data-data yang tidak lengkap dan *redundancy*, maka didapatkan 753 data. Untuk mendapatkan hasil yang akurat, peneliti menggunakan 6 skenario pengujian. Rangkuman mengenai skenario tersebut terangkum pada Tabel 1.

Tabel 1. Skenario pengujian

| Scenario ke- | Max Iterasi | Nilai Threshold | Accept ratio (ar) | Reject ratio (rr) | Jari-jari (r) |
|--------------|-------------|-----------------|-------------------|-------------------|---------------|
| 1 | 100 | 10^{-3} | 0,5 | 0,15 | 0,09 |
| 2 | 75 | 10^{-3} | 0,5 | 0,15 | 0,05 |
| 3 | 100 | 10^{-4} | 0,5 | 0,15 | 0,09 |
| 4 | 75 | 10^{-4} | 0,5 | 0,15 | 0,05 |
| 5 | 100 | 10^{-5} | 0,5 | 0,15 | 0,09 |
| 6 | 75 | 10^{-5} | 0,5 | 0,15 | 0,05 |

Skenario pada Tabel 1, dilakukan untuk algoritma K-Means dan SFCM. Namun nilai pada *accept ratio* (ar), *reject ratio* (rr) dan jari-jari (r) hanya digunakan untuk pengujian pada algoritma SFCM. Untuk mendapatkan jumlah *cluster* terbaik, peneliti melakukan perhitungan secara SFCM terlebih dahulu, kemudian jumlah *cluster* tersebut digunakan sebagai inputan manual untuk menghitung jumlah kelompok pada algoritma K-Means. Rangkuman hasil pengujian untuk algoritma K-Means terangkum pada Tabel 2.

Tabel 2. Hasil pengujian pada algoritma K-Means

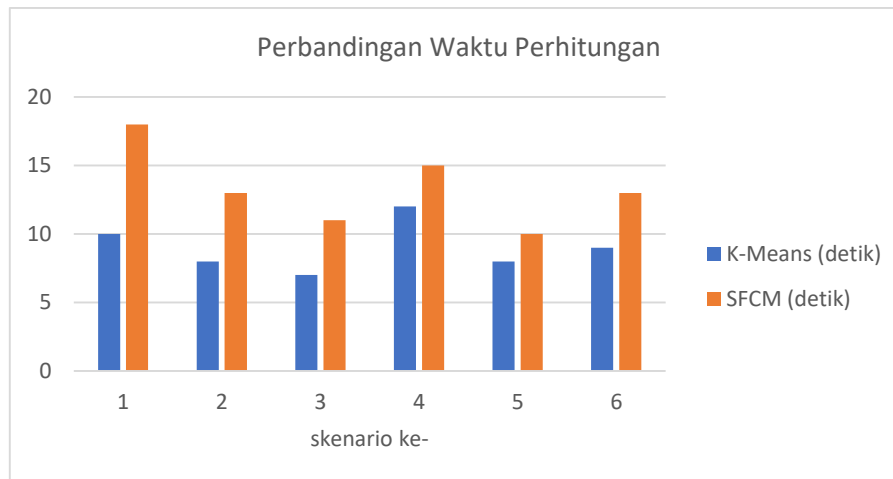
| <i>skenario ke-</i> | <i>Jumlah cluster</i> | S_w | S_b | S_i | <i>Presisi</i> | <i>Recall</i> | <i>Akurasi</i> |
|---------------------|-----------------------|---------------|---------------|---------------|----------------|---------------|----------------|
| skenario ke-1 | 3 | 0,3100 | 0,4500 | 0,7089 | 52% | 81% | 84% |
| skenario ke-2 | 4 | 0,2671 | 0,8912 | 0,3197 | 53% | 80% | 87% |
| skenario ke-3 | 5 | 0,3310 | 0,7045 | 0,7898 | 88% | 75% | 88% |
| skenario ke-4 | 4 | 0,9872 | 0,8312 | 0,5938 | 89% | 72% | 87% |
| skenario ke-5 | 5 | 0,8495 | 0,2438 | 0,7030 | 50% | 76% | 81% |
| skenario ke-6 | 2 | 0,9643 | 0,6331 | 0,4616 | 43% | 71% | 71% |

Dari Tabel 2, terlihat bahwa nilai akurasi tertinggi terdapat pada pengujian skenario ke-3 dengan jumlah *cluster* terbaik adalah 5, prosentase presisi skenario ini adalah 88% yang berarti bahwa algoritma K-Means mampu mengelompokkan jumlah rumah tangga miskin yang benar dari keseluruhan rumah tangga yang akan diprediksi masuk ke kelompok miskin. Sedangkan *recall* (*sensitivity*) algoritma K-Means dalam mengelompokkan rumah tangga miskin adalah 75% yang berarti bahwa algoritma ini mampu membandingkan dengan akurat prosentase kelompok rumah tangga yang diprediksi miskin dengan kelompok rumah tangga yang sebenarnya miskin. Namun apabila dilihat dari simpangan baku *cluster* dan simpangan baku antar *cluster*, maka pengujian pada skenario ke-1 memiliki pengelompokkan yang lebih baik dengan jumlah *cluster* 3. Sedangkan hasil pengujian untuk algoritma SFCM berdasarkan Tabel 1 terangkum pada Tabel 3.

