

PENGENALAN POLA TIDAK TERBIMBING DENGAN METODE FUZZY SUBSTRUCTIVE CLUSTERING

Putu Manik Prihatini

Jurusan Teknik Elektro, Politeknik Negeri Bali
Bukit Jimbaran, P.O.Box 1064 Tuban Badung – BALI
Phone:+62-361-701981, Fax:+62-361-701128 E-mail: manikprihatini@gmail.com

Abstrak: Klustering termasuk dalam pengenalan pola tidak terbimbing, vektor fitur pelatihan dan kelas-kelas tujuannya belum diketahui. *Fuzzy subtractive clustering* (FSC) adalah teknik klustering yang didasarkan atas ukuran densitas titik-titik data dalam suatu ruang. Metode FSC diterapkan pada penelitian ini untuk melakukan pengelompokan industri kecil dengan dua kriteria yaitu modal awal dan rata-rata penjualan. Proses diawali dengan melakukan normalisasi, kemudian menghitung potensi awal, mencari titik dengan potensi awal tertinggi, menentukan pusat kluster, mengembalikan pusat kluster ke bentuk semula, serta menghitung sigma kluster dan derajat keanggotaan. Berdasarkan parameter jari-jari sebesar 0,3; *squash factor* sebesar 0,25; *accept_ratio* sebesar 0,5; *reject_ratio* sebesar 0,15; *XMin* sebesar [0;0]; dan *XMax* sebesar [30.000.000; 40.000.000], proses klustering dihentikan pada iterasi ke-3. Hasil akhir klustering mengelompokkan industri ke dalam 3 kelompok yaitu Kluster 1 (Industri ke-1, 3, 4 dan 6), Kluster 2 (Industri ke-2, 8, 9 dan 10), dan Kluster 3 (Industri ke-5 dan 7).

Kata Kunci: pola, klustering, FSC

UNSUPERVISED PATTERN RECOGNITION BY FUZZY SUBSTRUCTIVE CLUSTERING

Abstract: Clustering is part of the unsupervised pattern recognition, training features vector and destination classes have not been recognized. *Fuzzy subtractive clustering* (FSC) is a clustering technique based on the size of data point's density within a space. FSC method was implemented in the study to perform clustering of small scale industry with two criteria, such as initial capital and sale average. The process was initiated by normalizing, then calculated the initial potential, looking for a point with the highest initial potential, determine the center of the cluster, restoring the cluster center to its original shape, and calculates the sigma cluster and the degree of membership. Based on radius of 0.3; *squash factor* of 0.25; *accept_ratio* of 0.5; *reject_ratio* of 0.15; *Xmin* of [0; 0]; and *Xmax* of [30,000,000; 40,000,000], clustering process is stopped in the third iteration. The final result of clustering could group industrial data into three clustering groups, that are Cluster 1 (Industry 1, 3, 4 and 6), Cluster 2 (Industry 2, 8, 9 and 10), and Cluster 3 (Industry 5 and 7).

Key words: pattern, clustering, FSC

I. PENDAHULUAN

Klustering merupakan bidang ilmu dalam pengenalan pola, yang termasuk dalam pengenalan pola tidak terbimbing, vektor fitur pelatihan dan kelas-kelas tujuannya belum diketahui. Klustering mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan. Pada klustering, tipe keluaran yang ingin dicapai tidak perlu ditentukan diawal, cukup menentukan jumlah kelompok yang ingin dicapai, maka metode klustering menentukan kelompok-kelompok yang terbentuk. Ada banyak metode klustering yang dapat digunakan dalam pengenalan pola. Salah satunya adalah klustering dengan metode *Fuzzy C-Means* (FCM) yang telah dilakukan dalam penelitian terdahulu (Manik, 2009).

