

Application of Fuzzy C-Means in Grouping Districts/Cities Based on Health Service Facilities in East Java

Wardatul Maghfiroh¹, Nurissaidah Ulinnuha²

^{1,2}Program Studi Matematika Universitas Islam Sunan Ampel Surabaya

¹viviro13elf@gmail.com , ²nuris.ulinnuha@uinsby.ac.id

Abstract— Health is a very important thing for every human being because without good health, then humans will be difficult to do activities. We need health facilities that can support human health or society. This study discussed use of clustering algorithm in grouping districts or cities in East Java according to the number of health care facilities using Fuzzy C-Means. The data source of this research got from Central Bureau of Statistics of East Java. The cluster results obtained then validated with silhouette coefficient and purity. With the centroid gained in the last iteration, four districts/cities were included in the first cluster, 26 districts/cities included in the second cluster, and 8 districts/cities included in the third cluster. The results of clustering validation is the value of silhouette coefficient of 0.695 and the purity value of 1. This can be a suggestion to the East Java provincial government, districts / municipalities that are more concerned with having the number of health facilities based on the cluster that has been done.

Keywords— data mining; health facilities; clustering; fuzzy c-means; silhouette coefficient; purity

Abstrak— Kesehatan adalah hal yang sangat penting bagi setiap manusia karena tanpa kesehatan yang baik, maka manusia akan sulit untuk melakukan aktivitas. Untuk menunjang kesehatan pada manusia diperlukan suatu fasilitas kesehatan yang dapat menunjang kesehatan manusia atau masyarakat. Dalam penelitian ini membahas pemanfaatan algoritma *clustering* dalam mengelompokkan kabupaten atau kota di Jawa Timur menurut jumlah fasilitas pelayanan kesehatan menggunakan *Fuzzy C-Means*. Sumber data penelitian ini berasal dari BPS Jawa Timur. Hasil *cluster* yang didapat kemudian divalidasi dengan *silhouette coefficient* dan *purity*. Dengan *centroid* yang telah didapatkan pada iterasi terakhir, diperoleh empat kabupaten/kota termasuk dalam *cluster* pertama, 26 kabupaten/kota termasuk dalam *cluster* kedua, dan 8 kabupaten/kota termasuk dalam *cluster* ketiga. Dari validasi *clustering* yang dilakukan, didapatkan nilai *silhouette coefficient* sebesar 0,695 dan nilai *purity* sebesar 1. Hal ini dapat dijadikan masukan pada pemerintah provinsi Jawa Timur, kabupaten /kota yang menjadi perhatian lebih yang memiliki jumlah fasilitas kesehatan berdasarkan *cluster* yang telah dilakukan.

Kata kunci— data mining, fasilitas kesehatan; clustering; fuzzy c-means; silhouette coefficient; purity

I. PENDAHULUAN

Menurut Depkes RI (2009) pelayanan kesehatan merupakan upaya yang diselenggarakan sendiri atau secara bersama-sama dalam suatu organisasi untuk memelihara dan meningkatkan kesehatan, mencegah, dan menyembuhkan penyakit serta memulihkan kesehatan, perorangan, keluarga, kelompok ataupun masyarakat [1]. Pelayanan kesehatan tidak hanya dilakukan oleh organisasi non pemerintah saja, pemerintah juga melakukan suatu pelayanan kesehatan. Dan pelayanan kesehatan dalam suatu pemerintah kota/kabupaten adalah sebagai tolak ukur kesuksesan pemerintah dalam menyejahterakan masyarakatnya. Jika pelayanan kesehatan oleh pemerintah semakin meningkat maka kesejahteraan masyarakat juga semakin meningkat, begitupun sebaliknya. Dalam upaya meningkatkan suatu pelayanan kesehatan terdapat beberapa faktor yang harus ditingkatkan, salah satunya adalah faktor fasilitas kesehatan. Menurut Undang-Undang Nomor 36 Tahun 2009 tentang Kesehatan, menyatakan bahwa fasilitas kesehatan adalah suatu alat atau tempat yang digunakan untuk menyelenggarakan upaya pelayanan kesehatan, baik promotif, preventif (pencegahan), kuratif (penyembuhan), maupun rehabilitatif yang dilakukan oleh pemerintah, pemerintah daerah, atau masyarakat.