Tabel 3. Hasil pengujian algoritma SFCM

| <i>No</i> | <i>jumlah cluster</i> | <i>Presisi</i> | <i>Recall</i> | <i>Akurasi</i> | S_i | S_w | S_b |
|---------------|-----------------------|----------------|---------------|----------------|---------------|-------------|--------------|
| skenario ke-1 | 3 | 92% | 51,0% | 89,7% | 0,3871 | 0,2567 | 0,291 |
| skenario ke-2 | 4 | 88% | 60,2% | 85,8% | 0,2564 | 0,2713 | 0,314 |
| skenario ke-3 | 5 | 97% | 45,0% | 94,1% | 0,4981 | 0,3981 | 0,4511 |
| skenario ke-4 | 4 | 96% | 39,8% | 94,8% | 0,5431 | 0,15 | 0,256 |
| skenario ke-5 | 5 | 91% | 62,0% | 89,2% | 0,2719 | 0,3391 | 0,6438 |
| skenario ke-6 | 2 | 82% | 40,3% | 80,1% | 0,3981 | 0,2671 | 0,8912 |

Berdasarkan Tabel 3, maka skenario 3 dan 4 memiliki nilai akurasi yang tinggi sekitar 94% bila dibandingkan skenario lainnya. Pada algoritma SFCM jumlah *cluster* yang terbentuk tidak akan dimasukkan pada tahap awal, tetapi dihitung berdasarkan kerapatan data terhadap objek sekitar sehingga didapatkan jumlah *cluster* optimal yang terbentuk. Apabila kita lihat, nilai *silhouette coefficient* yang dihasilkan pada skenario ke-4 ternyata lebih mendekati 1 bila dibandingkan dengan skenario lainnya, sehingga dapat dikatakan pengelompokkan pada skenario 4 lebih stabil bila dibandingkan yang lainnya. Selain itu banyak/sedikitnya jumlah *cluster* yang terbentuk, ternyata tidak berpengaruh terhadap nilai akurasi, rerata S_w/S_b ataupun *silhouette coefficient*. Sedangkan apabila dilihat dari sisi lamanya waktu pemrosesan data untuk mengelompok, maka algoritma K-Means memiliki waktu yang lebih cepat bila dibandingkan dengan algoritma SFCM. Rangkuman mengenai perbandingan waktu pemrosesan pengelompokkan data tersebut terangkum pada grafik Gambar 2. Berdasarkan Gambar 2, maka rata-rata waktu yang dibutuhkan algoritma K-Means dalam mengelompok adalah 9s sedangkan untuk algoritma SFCM adalah 13,3s. Kedua algoritma ini memiliki waktu yang relative lebih cepat karena data yang digunakan dalam berkelompok umumnya kecil yang hanya berkisar antara nilai 0-100.



Gambar 2. Perbandingan waktu yang dihasilkan

Namun apabila hasil pengujian pada Tabel 2 dan Tabel 3 kita bandingkan, maka terlihat bahwa algoritma SFCM mampu mengelompokkan data rumah tangga miskin lebih baik bila dibandingkan dengan algoritma K-Means. Hal ini didasarkan atas nilai akurasi tertinggi, nilai *silhouette coefficient*, serta rerata rasio S_w/S_b terkecil semuanya tepat berada pada skenario ke-4 algoritma SFCM. Akurasi yang tinggi dibutuhkan untuk memprosentasikan kebenaran yang dihasilkan oleh algoritma yang digunakan. Dimana semakin tinggi nilai akurasi yang dihasilkan maka semakin tinggi pula keakuratan algoritma tersebut bekerja, dan hal ini terdapat pada skenario ke-4 yaitu sekitar 94,8%. Selain itu nilai *silhouette coefficient* yang dihasilkan pada skenario 4 juga memiliki nilai terbesar (0,5431) dan hampir mendekati 1 yang berarti bahwa melalui pengujian skenario ke-4 data dapat berkelompok dengan baik melalui rerata nilai simpangan baku kelompok dan simpangan baku antar kelompok terkecil yaitu 0,203

5. KESIMPULAN

Berdasarkan 6 skenario pengujian yang telah dilakukan pada 753 data rumah tangga miskin, maka didapatkan bahwa algoritma SFCM mampu melakukan pengelompokkan rumah tangga miskin lebih baik bila dibandingkan dengan algoritma K-Means. Hal ini terlihat dari persentase nilai akurasi yang dihasilkan oleh SFCM lebih besar bila dibandingkan dengan K-Means yaitu 94,8% dengan cluster yang terbentuk maksimal berjumlah 4. Namun apabila dari sisi waktu pemrosesan pengelompokkan, maka algoritma K-Means memiliki waktu yang lebih cepat bila dibandingkan dengan SFCM, yaitu sekitar 9 detik.