Selain metode FCM, metode klustering lainnya adalah *Fuzzy Subtractive Clustering* (FSC). Perbedaan antara FSC dengan FCM adalah jika pada metode FCM pusat kluster bisa jadi bukan merupakan salah satu dari data yang dikluster, tetapi pada FSC, suatu pusat kluster pasti merupakan salah satu dari data yang dikluster, yaitu data dimana derajat keanggotaannya pada kluster tersebut sama dengan 1. Penjumlahan derajat keanggotaan pada FCM selalu bernilai sama dengan 1, tetapi pada FSC, penjumlahan semua derajat keanggotaannya belum tentu (bahkan jarang) bernilai sama dengan 1.

Melalui penelitian ini, penulis ingin membahas mengenai penerapan teknik klustering pada contoh kasus pengelompokan industri kecil dengan *Fuzzy Subtractive Clustering*, serta membandingkan hasilnya dengan metode FCM dari penelitian terdahulu.

II. FUZZY SUBTRACTIVE CLUSTERING

Subtractive clustering didasarkan atas ukuran densitas (potensi) titik-titik data dalam suatu ruang (variabel). Konsep dasar dari subtractive clustering adalah menentukan daerah-daerah dalam suatu variabel yang memiliki densitas tinggi terhadap titik-titik disekitarnya. Titik dengan jumlah tetangga terbanyak dipilih sebagai pusat kluster. Titik yang sudah terpilih sebagai pusat kluster kemudian dikurangi densitasnya. Algoritma memilih titik lain yang memiliki tetangga terbanyak untuk dijadikan pusat kluster yang lain. Hal ini dilakukan berulang-ulang hingga semua titik diuji.

Pada FSC digunakan dua faktor pembandingan yaitu *accept ratio* dan *reject ratio*, merupakan bilangan pecahan yang bernilai 0 sampai 1. *Accept ratio* merupakan batas bawah dimana suatu titik data yang menjadi kandidat pusat kluster diperbolehkan untuk menjadi pusat kluster. *Reject ratio* merupakan batas atas dimana suatu titik data yang menjadi kandidat pusat kluster diperbolehkan untuk menjadi pusat kluster.

Algoritma FSC dapat diuraikan sebagai berikut.

1. Input data yang akan dikluster, X , berupa matriks berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data), dimana X_{ij} = data sampel ke- i ($i = 1, 2, \dots, n$), atribut ke- j ($j = 1, 2, \dots, m$)
2. Tetapkan nilai:
 - a. R_j (jari-jari setiap atribut data); $j = 1, 2, \dots, m$
 - b. q (*squash factor*)
 - c. *accept_ratio*
 - d. *reject_ratio*
 - e. $XMin$ (minimum data yang diperbolehkan)
 - f. $XMax$ (maksimum data yang diperbolehkan)

3. Normalisasi

$$X_{ij} = \frac{X_{ij} - XMin_j}{XMax_j - XMin_j}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m$$

4. Tentukan potensi awal tiap-tiap titik data
 - a. $i = 1$
 - b. Kerjakan hingga $i = n$,
 - 1) $T_j = X_{ij} \quad j = 1, 2, \dots, m$
 - 2) Hitung:

$$Dist_{kj} = \left(\frac{T_j - X_{kj}}{r} \right) \quad j = 1, 2, \dots, m; \quad k = 1, 2, \dots, n$$

- 3) Potensi awal:

Jika $m = 1$, maka:

$$D_i = \sum_{k=1}^n e^{-4(D_{ik})^2}$$

Jika $m > 1$, maka:

