ANALISIS PERBANDINGAN METODE *FUZZY C-MEANS* DAN *SUBTRACTIVE FUZZY C-MEANS*

Baiq Nurul Haqiqi[1], Robert Kurniawan[2]
[1,2]Jurusan Komputasi Statistik, Sekolah Tinggi Ilmu Statistik (STIS)
Email: [1] qiqi0693@naver.com, [2] robertk@stis.ac.id

Abstract

Fuzzy C-Means (FCM) is one of the most frequently used clustering method. However FCM has some disadvantages such as number of clusters to be prespecified and partition matrix to be randomly initiated which makes clustering result becomes inconsistent. Subtractive Clustering (SC) is an alternative method that can be used when number of clusters are unknown. Moreover, SC produces consistent clustering result. A hybrid method of FCM and SC called Subtractive Fuzzy C-Means (SFCM) is proposed to overcome FCM's disadvantages using SC. Both SFCM and FCM are implemented to cluster generated data and the result of the two methods are compared. The experiment shows that generally SFCM produces better clustering result than FCM based on six validity indices.

Keywords : Clustering, Fuzzy C-Means, Subtractive Clustering, Subractive Fuzzy C-Means

1. Pendahuluan

Pengelompokan (*clustering*) merupakan salah satu teknik yang paling penting dalam *data mining*[8]. Salah satu metode pengelompokan yang paling sering digunakan adalah *Fuzzy C-Means* (FCM). Metode FCM memiliki beberapa kelemahan, antara lain membutuhkan banyaknya kelompok dan matriks keanggotaan kelompok yang ditetapkan sebelumnya[6]. Biasanya, matriks keanggotaan kelompok awal diinisialisasikan secara acak yang menyebabkan metode FCM memiliki masalah inkonsistensi.

Alternatif metode pengelompokan lainnya yang dapat digunakan jika jumlah kelompok tidak diketahui sebelumnya adalah metode *Substractive Clustering* (SC). SC memperoleh hasil yang lebih konsisten dibandingkan dengan FCM[2]. Selain itu, SC memiliki kecepatan yang lebih baik dibandingkan FCM, namun SC memiliki akurasi yang lebih rendah dibandingkan dengan FCM[4].

Untuk menjembatani kekurangan dan kelebihan kedua metode tersebut diusulkan sebuah metode baru yang merupakan penggabungan (*hybrid*) dari keduanya yang disebut *Subtractive Fuzzy C-Means* (SFCM). Metode ini digunakan oleh Liu, Xiao, Wang, Shi, dan Fang[7] dalam penelitiannya dan menyimpulkan bahwa SFCM secara umum memberikan solusi yang lebih baik dibandingkan dengan FCM serta memberikan tingkat kecepatan yang lebih tinggi dalam hal konvergensi fungsi objektif. Kelebihan metode SFCM lainnya ditunjukkan oleh Hossen, Rahman, Sayeed, Samsuddin, dan Rokhani[4] dengan menggunakan indeks validitas *Partition Coefficient* (PC) yang menyimpulkan bahwa SFCM dapat meningkatkan kecepatan, mengurangi jumlah iterasi, menghasilkan partisi data yang lebih stabil, dan lebih akurat.

Penelitian ini bertujuan untuk membandingkan kehandalan dari hasil pengelompokan menggunakan metode FCM dan SFCM, yang berdasarkan studi literatur, bahwa belum pernah ada penelitian lain yang membandingkan antara kedua metode tersebut. Pada penelitian ini digunakan data bangkitan dengan distribusi Uniform (0,1). Hasil pengelompokan terbaik dipilih berdasarkan berdasarkan indeks validitas *Partition*

Coefficent (PC), *Modified Partition Coefficient* (MPC), *Classification Entropy* (CE), *Partition Index* (PI), *Fukuyama Sugeno Index* (FS), dan *Xie Beni Index* (XB).

2. Tinjauan Pustaka

2.1. *Fuzzy C-Means* (FCM)

Metode ini ditemukan pertama kali oleh Dunn pada tahun 1973 kemudian dikembangkan lagi oleh Bezdek pada tahun 1981. Ide dasar dari metode ini mirip dengan metode *K-Means*. FCM didasarkan pada logika *fuzzy*, setiap titik data dimasukkan ke suatu kelompok berdasarkan nilai keanggotaannya pada kelompok tersebut.