Fasilitas kesehatan yang ada pada tiap kabupaten/kota di Jawa Timur ini masih terbilang belum memadai. Pemerintah Jawa Timur juga belum mengetahui

daerah kabupaten/kota mana saja yang fasilitas kesehatannya sudah terbilang memadai atau yang belum memadai. Untuk mengatasi permasalahan tersebut maka diperlukan suatu metode agar dapat mengetahui kabupaten/kota mana yang termasuk dalam fasilitas kesehatan memadai atau belum memadai. Dalam penelitian ini akan menggunakan analisis kluster. Analisis kluster adalah suatu metode yang digunakan dengan tujuan mengelompokkan objek-objek berdasarkan karakteristik yang dimiliki. Dengan analisis kluster akan didapat kelompok-kelompok kabupaten/kota yang memiliki fasilitas yang memadai atau yang kurang memadai. Dengan mengetahui kelompok-kelompok tersebut maka pemerintah Jawa Timur dapat mengetahui kabupaten/kota mana yang harus ditingkatkan fasilitas kesehatannya.

Dalam penelitian Aditya Ramadhan dkk, yang meneliti tentang perbandingan *K-Means* dan *Fuzzy C-means* untuk pengelompokan Data *User Knowledge Modeling* menghasilkan bahwa metode *Fuzzy C-Means* adalah metode yang lebih baik untuk melakukan *clustering* pada data user knowledge modelling dikarenakan nilai validitasnya lebih mendekati nilai 1[2]. Juga dalam penelitian Trissy Anjar Risqiyani dkk, yang berjudul Pengelompokan Kabupaten Kota di Provinsi Jawa Tengah dengan *Fuzzy C-Means Clustering* dengan studi kasus gizi buruk, faktor sarana dan tenaga kesehatan serta faktor kependudukan di Jawa Tengah pada tahun 2014 menghasilkan hasil evaluasi kluster yang menunjukkan nilai *Sum Square Error* (SSE) dengan metode

SOM memiliki nilai yang lebih kecil dari hasil *clustering* dengan metode *Fuzzy C-Means* [3]. R. Handoyo dkk menggunakan perbandingan metode Single Linkage Dan K - Means untuk pengelompokan dokumen, pada penelitian tersebut mengemukakan metode K-means dapat dikembangkan untuk memberikan inisialisasi awal dalam menentukan start cluster value [4]. Dari ketiga penelitian tersebut dapat disimpulkan bahwa dalam meng-*cluster* lebih bagus menggunakan metode *Fuzzy C-Means*. Dalam penelitian ini, digunakan metode *Fuzzy C-Means* dalam mengelompokkan kabupaten/kota di Jawa Timur berdasarkan jumlah fasilitas kesehatan yang ada.

II. TINJAUAN PUSTAKA

A. Fasilitas Pelayanan Kesehatan

Menurut, peraturan pemerintah RI No 47 Tahun 2016, Fasilitas Pelayanan Kesehatan merupakan suatu alat atau tempat yang digunakan untuk menyelenggarakan upaya pelayanan kesehatan, baik promotif, preventif, kuratif maupun rehabilitatif yang dilakukan oleh pemerintah pusat, pemerintah daerah, dan masyarakat. Jenis fasilitas pelayanan kesehatan adalah sebagai berikut :

- a. Tempat praktik mandiri Tenaga Kesehatan'
- b. Pusat kesehatan masyarakat,
- c. Klinik,
- d. Rumah sakit,
- e. Apotek,
- f. Unit transfusi darah,
- g. Laboratorium kesehatan,
- h. Optikal,
- i. Fasilitas pelayanan kedokteran untuk kepentingan hukum, dan
- j. Fasilitas pelayanan kesehatan tradisional

Untuk menunjang pelayanan kesehatan, penelitian Suhartoyo & Wijaya, 2016 menggunakan metode Neuro-Fuzzy dan teknik COCOMO II untuk mengetahui estimasi biaya pembangunan perangkat lunak rumah sakit, dimana semakin tinggi tipe rumah sakit akan mengeluarkan biaya yang besar meskipun biaya pembangunan rumah sakit sama [5].