$$Us_k = \sum_{j=1}^m Dist_{kj}^2$$

$$D_i = \sum_{k=1}^n e^{-4(Us_k)}$$

- 4) $i = i + 1$
5. Cari titik dengan potensi tertinggi
 - a. $M = \max[D_i \mid i = 1, 2, \dots, n]$
 - b. $h = i$, sedemikian hingga $D_i = M$
6. Tentukan pusat kluster dan kurangi potensinya terhadap titik-titik disekitarnya
 - a. Center = []
 - b. $C = 0$ (jumlah kluster)
 - c. Kondisi = 1
 - d. $Z = M$
 - e. $V_j = X_{hj}; j = 1, 2, \dots, m$
 - f. Kerjakan jika (Kondisi $\neq 0$) dan ($Z \neq 0$)
 - 1) Rasio = Z/M
 - 2) Jika Rasio $\leq reject_ratio$, maka Kondisi = 0 (calon pusat baru tidak diterima sebagai pusat kluster)
 - 3) Jika Rasio $> accept_ratio$, maka Kondisi = 1 (calon pusat baru diterima sebagai pusat kluster)
 - 4) Jika $reject_ratio < Rasio \leq accept_ratio$ (calon baru akan diterima sebagai pusat jika keberadaannya akan memberikan keseimbangan terhadap data-data yang letaknya cukup jauh dengan pusat kluster yang telah ada), maka kerjakan:
 - a) $Md = -1$
 - b) Kerjakan untuk $i = 1$ sampai $i = C$:

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

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

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

- c) $Smd = \sqrt{Md}$
- d) Jika ($Rasio + Smd \geq 1$), maka Kondisi = 1 (data akan diterima sebagai pusat kluster)
- e) Jika ($Rasio + Smd < 1$), maka Kondisi = 2 (data tidak akan dipertimbangkan kembali sebagai pusat kluster)
- 5) Jika Kondisi = 1, kerjakan:
 - a) $C = C + 1$
 - b) Center = V
 - c) Kurangi potensi dari titik-titik di dekat pusat kluster

$$S_{ij} = \frac{V_j \cdot X_{ij}}{r_j \cdot q} \quad j = 1, 2, \dots, m; \quad i = 1, 2, \dots, n$$

$$ST_i = \sum_{j=1}^2 S_{ij}^2$$

$$De_i = M \cdot e^{-4 \cdot ST_i}; \quad i = 1, 2, \dots, n$$

$$D_i = D_i - Dc_i$$

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 $D_i = Z$

6) Jika Kondisi = 2, kerjakan:

a) $D_h = 0$

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

c) Pilih $h = i$, hingga $D_i = Z$

7. Kembalikan pusat kluster dari bentuk ternormalisasi ke bentuk semula

$$Center_{ij} = Center_{ij} \cdot (XMax_j - XMin_j) + XMin_j$$

8. Hitung nilai sigma kluster

$$\sigma_j = r_j \cdot \frac{(XMax_j - XMin_j)}{\sqrt{8}}$$

Hasil dari algoritma FSC berupa matriks pusat kluster (C) dan sigma (σ) akan digunakan untuk menentukan nilai parameter fungsi keanggotaan Gauss. Derajat keanggotaan suatu titik data X_i pada kluster ke- k adalah:

$$\mu_{ki} = e^{-\frac{\sum_{j=1}^m (X_{ij} - C_{kj})^2}{2\sigma_j^2}}$$

III. DATA TEKNIS OBJEK

Penerapan teknik klustering berbasis prototipe metode FSC dilakukan pada contoh kasus yang digunakan pada penelitian terdahulu dengan metode FCM (Kusumadewi, 2009). Pemerintah daerah suatu kabupaten mendata sejumlah industri kecil di lingkungannya berdasarkan beberapa kriteria untuk mengelompokkan industri tersebut ke dalam beberapa kluster, sehingga lebih mudah merencanakan upaya pengembangan terhadap industri-industri tersebut.

Proses klustering yang dilakukan dengan metode FSC menggunakan data input dan parameter yaitu:

$n = 10$ (ada 10 industri)

$m = 2$ (ada dua kriteria yaitu modal awal dan rata-rata penjualan dari setiap industri)

Jari-jari (r) = 0,3

q (*Squash factor*) = 0,25

accept_ratio = 0,5

Modal	Rata-Rata Penjualan
15,000,000	25,000,000
20,000,000	26,420,000
17,820,000	22,052,000
16,205,000	18,500,000
8,000,000	15,200,000
14,260,000	19,640,000
7,025,000	15,230,000
25,032,000	34,000,000
24,320,100	35,100,000
25,602,100	38,200,000

X =

Gambar 1. Data Input Penelitian

reject_ratio= 0,15

XMin = [0;0]

XMax = [30.000.000; 40.000.000]

Keluaran yang dihasilkan dari proses klustering adalah industri-industri yang masuk dalam kelompok kluster (1 atau 2 atau 3) sesuai dengan parameter jumlah kluster.