Algoritma FCM adalah sebagai berikut[5]:

1) Menentukan banyak kelompok (c), *fuzzifier* (m), maksimum iterasi (*MaxIter*), perubahan nilai fungsi objektif terkecil yang diharapkan (ε), fungsi objektif awal ($P_0 = 0$), dan iterasi awal ($t = 1$)
2) Membangkitkan bilangan random u_{ik} dengan i merupakan banyak data dan k merupakan banyak kelompok sebagai elemen-elemen awal matriks keanggotaan awal U.
3) Menghitung pusat kelompok ke-i dengan persamaan:

$$p_i = \frac{\sum_{k=1}^{N} (u_{ik})^m x_k}{\sum_{k=1}^{N} (u_{ik})^m} \tag{1}$$

   dengan u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i, X_k adalah objek data ke-k, N adalah banyaknya objek penelitian, dan m adalah *fuzzifier*.
4) Menghitung fungsi objektif pada iterasi ke-t dengan persamaan:

$$J(P,U,X,c,m) = \sum_{i=1}^{c} \sum_{k=1}^{N} (u_{ik})^m d_{ik}^2 (x_k, p_i) \tag{2}$$

   dengan c adalah banyak kelompok yang diinginkan, N adalah banyak objek penelitian, u_{ik} adalah nilai keanggotaan objek ke-k pada kelompok ke-i yang merupakan bagian dari matriks U, m adalah *fuzzifier*, dan $d_{ik}^2(x_k, p_i)$ adalah jarak antara vektor pengamatan ke-k dengan pusat kelompok ke-i.
5) Menghitung perubahan matriks keanggotaan dengan persamaan:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}^2}{d_{jk}^2} \right)^{\frac{1}{m-1}}} \tag{3}$$

   dengan u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i, d_{ik}^2 adalah jarak antara objek ke-k dengan pusat kelompok ke-i, d_{jk}^2 adalah jarak antara objek ke-k dengan pusat kelompok ke-j, dan m adalah *fuzzifier*.
6) Cek kondisi berhenti
   - Jika $|\mathrm{J_t} - \mathrm{J_{t-1}}| < \varepsilon$ atau $t > MaxIter$ maka berhenti;
   - Jika tidak : $t = t + 1$, ulangi langkah ke-3

2.2. *Subtractive Fuzzy C-Means* (SFCM)

Subtractive Clustering (SC) merupakan metode pengelompokan yang diperkenalkan oleh Stephen L. Chiu pada tahun 1994. Metode ini merupakan modifikasi dari *Mountain Method* (MM) yang diperkenalkan oleh Yager dan Filev pada tahun 1992.

Dimisalkan terdapat *n* buah titik data $\{x_1, x_2,.., x_n\}$ dalam sebuah ruang berdimensi *M*. Setiap titik data dipertimbangkan sebagai kandidat pusat kelompok yang nilai potensinya dihitung dengan persamaan[3]:

$$P_i = \sum_{j=1}^{n} e^{-\frac{4}{r_\alpha^2}.\left\|x_i - x_j\right\|^2} \tag{4}$$

P_i adalah nilai potensi untuk titik data ke-i atau potensi dari nilai x_i, $\|.\|$ menotasikan jarak Euclidean, dan r_α adalah sebuah konstanta positif. Konstanta r_α merupakan radius atau jari-jari yang mendefinisikan sebuah lingkungan tetangga. Data yang memiliki potensi yang tinggi adalah data yang memiliki jumlah tetangga paling banyak.

Setelah nilai potensi dari semua titik data dihitung, titik data dengan nilai potensi paling tinggi dipilih sebagai pusat kelompok pertama. Misalkan x_1^* adalah data yang terpilih sebagai pusat kelompok pertama dan P_1^* sebagai nilai potensinya. Nilai potensi untuk setiap titik data diperbaharui dengan persamaan[7]:

$$P_i = P_i - P_1^* e^{-\frac{4}{r_\beta^2}.\left\|x_i - x_1^*\right\|^2} \tag{5}$$

dimana r_β merupakan konstanta positif.