B. Data Mining

Data mining adalah proses mengekstrak atau menggali dari data yang sangat besar. Data yang diolah dengan teknik data mining ini kemudian menghasilkan suatu pengetahuan baru yang bersumber dari data lama. Hasil dari pengolahan data dapat digunakan untuk menentukan di masa depan. Penugasan data mining dapat diklasifikasikan menjadi dua kategori, deskriptif dan prediktif. Kategori deskriptif digunakan untuk mencari karakteristik pola data secara umum pada *database*. Metode data mining yang termasuk kategori deskriptif yaitu *cluster analysis* [6].

Data preprocessing bertujuan untuk menghasilkan data yang berkualitas sehingga hasil pengolahan data tersebut diharapkan dapat berkualitas. *Data preprocessing* mencakup

data clearing, seperti menangani *missing value* dan data *noisy* atau *outlier*. Pembersihan data (*data cleaning*) berfungsi untuk menangani *missing value*, memperhalus noise dan memperbaiki data data yang tidak konsisten. Transformasi data dapat dilakukan dengan agregasi dan normalisasi. Agregasi dilakukan pada data ketika operasi perhitungan kumulatif. Normalisasi data dilakukan dengan cara mengubah nilai atribut ke dalam rentang nilai yang lebih kecil dan telah ditentukan.

C. Clustering

Clustering dengan melakukan analisis pengelompokan merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang memiliki kesamaan data dengan data kelompok lain. Kelas pada *clustering* disebut sebagai *cluster*. *Cluster* adalah suatu koleksi data yang memiliki kemiripan pada satu *cluster* dan memiliki ketidakmiripan dengan *cluster* yang lain [7].

Clustering adalah salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*) dan bertujuan untuk mencari karakteristik pola data secara umum pada database. Pendekatan yang dapat digunakan untuk mengembangkan *clustering* adalah *partitioning*. Metode *partitioning* akan mengalokasikan data ke dalam sebuah cluster sehingga tiap cluster memiliki minimal satu data dan tiap data termasuk pada salah satu grup [6].

D. Fuzzy C-Means

Fuzzy C-Means adalah suatu metode pengklasteran data yang ditentukan oleh derajat keanggotaan. Konsep dasar *Fuzzy C-Means*, pertama kali adalah menentukan pusat *cluster*, yang akan menandai lokasi rata – rata untuk tiap *cluster* [8]. Pada kondisi awal, pusat *cluster* masih belum akurat. Tiap – tiap titik data memiliki derajat keanggotaan untuk tiap *cluster*. Berikut adalah algoritma *Fuzzy C-Means* [9]:

1. Input data yang akan di-*cluster* X , berupa matriks ukuran $n \times m$. Dimana n merupakan jumlah sampel data, m merupakan atribut setiap data, X_{ij} merupakan data sampel ke- i untuk $i=1,2,\dots,n$ dan atribut ke- j untuk $j=1,2,\dots,m$.
2. Tentukan :
 - a. jumlah *cluster* = c
 - b. pangkat = w
 - c. maksimum iterasi = $MaxIter$
 - d. error terkecil yang diharapkan = **Error!**
 - e. fungsi obyektif, $P_0 = 0$
 - f. iterasi awal, $t = 1$
3. Bangkitkan bilangan acak (**Error! Reference source not found.**) dengan persamaan (1). Dimana $i = 1,2,\dots,n$ dan $k = 1,2,\dots,c$. Elemen – elemen dari matriks partisi awal U dengan menghitung jumlah setiap kolom.
Error! Reference source not found.
(1)
dengan $j=1,2,\dots,n$ dihitung dengan persamaan (2).