IV. HASIL DAN PEMBAHASAN

Langkah 1

Lakukan normalisasi terhadap data ke-1 sampai dengan data ke-10. Contoh pada data pertama.

Atribut Pertama:

$$X_{11} = \frac{X_{11} - XMin_1}{XMax_1 - XMin_1} = \frac{15.000.000 - 0}{30.000.000 - 0} = \frac{15.000.000}{30.000.000} = 0,5000$$

Atribut Kedua:

$$X_{12} = \frac{X_{12} - XMin_2}{XMax_2 - XMin_2} = \frac{25.000.000 - 0}{40.000.000 - 0} = \frac{25.000.000}{40.000.000} = 0,6250$$

Langkah di atas dilakukan juga terhadap data ke-4 sampai data ke-10, sehingga hasil akhir normalisasi seperti pada Gambar 2.

Normalisasi

No	1	2
1	0.5000	0.6250
2	0.6667	0.6605
3	0.5940	0.5513
4	0.5402	0.4625
5	0.2667	0.3800
6	0.4753	0.4910
7	0.2342	0.3808
8	0.8344	0.8500
9	0.8107	0.8775
10	0.8534	0.9550

Gambar 2. Hasil Normalisasi

Langkah 2

Menghitung potensi awal untuk data ke-1 sampai data ke-10. Contoh pada data pertama terhadap data pertama.

Atribut Pertama:

$$Dist_{11} = \left(\frac{T_1 - X_{11}}{r} \right) = \frac{0,5000 - 0,5000}{0,3} = 0$$

Atribut Kedua:

$$Dist_{12} = \left(\frac{T_2 - X_{12}}{r} \right) = \frac{0,6250 - 0,6250}{0,3} = 0$$

Jadi, nilai total Dist untuk data pertama terhadap data pertama:

$$D_{s1} = \sum_{j=1}^2 Dist_{1j}^2 = Dist_{11}^2 + Dist_{12}^2 = 0^2 + 0^2 = 0$$

Perhitungan nilai total Dist dilakukan juga terhadap data ke-2 sampai data ke-10, sehingga hasil akhir potensi awal untuk data pertama seperti pada Gambar 3.

$$D_1 = \sum_{k=1}^{10} e^{-4(D_{s_k})} = e^{-4(0,0000)} + e^{-4(0,3226)} + e^{-4(0,1585)} + \dots = 2,5424$$

Dist _{kj}			
Dist _{k1}	Dist _{k2}	Ds = ∑ Dist _{kj} ²	e ^{-4(Ds)}
0.0000	0.0000	0.0000	1.0000
-0.5556	-0.1183	0.3226	0.2751
-0.3133	0.2457	0.1585	0.5304
-0.1339	0.5417	0.3113	0.2878
0.7778	0.8167	1.2719	0.0062
0.0822	0.4467	0.2063	0.4382
0.8861	0.8142	1.4481	0.0031
-1.1147	-0.7500	1.8050	0.0007
-1.0356	-0.8417	1.7808	0.0008
-1.1780	-1.1000	2.5977	0.0000
D1 =			2.5424

Gambar 3. Potensi Awal Data Pertama

Langkah 2 di atas dilakukan juga terhadap data ke-2 sampai data ke-10, sehingga hasil akhir potensi awal untuk seluruh data seperti pada Gambar 4.