Setelah nilai potensi tiap titik data diperbaharui, titik data dengan nilai potensi terbesar dipilih sebagai pusat kelompok kedua. Selanjutnya nilai potensi tiap titik data kembali diperbaharui. Proses ini akan terus berlanjut sampai diperoleh jumlah kelompok yang cukup.

Subtractive Fuzzy C-Means (SFCM) merupakan penggabungan antara metode pengelompokan *Subtractive Clustering* (SC) dan *Fuzzy C-Means* (FCM). SC digunakan untuk menentukan jumlah kelompok dan matriks keanggotaan awal FCM sehingga inisialisasi secara random tidak perlu dilakukan. Secara garis besar algoritma metode SFCM adalah sebagai berikut:

1) Menghitung potensi awal tiap titik data dengan Persamaan (4)
2) Menetapkan titik data dengan potensi paling tinggi sebagai pusat kelompok pertama dan P_1^* sebagai nilai potensinya.
3) Perbaharui potensi tiap titik data dengan Persamaan (5)
4) Jika rasio potensi titik data tertinggi terbaru dengan potensi titik data tertinggi sebelumnya lebih besar dari *accept ratio* maka titik data dengan potensi tertinggi terbaru ditetapkan menjadi pusat kelompok dan tahap 3 diulangi kembali. Jika rasio kurang dari *reject ratio* maka iterasi dihentikan berlanjut ke tahap 5.
5) Menghitung matriks keanggotaan awal FCM berdasarkan pusat kelompok yang didapatkan melalui metode SC menggunakan Persamaan (3). Jika terdapat jarak d_{ik}^2 yang bernilai nol maka u_{ik} (nilai keanggotaan data k pada kelompok ke-i) akan bernilai 1 dan u_{jk} (nilai keanggotaan data k pada kelompok lainnya) akan bernilai 0[7].
6) Melakukan tahapan 3 sampai dengan 6 pada algoritma FCM hingga ditemukannya matriks keanggotaan dan pusat kelompok terakhir.

2.3. Indeks Validitas

Beberapa indeks validitas yang sering digunakan dalam penelitian-penelitian adalah :

1) *Partition Coefficient* (PC)

Indeks ini mengukur jumlah *overlapping* antarkelompok. Persamaan indeks PC oleh Bezdek sebagai berikut[9] :

$$PC(c) = \frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^2 \tag{6}$$

dengan *N* adalah banyak objek penelitian, c adalah banyak kelompok, dan u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i. Indeks ini memiliki rentang 1/c sampai 1. Jumlah kelompok yang optimal ditunjukkan oleh nilai PC yang paling besar.

2) *Classification Entropy* (CE)

CE hanya mengukur kekaburan (*fuzziness*) dari partisi kelompok. Persamaan indeks dapat dituliskan sebagai berikut[9]:

$$CE(c) = -\frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}\ln(u_{ik}) \tag{7}$$

dengan *N* adalah banyak objek penelitian, *c* adalah banyak kelompok, dan u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i. Indeks ini memiliki rentang 0 sampai ln(c). Indeks CE yang semakin kecil menunjukkan pengelompokan yang lebih baik.

3) *Partition Index (PI)/Separation and Compactness* (SC)

PI merupakan rasio antara jumlah kepadatan dan pemisahan kelompok-kelompok. Indeks ini ditulis menjadi persamaan sebagai berikut[9]:

$$PI(c) = \sum_{i=1}^{c}\frac{\sum_{k=1}^{N}(u_{ik})^m\|x_k - v_i\|^2}{N_i\sum_{j=1}^{c}\|v_j - v_i\|^2} \tag{8}$$

dengan *N* adalah banyak objek penelitian, N_i adalah banyak objek penelitian kelompok ke-i, *c* adalah banyak kelompok, u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i, *m* adalah *fuzzifier*, $\|x_k - v_i\|$ adalah jarak euclidean titik data (x_k) dengan pusat kelompok v_i, dan $\|v_j - v_i\|$ jarak euclidean antar pusat kelompok. Nilai SC yang rendah mengindikasikan partisi kelompok yang lebih baik.