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (2)$$

4. Hitung pusat cluster V_{kj} ke k dengan persamaan (3). Error! Reference source not found.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (3)$$

5. Hitung fungsi objektif pada iterasi $ke-t$ dengan persamaan (4).

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (4)$$

6. Hitung perubahan matriks partisi dengan persamaan (5).

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \quad (5)$$

dimana $i = 1, 2, \dots, n$ dan $k = 1, 2, \dots, c$

7. Cek kondisi berhenti :
- Jika **Error! Reference source not found.** atau $(t > \text{MaxIter})$ maka berhenti.
 - Jika tidak : $t = t+1$, ulangi langkah ke-4.

E. Validitas Clustering

Untuk menguji validitas hasil clustering dengan pendekatan *Fuzzy C-Means*, dalam penelitian ini digunakan dua metode evaluasi yaitu *Silhouette Coefficient* dan *Purity*.

Silhouette Coefficient (SC) digunakan untuk melihat kualitas dan kekuatan cluster, seberapa baik suatu objek ditempatkan dalam suatu cluster. Semakin nilai SC mendekati nilai 1 maka cluster tersebut semakin baik. Langkah menghitung SC adalah sebagai berikut:

1. Hitung a_i dimana nilai a_i berdasarkan persamaan (6) untuk mengukur seberapa tidak mirip sebuah data dengan cluster yang diikutinya, nilai yang semakin kecil menandakan semakin tepatnya data tersebut berada dalam cluster tersebut:

$$a_i^j = \frac{1}{m_j - 1} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_i^j, x_r^j), \quad i = 1, 2, \dots, m_j \quad (6)$$

Dimana d adalah jarak data $ke-i$ dengan data $ke-r$ dalam satu cluster j , sedangkan m_j adalah jumlah data dalam cluster $ke-j$.

2. Hitung nilai cluster b_i dengan persamaan (7), dimana nilai b_i yang besar menandakan seberapa jeleknya data terhadap cluster yang lain:

$$b_i^j = \min \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ r \neq i}}^{m_n} d(x_i^j, x_r^n) \right\}, \quad i = 1, 2, \dots, m_n \quad (7)$$

3. Untuk mendapatkan nilai SC data $ke-i$ menggunakan persamaan (8), dimana nilai SI yang didapat dalam rentang $[-1, +1]$.

$$SI_i^j = \frac{a_i^j - b_i^j}{\max\{a_i^j, b_i^j\}} \quad (8)$$

4. Untuk nilai SI dari sebuah cluster didapatkan dengan menghitung rata-rata nilai SI semua data yang bergabung dalam cluster tersebut, seperti pada persamaan (9).

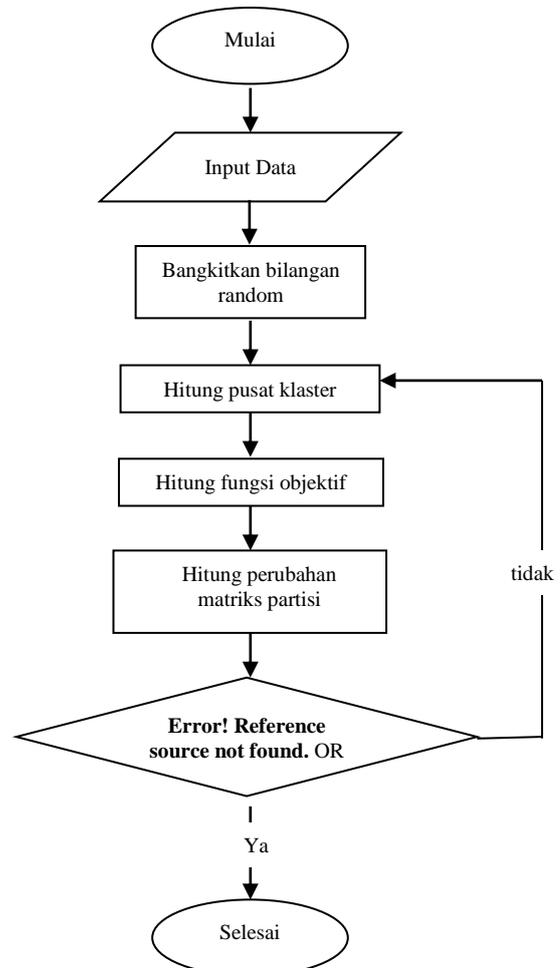
$$SI_j = \frac{1}{m_j} \sum_{i=1}^{m_j} SI_i^j \quad (9)$$