DAFTAR PUSTAKA

- [1] BPS, "Kabupaten Brebes Dalam Angka," in *Kabupaten Brebes Dalam Angka*, 2019th ed., Brebes: BPS Kabupaten Brebes, 2019.
- [2] Badan Ketahanan Pangan, "Badan ketahanan pangan kabupaten cirebon DENGAN VECTORIAL PROJECT ANALYSIS (V P A) Vectorial Project Analysis (VPA)," in *Monotoring dan Evaluasi Kemiskinan*, III., vol. 5, Badan Ketahanan Pangan, Ed. Kabupaten Cirebon: Badan Ketahanan Pangan Kabupaten Cirebon, 2016.
- [3] A. Ulfah and S. 'uyun, "Analisis Kinerja Algoritma Fuzzy C-Means Dan K-Means Pada Data Kemiskinan," *Jatiti*, vol. Vol. 1, no. No. 2, pp. 139–148, 2015.
- [4] S. S. P. Luka, I. M. Candiasa, and K. Y. E. Aryanto, "ANALISIS PEMBENTUKAN KELOMPOK DISKUSI PANEL SISWA MENGGUNAKAN ALGORITMA FUZZY C-MEANS DAN K-MEANS," *J. Pendidik. Teknol. dan Kejuru.*, vol. 16, no. 2, pp. 267–277, 2019.

-
- [5] B. N. Haqiqi and R. Kurniawan, "Analisis perbandingan metode," *Media Stat.*, vol. 2, no. 2, pp. 1–15, 2015.
 - [6] B. Dimas Sanjaya, S. A Kadir, and F. Bahri, "Analisis Kemiskinan di Kota Pagar Alam," *J. Ekon. Pembang.*, vol. 16, no. 2, pp. 81–93, 2019.
 - [7] P. Anita Rahman, Firman, and Rusdinal, "Kemiskinan Dalam Perspektif Ilmu Sosiologi," *J. Pendidik. Tambusai*, vol. 3, no. 6, pp. 75–84, 2019.
 - [8] BKP, "Tahapan Tingkat Kesejahteraan Rumahtangga Miskin Pangan," Kabupaten Cirebon, 2016.
 - [9] R. W. Sari, A. Wanto, and A. P. Windarto, "Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 224–230, 2018.
 - [10] A. Asroni, H. Fitri, and E. Prasetyo, "Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokkan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik)," *Semesta Tek.*, vol. 21, no. 1, pp. 60–64, 2018.
 - [11] Rahmawati, Wartono, and A. Y. Putri, "Pengklasteran Tingkat Pendidikan Pegawai Menggunakan Metode Fuzzy Subtractive Clustering (Studi Kasus : Badan Kepegawaian dan Pelatihan Daerah Provinsi Riau)," vol. 5, no. 1, pp. 79–89, 2019.
 - [12] M. Goreti, Y. N. N, and S. Wahyuningsih, "Perbandingan Hasil Analisis Cluster dengan Menggunakan Metode Single Linkage dan Metode C-Means," *J. EKSPONENSIAL*, vol. 7, no. 1, pp. 9–16, 2016.
 - [13] B. K. Amijaya, M. T. Furqon, and C. Dewi, "Clustering Titik Panas Bumi Menggunakan Algoritme Affinity Propagation," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 10, pp. 3835–3842, 2018.
 - [14] R. F. Amanullah, A. Pujiyanto, B. T. Pratama, and Kusriani, "Deteksi Motif Batik Menggunakan Ekstraksi Tekstur dan Jaringan Syaraf Tiruan," *CSRID (Computer Sci. Res. Its Dev. J.)*, vol. 10, no. 2, pp. 31–36, 2018.

Biodata Penulis

Wiji Widayani, M.Kom, lahir di Majalengka pada tahun 1983. Penulis pertama memperoleh gelar S.Kom di Jurusan Sistem Informasi STMIK Amikom Yogyakarta pada Tahun 2006. Kemudian melanjutkan pendidikan S2 pada STMIK Amikom Yogyakarta (sekarang telah menjadi Universitas) dan lulus pada Tahun 2015. Konsentrasi penelitian yang penulis pertama dalam yaitu bidang Artificial Intellience dan Rekayasa Perangkat Lunak. Saat ini penulis adalah salah satu dosen di Jurusan Sistem Informasi pada Universitas Amikom Yogyakarta.

Harliana, ST., M.Cs, lahir di Cirebon pada tahun 1986. Penulis kedua memperoleh gelar S.T di Jurusan Teknik Informatika Universitas Muhammadiyah Cirebon pada Tahun 2008. Kemudian melanjutkan pendidikan S2 pada Universitas Gadjah Mada Yogyakarta dan lulus pada Tahun 2012. Konsentrasi penelitian yang penulis pertama dalam yaitu bidang *Artificial Intellience*, data mining, dan *image processing*. Saat ini penulis adalah salah satu dosen di Jurusan Teknik Informatika pada Universitas Muhadi Setiabudi Brebes.