4) *Fukuyama Sugeno Index* (FS)

Persamaan indeks oleh Fukuyama dan Sugeno sebagai berikut[9]:

$$FS(c) = \sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik})^m\|x_k - v_i\|^2 - \sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik})^m\|v_i - \bar{v}\|^2 \tag{9}$$

dengan *N* adalah banyak objek penelitian, c adalah banyak kelompok, u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i, *m* adalah *fuzzifier*, $\|x_k - v_i\|$ adalah jarak euclidean titik data (x_k) dengan pusat kelompok v_i, $\|v_i - \bar{v}\|$ adalah jarak euclidean pusat kelompok v_i dengan rata-rata pusat kelompok. Nilai FS yang rendah mengindikasikan partisi kelompok yang lebih baik.

5) *Xie and Beni's Index (XB)*

XB bertujuan untuk menghitung rasio total variasi di dalam kelompok dan pemisahan kelompok. Indeks ini dituliskan sebagai berikut[9] :

$$XB(c) = \frac{\sum_{i=1}^{c}\sum_{k=1}^{N}(u_{ik})^m\|x_k - v_i\|^2}{N\min_{i,k}\|v_k - v_i\|^2} \tag{10}$$

dengan *N* adalah banyak objek penelitian, c banyak kelompok, u_{ik} adalah nilai keanggotaan objek ke-k dengan pusat kelompok ke-i. *m* adalah *fuzzifier*, $\|x_k - v_i\|$ adalah jarak euclidean titik data (x_k) dengan pusat kelompok v_i,dan $\|v_k - v_i\|$ adalah jarak euclidean antar pusat kelompok. Nilai XB yang rendah mengindikasikan partisi kelompok yang lebih baik.

6) *Modified Partition Coefficient (MPC).*

Indeks ini diajukan oleh Dave (1996) untuk mengatasi kekurangan PC dan CE. Nilai PC dan CE memiliki kecenderungan berubah secara monoton seiring dengan berubahnya nilai c (Wang dan Zhang, 2007). Persamaan indeks ini dituliskan sebagai berikut[9]:

$$MPC(c) = 1 - \frac{c}{c-1}\left(1 - PC\right) \tag{11}$$

dengan *c* adalah banyak kelompok dan PC adalah indeks PC.

3. Metodologi Penelitian

3.1. Sumber Data

Data yang digunakan pada penelitian ini merupakan data bangkitan berdistribusi Uniform dengan rentang 0-1 yang diperoleh melalui fungsi *runif()* pada *software* R. Jumlah data bangkitan yang digunakan adalah sebanyak 20 dan 100 data yang masing-masing terdiri dari tiga variabel

3.2. Metode Penelitian

Metode analisis penelitian ini adalah sebagai berikut:

1) Melakukan pengelompokan menggunakan metode FCM dan SFCM pada data bangkitan 20 dan 100 data.
2) Menghitung enam indeks validitas hasil pengelompokan metode FCM dan SFCM
3) Menentukan jumlah kelompok optimal FCM dan SFCM berdasarkan enam indeks validitas
4) Membandingkan hasil pengelompokan FCM dan SFCM pada jumlah kelompok optimal.

4. Hasil dan Pembahasan

4.1. Data Bangkitan 20 Data

Pengelompokan metode FCM dilakukan pada jumlah kelompok sebanyak dua sampai dengan sepuluh. Parameter FCM lainnya ditetapkan sama untuk tiap jumlah kelompok yaitu *fuzzifier* sebesar 2, iterasi maksimum sebanyak 1000 kali, dan ε sebesar 10^{-5}. Indeks validitas yang didapatkan untuk tiap jumlah kelompok seperti tercantum pada Tabel 1.