Purity (kemurnian) digunakan untuk suatu cluster yang direpresentasikan sebagai anggota cluster yang paling banyak sesuai (cocok) disuatu cluster [8]. *Purity* dapat dihitung dengan persamaan (10).

$$Purity(j) = \frac{1}{n_j} \max(n_{ij}) \quad (10)$$

Total nilai *Purity* dapat dihitung dengan persamaan (11).

$$Purity = \sum_{i=0}^j \frac{n_j}{n} Purity(j) \quad (11)$$



Gambar 1. Flowchart *Fuzzy C-Means*

III. METODOLOGI PENELITIAN

A. Metode Pengumpulan Data

Metode pengumpulan data pada penelitian dimaksudkan untuk mendapatkan data yang relevan, akurat, dan realistis. Metode yang digunakan dalam penelitian ini adalah metode studi pustaka yang diperoleh dari instansi terkait dan buku referensi. Sumber data yang digunakan dalam penelitian ini adalah data Jumlah Fasilitas Pelayanan Kesehatan tiap Kab / Kota di Jawa Timur yang diperoleh dari situs <http://www.jatim.bps.go.id>. Variabel yang digunakan adalah Rumah Sakit, Puskesmas, Posyandu, Polindes dan Klinik Keluarga Berencana (KKB).

B. Pengolahan data

Data yang telah didapat selanjutnya diolah terlebih dahulu. Selanjutnya data akan *di clustering*. Dalam menentukan *cluster* berdasarkan data yang telah ada, dibutuhkan sebuah flowchat untuk memudahkan dalam menentukan alur perhitungan. Juga untuk menemukan hasil dari penerapan *cluster* terhadap data yang akan diproses. Data dari Jumlah Fasilitas Pelayanan Kesehatan tiap Kabupaten/kotadi Jawa Timur adak *di cluster* ke dalam 3 *cluster* yaitu *cluster* tingkat fasilitas kesehatan tinggi, *cluster* tingkat fasilitas kesehatan sedang dan *cluster* tingkat fasilitas kesehatan rendah. Setelah didapatkan hasil *cluster*, akan dianalisis hasil tersebut menggunakan *Silhouette Index* (intercluster) dan *Purity* (intra-cluster). Gambar 1 menunjukkan *flowchart* dalam menentukan *cluster* dengan *Fuzzy C-Means*.

TABEL I

DATA JUMLAH FASILITAS KABUPATEN/KOTA DI JAWA TIMUR

No	Kab/Kota	Rumah Sakit	Puskesmas	Posyandu	Polindes	Klinik Keluarga Berencana
1	Pacitan	3	24	823	118	244
2	Ponorogo	6	31	1131	174	284
3	Trenggalek	4	22	857	104	440
4	Tulungagung	12	31	1239	167	485
5	Blitar	10	24	1469	117	408
6	Kediri	11	37	1727	321	482
7	Malang	24	39	2838	390	248
8	Lumajang	6	25	1292	108	205
9	Jember	12	49	2876	73	731
10	Banyuwangi	13	45	2286	70	788
11	Bondowoso	4	25	1059	0	386
12	Situbondo	5	17	924	96	282
13	Probolinggo	6	33	1312	70	273
14	Pasuruan	6	33	1893	189	519
15	Sidoarjo	26	26	1783	126	716
16	Mojokerto	11	27	1275	27	391
17	Jombang	13	34	1563	216	473
18	Nganjuk	6	20	1319	202	725
19	Madiun	4	26	878	73	259
20	Magetan	6	22	929	40	262
21	Ngawi	3	24	1202	169	485
22	Bojonegoro	10	36	1614	358	580
23	Tuban	4	33	1432	280	306
24	Lamongan	11	33	1742	252	1231
25	Gresik	14	32	1507	48	373
26	Bangkalan	4	22	1091	216	418
27	Sampang	1	21	1020	220	80
28	Pamekasan	5	20	939	162	220
29	Sumenep	3	30	1476	313	270
30	Kota Kediri	12	9	336	0	131
31	Kota Blitar	5	3	165	0	22