Potensi Awal:	
D1	2.5424
D2	1.9930
D3	3.0759
D4	2.8315
D5	2.0727
D6	2.8756
D7	2.0136
D8	2.6062
D9	2.7001
D10	2.3135

Gambar 4. Hasil Akhir Potensi Awal

Langkah 3

- Mencari titik dengan potensi awal tertinggi
- $M = \max[D_i \mid i=1,2,\dots,n] = 3,0759$ yaitu potensi pada data ke-3;
 - $h = i = 3$, sedemikian hingga $D_3 = M = 3,0759$;

Langkah 4

Tentukan pusat kluster dan kurangi potensinya terhadap titik-titik disekitarnya

- Center = []
- $C = 0$
- Kondisi = 1
- $Z = M = 3,0759$

Karena Kondisi $\neq 0$ dan $Z \neq 0$, maka lakukan perhitungan untuk **#Iterasi 1**

- Menentukan pusat kluster:
Potensi tertinggi terletak pada data ke-3, maka hasil normalisasi data ke-3 menjadi pusat kluster, yaitu:
 $V_1 = 0,5940$
 $V_2 = 0,5513$
- Menentukan apakah calon pusat kluster diterima atau tidak sebagai pusat kluster
 Rasio = $Z / M = 3,0759 / 3,0759 = 1$
 Rasio > accept_ratio (0,5), maka Kondisi = 1.
- Kondisi = 1 berarti calon pusat kluster diterima sebagai pusat kluster, maka:
 - $C = C + 1 = 1$
 - Center₁ = 0,5940,5513
 - Kurangi potensi dari titik-titik di dekat pusat kluster dari data pertama sampai dengan data kesepuluh. Contoh pada data pertama

Data Pertama

Potensi titik data pertama atribut pertama terhadap pusat kluster

$$S_{11} = \frac{V_1 - X_{11}}{r_j * q} = \frac{0,5940 - 0,5000}{0,3 * 1,25} = 0,2507$$

Potensi titik data pertama atribut kedua terhadap pusat kluster

$$S_{12} = \frac{V_2 - X_{12}}{r_j * q} = \frac{0,5513 - 0,6250}{0,3 * 1,25} = -0,1965$$

Total kedua nilai pengurangan potensi adalah:

$$ST_1 = \sum_{j=1}^2 S_{1j}^2 = ST_{11}^2 + ST_{12}^2 = 0,2507^2 + -0,1965^2 = 0,1015$$

Nilai pengurangan potensi untuk data pertama adalah:

$$Dc_1 = M * e^{-4[ST_1]} = 3,0759 * e^{-4*0,1015} = 2,0498$$

Nilai potensi baru dihitung dengan mengurangi potensi lama data pertama dengan nilai pengurangan potensi data pertama, yaitu:

$$D_1 = D_1 - Dc_1 = 2,5424 - 2,0498 = 0,4925$$

Langkah di atas dilakukan juga terhadap data ke-2 sampai data ke-10, sehingga hasil akhir potensi baru untuk seluruh data seperti pada Gambar 5.

Potensi Baru:

D1	0.4925
D2	0.1075
D3	0.0000
D4	0.5681
D5	2.0094
D6	1.0174
D7	1.9797
D8	2.5592
D9	2.6608
D10	2.3091

Gambar 5. Hasil Akhir Potensi Baru

Mencari titik dengan potensi baru tertinggi

- a) $Z = \max[D_i | i=1,2,\dots,n] = 2,6608$ yaitu potensi pada data ke-9
- b) $h = i = 9$, sedemikian hingga $D_9 = Z = 2,6608$

Karena Kondisi $\neq 0$ dan $Z \neq 0$, maka lakukan perhitungan untuk **#Iterasi 2** dengan langkah yang sama dengan point c sebelumnya, sehingga diperoleh:

- a) $Z = \max[D_i | i=1,2,\dots,n] = 2,0094$ yaitu potensi pada data ke-5
- b) $h = i = 5$, sedemikian hingga $D_5 = Z = 2,0094$

Karena Kondisi $\neq 0$ dan $Z \neq 0$, maka lakukan perhitungan untuk **#Iterasi 3** dengan langkah yang sama dengan point c sebelumnya, sehingga diperoleh:

- a) $Z = \max[D_i | i=1,2,\dots,n] = 0,3877$ yaitu potensi pada data ke-6
- b) $h = i = 6$, sedemikian hingga $D_6 = Z = 0,3877$

Karena Kondisi $\neq 0$ dan $Z \neq 0$, maka lakukan perhitungan untuk **#Iterasi 4** dengan langkah yang sama dengan point c sebelumnya, sehingga diperoleh:

$$\begin{aligned} \text{Rasio} &= Z / M \\ &= 0,3877 / 3,0759 \\ &= 0,1261 \end{aligned}$$

Rasio \leq reject_ratio (0,5), maka Kondisi = 0

Kondisi = 0 berarti calon pusat kluster tidak diterima sebagai pusat kluster, maka proses dihentikan dengan jumlah kluster yaitu 3 dan pusat kluster seperti pada Gambar 6.

Cluster	Normalisasi		Data ke-	Hasil DeNormalisasi	
	1	2		1	2
1	0.5940	0.5513	3	17,820,000	22,052,000
2	0.8107	0.8775	9	24,320,100	35,100,000
3	0.2667	0.3800	5	8,000,000	15,200,000

Gambar 6. Pusat Kluster

Langkah 5

Mengembalikan pusat kluster dari bentuk ternormalisasi ke bentuk semula. Contoh pada kluster pertama.

Atribut Pertama

$$\begin{aligned} Center_{11} &= Center_{11} * (XMax_1 - XMin_1) + XMin_1 \\ Center_{11} &= 0,5940 * (30.000.000 - 0) + 0 = 17.820.000 \end{aligned}$$

Atribut Kedua

$$\begin{aligned} Center_{12} &= Center_{12} * (XMax_1 - XMin_1) + XMin_1 \\ Center_{12} &= 0,5513 * (40.000.000 - 0) + 0 = 22.052.000 \end{aligned}$$

Lakukan hal yang sama untuk kluster kedua dan ketiga.

Langkah 6

- a. Menghitung nilai sigma kluster

$$\sigma_1 = r * \frac{(XMax_1 - XMin_1)}{\sqrt{8}} = 0,3 * \frac{30.000.000 - 0}{\sqrt{8}} = 3.181.980,5$$

$$\sigma_2 = r * \frac{(XMax_2 - XMin_2)}{\sqrt{8}} = 0,3 * \frac{40.000.000 - 0}{\sqrt{8}} = 4.242.640,69$$

- b. Menghitung derajat keanggotaan suatu titik data X_i pada kluster ke- k . Contoh pada data pertama kluster pertama.
Kluster Pertama

$$\begin{aligned} \mu_{11} &= e^{-\left[\frac{(X_{11} - C_{11})^2}{2\sigma_1^2} + \frac{(X_{12} - C_{12})^2}{2\sigma_2^2}\right]} \\ \mu_{11} &= e^{-\left[\frac{(15.000.000 - 17.820.000)^2}{2 * 3.181.980,5^2} + \frac{(25.000.000 - 22.052.000)^2}{2 * 4.242.640,69^2}\right]} \\ \mu_{11} &= 0,5304 \end{aligned}$$

Lakukan hal yang sama untuk kluster kedua dan ketiga.

Langkah di atas dilakukan juga terhadap data ke-4 sampai data ke-10, sehingga hasil akhir derajat keanggotaan untuk seluruh data seperti pada Gambar 7.