Tabel 1. Indeks Validitas Hasil Pengelompokan FCM Data Bangkitan 20

c	PC	MPC	CE	FS	XB	PI
2	0,6635	0,3271	0,5115	1,2777	0,4598	0,9412
3	0,5860	0,3789	0,7346	0,0304	0,2412	0,3088
4	0,5818	0,4424	0,8195	-1,0955	0,1603	0,1210
5	0,5540	0,4425	0,9174	-1,4625	0,2047	0,0761
6	0,5438	0,4526	0,9842	-1,6484	0,1528	0,0624
7	0,5596	0,4861	0,9988	-1,8248	0,1936	0,0495
8	0,5514	0,4873	1,0456	-1,8670	0,1734	0,0418
9	0,5751	0,5220	1,0245	-1,9909	0,1371	0,0376
10	0,6010	0,5567	0,9881	-2,1677	0,1306	0,0328

Dari Tabel 1 dijelaskan bahwa jumlah kelompok sebanyak sepuluh mempunyai indeks MPC, FS, XB, dan PI terbaik sehingga dapat disimpulkan bahwa berdasarkan metode FCM jumlah kelompok optimal pada data tersebut sebanyak sepuluh kelompok.

Hal yang sama dilakukan menggunakan metode SFCM. Untuk metode SFCM radius diubah-ubah sampai didapatkan jumlah kelompok sebanyak dua sampai dengan sepuluh. Parameter SFCM lainnya ditetapkan sama untuk tiap radius yaitu *squash factor* sebesar 1.5, *accept ratio* sebesar 0.5, *reject ratio* sebesar 0.15, *fuzzifier* sebesar 2, iterasi maksimum sebanyak 1000 kali, dan ε sebesar 10^{-5}. Indeks validitas yang didapatkan untuk tiap jumlah kelompok seperti pada Tabel 2.

Tabel 2. Indeks Validitas Hasil Pengelompokan SFCM Data Bangkitan 20

r	c	PC	MPC	CE	FS	XB	PI
0,68	2	0,6635	0,3271	0,5115	1,2777	0,4598	0,9412
0,65	3	0,5860	0,3789	0,7346	0,0304	0,2412	0,3088
0,6	4	0,5818	0,4424	0,8195	-1,0955	0,1603	0,1210
0,55	5	0,5540	0,4425	0,9174	-1,4625	0,2047	0,0761
0,45	6	0,5533	0,4639	0,9643	-1,8919	0,2072	0,0535
0,44	7	0,5725	0,5013	0,9596	-2,3679	0,1624	0,0366
0,41	8	0,5898	0,5312	0,9607	-2,5842	0,1836	0,0308
0,4	9	0,6096	0,5608	0,9393	-2,7831	0,0970	0,0217
0,35	10	0,6183	0,5759	0,9369	-2,9008	0,1013	0,0168

Berdasarkan Tabel 2 terlihat bahwa jumlah kelompok sebanyak sepuluh memberikan hasil terbaik pada indeks MPC, FS, dan PI. Dapat disimpulkan bahwa berdasarkan metode SFCM jumlah kelompok optimal pada data tersebut sebanyak sepuluh kelompok. Perbandingan hasil pengelompokan optimal kedua metode tersebut seperti yang tercantum pada Tabel 3.

Tabel 3. Perbandingan Hasil FCM dan SFCM Data Bangkitan 20

Metode	Waktu Komputasi (detik)	Fungsi Objektif Terakhir	PC	MPC	CE	FS	XB	PI
FCM	14,8819	0,20634	0,6010	0,5567	0,988	-2,1677	0,1306	0,0328
SFCM	14,1240	0,16256	0,6183	0,5759	0,937	-2,9008	0,1013	0,0168

Tabel 3 menginformasikan bahwa metode SFCM memberikan hasil pengelompokan yang lebih baik berdasarkan keenam indeks. Selain itu dengan iterasi sebanyak 1000 kali, metode SFCM memberikan nilai fungsi objektif terakhir lebih kecil dibandingkan FCM. Dari segi waktu komputasi kecepatan metode SFCM memberikan hasil yang lebih cepat 0,7579 detik. Selain itu, berdasarkan dua tabel hasil perhitungan sebelumnya terlihat bahwa secara umum metode SFCM memberikan hasil pengelompokan yang lebih baik dibandingkan metode FCM. Jika dikaitkan dengan inkonsistensi FCM, pada jumlah kelompok dua sampai dengan lima FCM memberikan hasil yang stabil. Pada saat itulah nilai yang dihasilkan oleh FCM dan SFCM persis sama. Namun pada jumlah kelompok di atas lima FCM memberikan hasil yang bervariasi setiap kali pengelompokan dilakukan.