No	Kab/Kota	Rumah Sakit	Puskesmas	Posyandu	Polindes	Klinik Keluarga Berencana
32	Kota Malang	24	15	647	0	148
33	Kota Probolinggo	4	6	218	0	71
34	Kota Pasuruan	1	8	286	0	53
35	Kota Mojokerto	7	5	167	0	72
36	Kota Madiun	8	6	270	0	49
37	Kota Surabaya	62	62	2824	0	226
38	Kota Batu	5	5	189	12	56

Centroid	Rumah Sakit	Puskesmas	Posyandu	Polindes	KKB
C1	26,324	46,941	2612,121	142,769	520,813
C2	7,517	27,607	1300,000	159,546	398,783
C3	7,329	8,555	342,326	13,465	96,399

TABEL II
CENTROID ITERASI AKHIR

TABEL III
DERAJAT KEANGGOTAAN TIAP DATA PADA SETIAP CLUSTER

Data Ke	Derajat Keanggotaan data pada Cluster ke			Data Ke	Derajat Keanggotaan data pada Cluster ke		
	1	2	3		1	2	3
1	0,038	0,491	0,471	20	0,039	0,659	0,302
2	0,017	0,926	0,057	21	0,008	0,973	0,018
3	0,041	0,633	0,325	22	0,131	0,800	0,069
4	0,006	0,983	0,011	23	0,026	0,944	0,029
5	0,022	0,957	0,021	24	0,356	0,506	0,137
6	0,193	0,733	0,073	25	0,041	0,923	0,036
7	0,904	0,069	0,026	26	0,019	0,920	0,062
8	0,020	0,939	0,041	27	0,047	0,698	0,255
9	0,941	0,043	0,016	28	0,038	0,681	0,281
10	0,830	0,134	0,036	29	0,047	0,907	0,046
11	0,029	0,852	0,119	30	0,000	0,001	0,998
12	0,037	0,678	0,284	31	0,006	0,025	0,969
13	0,013	0,963	0,024	32	0,020	0,154	0,826
14	0,382	0,542	0,076	33	0,003	0,012	0,985

Data Ke	Derajat Keanggotaan data pada Cluster ke			Data Ke	Derajat Keanggotaan data pada Cluster ke		
	1	2	3		1	2	3
15	0,289	0,626	0,085	34	0,001	0,004	0,995
16	0,010	0,972	0,019	35	0,005	0,022	0,973
17	0,063	0,895	0,042	36	0,001	0,006	0,992
18	0,055	0,876	0,069	37	0,918	0,059	0,023
19	0,039	0,584	0,377	38	0,004	0,018	0,978