Data ke	Derajat keanggotaan pada cluster ke-			Data cenderung masuk di cluster ke-		
	1	2	3	1	2	3
1	0.5304	0.0008	0.0062	*		
2	0.4655	0.0491	0.0000		*	
3	1.0000	0.0011	0.0023	*		
4	0.6192	0.0000	0.0266	*		
5	0.0023	0.0000	1.0000			*
6	0.4550	0.0000	0.0835	*		
7	0.0009	0.0000	0.9541			*
8	0.0015	0.9430	0.0000		*	
9	0.0011	1.0000	0.0000		*	
10	0.0000	0.7060	0.0000		*	

Gambar 7. Derajat Keanggotaan FSC

Dari Gambar 7 di atas dapat diambil kesimpulan bahwa:

- a. Kelompok 1 (kluster 1), berisi industri-industri kecil ke-: 1, 3, 4, 6
- b. Kelompok 2 (kluster 2), berisi industri-industri kecil ke-: 2, 8, 9, 10
- c. Kelompok 3 (kluster 3), berisi industri-industri kecil ke-: 5, 7

Pada kasus yang sama dengan metode FCM menggunakan parameter berikut:

- a. Jumlah kluster, $c = 3$
- b. Pangkat, $m = 2$
- c. Maksimum iterasi, $MaxIter = 100$
- d. Error terkecil yang diharapkan $= \xi = 1.10^{-6}$
- e. Fungsi objektif awal, $= P_0 = 0$
- f. Iterasi awal, $= t = 1$

dilakukan proses perhitungan hingga iterasi ke-33 karena nilai $|P_{33} - P_{32}| = 0 < \xi = 1.10^{-6}$, diperoleh hasil berupa pusat kluster dan derajat keanggotaan untuk seluruh data seperti pada Gambar 8 dan Gambar 9.

V =

7,565,753	15,257,156
24,953,038	35,712,062
16,618,499	22,134,097

Gambar 8. Pusat Kluster FCM

No. Data	Derajat keanggotaan pada cluster ke-			Data cenderung masuk cluster ke-		
	1	2	3	1	2	3
1	0.0642	0.0451	0.8906			*
2	0.0297	0.1187	0.8516			*
3	0.0009	0.0005	0.9986			*
4	0.0582	0.0064	0.9354			*
5	0.9999	0.0000	0.0001	*		
6	0.0728	0.0052	0.9220			*
7	0.9999	0.0000	0.0001	*		
8	0.0003	0.9981	0.0016		*	
9	0.0000	0.9998	0.0002		*	
10	0.0007	0.9966	0.0027		*	

Gambar 9. Derajat Keanggotaan FCM

Dari Gambar 9 di atas dapat diambil kesimpulan bahwa:

- a. Kelompok 1 (kluster 1), berisi industri-industri kecil ke-: 5, 7
- b. Kelompok 2 (kluster 2), berisi industri-industri kecil ke-: 8, 9, 10
- c. Kelompok 3 (kluster 3), berisi industri-industri kecil ke-: 1, 2, 3, 4, 6

Terdapat perbedaan hasil yang cukup jauh antara hasil perhitungan dengan metode FSC dan FCM seperti pada Gambar 10. Hal ini dapat dipengaruhi oleh pemilihan nilai-nilai parameter yang digunakan pada kedua metode tersebut.

Data	Fuzzy Subtractive Clustering			Fuzzy C-Means		
	Derajat keanggotaan pada cluster ke-			Data cenderung masuk di cluster ke-		
	1	2	3	1	2	3
1	*					*
2		*				*
3	*					*
4	*					*
5			*	*		
6	*					*
7			*	*		
8		*			*	
9		*			*	
10		*			*	