4.2. Data Bangkitan 100 Data

Tahapan yang sama seperti pada pengelompokan data bangkitan 20 data dilakukan pada data bangkitan 100 data. Indeks validitas yang didapatkan untuk tiap jumlah kelompok menggunakan metode FCM seperti tercantum pada Tabel 4.

Tabel 4. Indeks Validitas Hasil Pengelompokan FCM Data Bangkitan 100

c	PC	MPC	CE	FS	XB	PI
2	0,6354	0,2708	0,5443	7,8340	0,4708	0,9433
3	0,5275	0,2913	0,8139	2,2282	0,3243	0,3807
4	0,4590	0,2787	1,0213	-0,3445	0,2823	0,2319
5	0,4214	0,2768	1,1645	-2,0994	0,2690	0,1405
6	0,3999	0,2799	1,2738	-3,1684	0,2065	0,1069
7	0,3862	0,2839	1,3571	-4,1055	0,3287	0,0821
8	0,3840	0,2960	1,4133	-4,9074	0,3823	0,0689
9	0,3716	0,2930	1,4876	-5,1729	0,2512	0,0576
10	0,3783	0,3093	1,5124	-6,0103	0,2148	0,0473

Berdasarkan Tabel 4 terlihat bahwa jumlah kelompok sebanyak sepuluh memberikan hasil terbaik pada indeks MPC, FS, dan PI sehingga dapat disimpulkan bahwa berdasarkan metode FCM jumlah kelompok optimal pada data tersebut sebanyak sepuluh kelompok. Sedangkan indeks validitas pengelompokan menggunakan metode SFCM untuk tiap jumlah kelompok didapatkan seperti tercantum pada Tabel 5.

Tabel 5. Indeks Validitas Hasil Pengelompokan SFCM Data Bangkitan 100

r	c	PC	MPC	CE	FS	XB	PI
0,79	2	0,6354	0,2708	0,5443	7,8340	0,4708	0,9433
0,75	3	0,5275	0,2913	0,8139	2,2282	0,3243	0,3807
0,7	3	0,5275	0,2913	0,8139	2,2282	0,3243	0,3807
0,65	3	0,5275	0,2913	0,8139	2,2282	0,3243	0,3807
0,6	3	0,5275	0,2913	0,8139	2,2282	0,3243	0,3807
0,55	4	0,4590	0,2787	1,0213	-0,3445	0,2823	0,2319
0,53	5	0,4214	0,2768	1,1645	-2,0994	0,2690	0,1405
0,5	6	0,3999	0,2799	1,2738	-3,1684	0,2065	0,1069
0,465	7	0,3916	0,2902	1,3504	-4,3800	0,1790	0,0831
0,45	8	0,3819	0,2937	1,4235	-5,0889	0,1800	0,0679
0,4	9	0,3756	0,2975	1,4758	-5,3989	0,2965	0,0555
0,37	10	0,3759	0,3066	1,5154	-6,0077	0,2478	0,0463

Pada Tabel 5 menjelaskan bahwa jumlah kelompok sebanyak sepuluh memberikan hasil terbaik pada indeks MPC, FS, dan PI sehingga disimpulkan bahwa berdasarkan metode SFCM jumlah kelompok optimal pada data tersebut sebanyak sepuluh kelompok. Perbandingan hasil pengelompokan optimal kedua metode tersebut tercantum pada Tabel 6.

Berdasarkan Tabel 6 dapat disimpulkan bahwa metode FCM memberikan hasil pengelompokan yang lebih baik berdasarkan lima indeks kecuali PI. Selain itu dengan iterasi sebanyak 1000 kali, metode FCM memberikan nilai fungsi objektif lebih kecil dibandingkan SFCM. Dari segi waktu komputasi, kecepatan metode FCM memberikan hasil yang lebih cepat 14,6904 detik. Namun, jika dilihat secara keseluruhan, metode

SFCM memberikan hasil pengelompokan yang lebih baik dibandingkan metode FCM. Hasil pengelompokan kedua metode sama pada jumlah kelompok dua sampai dengan enam. Namun, pada jumlah kelompok tujuh dan sembilan SFCM memberikan hasil yang lebih baik dibandingkan FCM. Pada jumlah kelompok delapan kedua metode memberikan hasil seri dimana masing-masing metode dinilai baik oleh tiga indeks berbeda.