TABEL IV
HASIL CLUSTER MENGGUNAKAN FCM

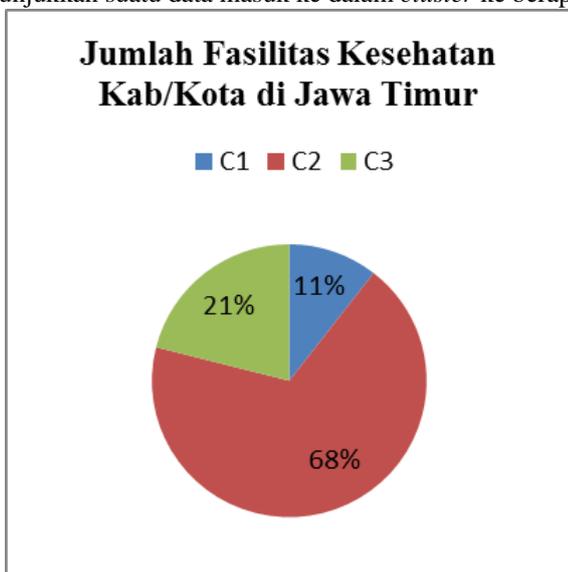
Data Ke	Hasil Cluster			Data Ke	Hasil Cluster		
	C1	C2	C3		C1	C2	C3
1		I		21		I	
2		I		22		I	
3		I		23		I	
4		I		24		I	
5		I		25		I	
6		I		26		I	
7	I			27		I	
8		I		28		I	
9	I			29		I	
10	I			30			I
11		I		31			I
12		I		32			I
13		I		33			I
14		I		34			I
15		I		35			I
16		I		36			I
17		I		37	I		
18		I		38			I
19		I					
20		I					

IV. HASIL DAN PEMBAHASAN

Data jumlah fasilitas pelayanan kesehatan setiap kabupaten/kota di Jawa Timur ditunjukkan oleh Tabel I, Setelah pengumpulan data, dilakukan *clustering* menggunakan *Fuzzy C-Means* dengan batasan jumlah *cluster* sebanyak 3 *cluster*, sehingga proses berhenti pada iterasi ke 48 dan diperoleh pusat *cluster/centroid* pada iterasi terakhir seperti pada Tabel II.

Dapat dilihat pada Tabel II bahwa *cluster* dapat dibagi menjadi tiga bagian yaitu *cluster* tingkat jumlah fasilitas kesehatan tinggi (C1), *cluster* tingkat jumlah fasilitas kesehatan sedang (C2), dan *cluster* tingkat jumlah fasilitas kesehatan rendah (C3). Selain didapatkan pusat *cluster* atau *centroid*, juga didapatkan derajat keanggotaan tiap data pada setiap *cluster* dengan FCM seperti pada Tabel III.

Tiap data memiliki derajat keanggotaan tertentu untuk menjadi anggota suatu *cluster*. Dengan derajat keanggotaan terbesar menunjukkan kecenderungan tertinggi tiap data untuk menjadi anggota suatu *cluster*. Tabel IV menunjukkan suatu data masuk ke dalam *cluster* ke berapa.



Gambar 2. Presentasi Jumlah Fasilitas Kesehatan Kabupaten/kotadi Jawa Timur

Dari hasil cluster pada Tabel IV dan Gambar 2 menunjukkan bahwa terdapat 4 kabupaten/kotayang termasuk dalam cluster tingkat pertama atau memiliki jumlah fasilitas kesehatan tinggi, 26 kabupaten/kotayang termasuk dalam cluster tingkat kedua atau memiliki jumlah fasilitas kesehatan sedang, dan 8 kabupaten/kotayang termasuk dalam cluster tingkat ketiga atau memiliki jumlah fasilitas kesehatan rendah.

Untuk menganalisis hasil cluster yang telah didapatkan peneliti menggunakan *Silhouette Coefficient* dan *Purity* sebagai berikut :

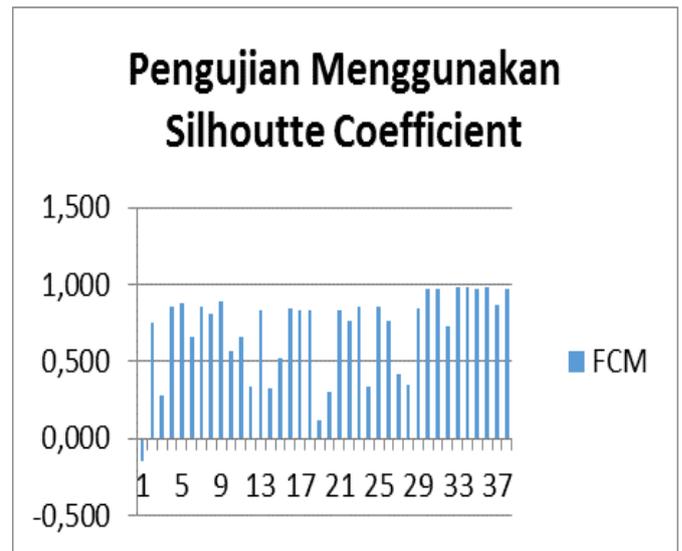
1. *Silhouette Coefficient*

Pengujian *Silhouette Coefficient* digunakan untuk melihat kualitas dan kekuatan *cluster*, seberapa baik