Gambar 10. Perbandingan FSC dan FCM

V. SIMPULAN DAN SARAN

5.1 Simpulan

Dari penelitian yang dilakukan dapat diperoleh kesimpulan bahwa penerapan teknik klustering di dalam pengenalan pola tidak terbimbing pada contoh kasus pengelompokan industri kecil dengan metode **Fuzzy Subtractive Clustering** telah dilakukan dengan menggunakan data berupa 10 industri kecil dengan 2 kriteria yaitu modal awal dan rata-rata penjualan. Berdasarkan parameter jari-jari sebesar 0,3; squash factor sebesar 0,25; accept_ratio sebesar 0,5; reject_ratio sebesar 0,15; XMin sebesar [0;0]; dan XMax sebesar [30.000.000; 40.000.000], proses klustering dihentikan pada iterasi ke-3. Hasil akhir klustering dengan metode FSC telah mengelompokkan industri-industri tersebut ke dalam 3 kelompok kluster yaitu Kluster 1 (Industri ke-1, 3, 6 dan 6), Kluster 2 (Industri ke-2, 8, 9 dan 10), dan Kluster 3 (Industri ke-5 dan 7). Hasil klustering terhadap data yang sama pada metode FCM dengan parameter *error rate* yang ditentukan sebesar 0,000016 proses klustering dihentikan pada iterasi ke-33. Hasil akhir klustering dengan metode FCM telah mengelompokkan industri-industri tersebut ke dalam 3 kelompok kluster yaitu Kluster 1 (Industri ke-5 dan 7), Kluster 2 (Industri ke-8, 9 dan 10), dan Kluster 3 (Industri ke-1, 2, 3, 4 dan 6). Perbedaan hasil klustering dipengaruhi oleh pemilihan nilai-nilai parameter yang digunakan pada kedua metode tersebut.

5.2 Saran

Untuk memperoleh hasil klustering yang lebih baik, disarankan untuk melakukan percobaan terhadap beberapa nilai parameter yang berbeda-beda pada metode klustering.

DAFTAR PUSTAKA

- [1] Kadda Beghdad Bey, Farid Benhammadi, Faouzi Sebbak. "Fuzzy Subtractive Clustering Based Prediction Approach for CPU Load Availability". *The Fourth International Conference on Cloud Computing, Grids and Virtualization, 2013*.
- [2] Kusrini, Luthfi Emha Taufiq. 2009. "Algoritma Data Mining". Yogyakarta: Andi Offset.
- [3] Kusumadewi Sri, Purnomo Hadi. 2010. "Aplikasi Logika Fuzzy untuk Pendukung Keputusan". Yogyakarta: Graha Ilmu.
- [4] Manik Prihatini Putu. "Klustering Berbasis Prototipe dengan Metode Fuzzy C-Means". *Jurnal LOGIC Vol. 15, No. 1, Maret 2015*.
- [5] Padmavathi G, Muthukumar M and Thakur Suresh Kumar. "Non linear Image segmentation using fuzzy c means clustering method with thresholding for underwater images". *IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 3, No 9, May 2010*.
- [6] Pereira R, Fagundes A, Melicio R, Mendes VMF, Figueiredo J., Quadrado J.C. "Fuzzy Subtractive Clustering Technique Applied to Demand Response in a Smart Grid Scope". *Procedia Technology Volume 17, 2014*.
- [7] Sanmorino Ahmad. "Clustering Batik Images using Fuzzy C-Means Algorithm Based on Log-Average Luminance". *Computer Engineering and Applications Vol. 1, No. 1, June 2012*.
- [8] Santosa Budi. 2007. "Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis". Yogyakarta: Graha Ilmu.
- [9] Shahram Mollaiy Berneti, "Design of Fuzzy Subtractive Clustering Model using Particle Swarm Optimization for the Permeability Prediction of the Reservoir". *International Journal of Computer Application Volume 29, Number 11, September 2011*.
- [10] Singh Tejwant, Mahajan Manish. "Performance Comparison of Fuzzy C Means with Respect to Other Clustering Algorithm". *IJARCSSE Volume 4, Issue 5, May 2014*.
- [11] Sreenivasarao Vuda, Vidyavathi S. "Comparative Analysis of Fuzzy C- Mean and Modified Fuzzy Possibilistic C -Mean Algorithms in Data Mining". *IJCST Vol.1, Issue 1, September 2010*.
- [12] Suganya R, Shanthi R. "Fuzzy C-Means Algorithm-A Review". *International Journal of Scientific and Research Publications, Volume 2, Issue 11, November 2012*.
- [13] Susanto Sani, Suryadi Dedy. 2010. "Pengantar Data Mining". Yogyakarta: Andi Offset.