Tabel 6. Perbandingan Hasil Optimum FCM dan SFCM Data Bangkitan 100

Metode	**Waktu Komputasi (detik)**	**Fungsi Objektif Terakhir**	**PC**	**MPC**	**CE**	**FS**	**XB**	**PI**
FCM	52,9887	1,8584	0,3783	0,3093	1,512	-6,0103	0,2148	0,0473
SFCM	67,9491	1,8659	0,3759	0,3066	1,515	-6,0077	0,2478	0,0463

5. Kesimpulan

Sesuai dengan tujuan dari penelitian, berdasarkan hasil uji coba dengan menggunakan data bangkitan menghasilkan kesimpulan bahwa pengelompokan menggunakan metode SFCM memberikan hasil lebih baik dibandingkan metode FCM pada data sebanyak 20, sedangkan FCM memberikan hasil lebih baik dibandingkan metode SFCM pada data sebanyak 100 dimana jumlah kelompok optimal yang diberikan oleh kedua metode pada kedua data bangkitan sama. Ketika hasil yang diberikan FCM pada tiap kali perulangan adalah konsisten, hasil yang diberikan kedua metode persis sama. Jika dilihat secara keseluruhan pada umumnya SFCM memberikan hasil pengelompokan yang lebih baik dibandingkan FCM pada jumlah kelompok sama. Untuk penelitian selanjutnya disarankan untuk membandingkan jarak yang dipakai dalam kedua metode tersebut, misalnya dengan jarak manhattan dan mahalanobis. Diharapkan dengan perbedaan jarak yang digunakan akan menghasilkan kehandalan yang lebih baik lagi.

DAFTAR PUSTAKA

1. Balasko, B., Abonyi, J. dan Feil, B., *Fuzzy Clustering and Data Analysis Toolbox (for Use With Matlab)*, [online] Available: http://www.fmt.vein.hu/softcomp/fclusttoolbox/.
2. Bataineh, K.M., Naji, M., dan Saqer,M., A Comparison Study between Various Fuzzy Clustering Algorithms, *Jordan Journal of Mechanical and Industrial Engineering*, 2011, Vol.5, No.4: 335-343.
3. Chiu, S.L., Fuzzy Model Indentification Based on Cluster Estimation, *Journal of Intelligent and Fuzzy Systems*, 1994, Vol. 2: 267-278.
4. Hossen, J., Rahman, A., Sayeed, S., Samsuddin, K., dan Rokhani, F, A Modified Hybrid Fuzzy Clustering Algorithm for Data Partitions, *Australian Journal of Basic and Applied Sciences*, 2011, Vol.5, No.8: 674-681.
5. Kusumadewi, S. dan Purnomo, H., *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*, Graha Ilmu, Yogyakarta, 2010.
6. Le, T. dan Altman, T., A new initialization method for the Fuzzy C-Means Algorithm using Fuzzy Subtractive Clustering, *Proc. International Conference on Information and Knowledge Engineering*, Las Vegas USA, 2011, Vol. 1: 144-150.
7. Liu, W.Y., Xiao, C.J., Wang, B.W., Shi, Y., dan Fang, S.F., Study on Combining Subtractive Clustering with Fuzzy C-Means Clustering, *International Conference on Machine Learning and Cybernetics*, 2003, Vol. 5: 2659 – 2662.

8. Suman dan Mittal, P., Comparison and Analysis of Various Clustering Methods in Data mining On Education Data Set Using the WEAK Tool, *International Journal of Emerging Trends & Technology in Computer Science*, 2014, Vol. 3, Issue 2.
9. Wang, W. dan Zhang, Y., On Fuzzy Cluster Validity Indices, *Fuzzy Sets System*, 2007, Vol. 158, No. 19: 2095-2117.