suatu objek ditempatkan dalam suatu *cluster*. Pengujian dilakukan pada hasil *clustering* menggunakan *Fuzzy C-Means*. Dalam pengujian hasil *cluster* ini menghasilkan nilai $s(i)$ seperti yang ditunjukkan oleh Tabel V.

TABEL V
 HASIL PERHITUNGAN NILAI SILHOUTTE COEFFICIENT

Data Ke	s(i)						
1	-	11	0,665	21	0,838	31	0,971
2	0,146	12	0,339	22	0,758	32	0,723
3	0,280	13	0,832	23	0,859	33	0,978
4	0,852	14	0,324	24	0,342	34	0,979
5	0,874	15	0,515	25	0,861	35	0,972
6	0,659	16	0,839	26	0,758	36	0,979
7	0,850	17	0,833	27	0,421	37	0,866
8	0,805	18	0,832	28	0,347	38	0,975
9	0,886	19	0,114	29	0,847		
10	0,572	20	0,300	30	0,971		



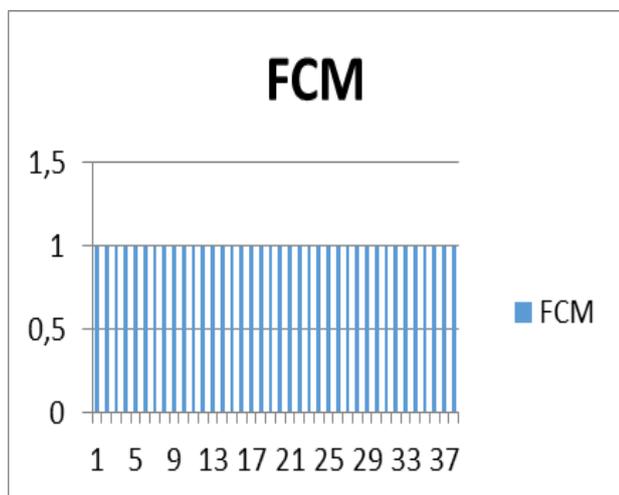
Gambar 2. Nilai Silhouette Coefficient untuk Hasil Clustering

Setelah didapat hasil $s(i)$ masing – masing data, akan diambil rata – rata untuk digunakan sebagai nilai *Silhouette Coefficient* dari hasil *clustering* menggunakan *Fuzzy C-Means*. Maka, nilai *Silhouette Coefficient* adalah 0,695. Karena nilai *Silhouette Coefficient* mendekati 1 maka hasil *cluster* sudah baik dalam kedekatan data di dalam satu *cluster*.

2. *Purity*

Pengujian *Purity* ini untuk melihat suatu *cluster* yang direpresentasikan sebagai anggota *cluster* yang paling cocok. Sebelum dilakukan pengujian *Purity* maka hasil *cluster* diberikan label terlebih dahulu. Untuk

pemberian label, data diproses pada matlab dan dirandom sebanyak 10 kali. Selanjutnya diambil hasil yang paling sering muncul dalam proses tersebut. sehingga didapatkan hasil *Purity* yang ditunjukkan oleh Gambar 4.



Gambar 4. Hasil perhitungan nilai *Purity* pada hasil cluster FCM

Dari hasil Gambar 4 menunjukkan bahwa hasil dari perhitungan *Purity* pada hasil *cluster* adalah 1. Yang berarti disetiap *cluster* selalu dihasilkan anggota kelompok yang selalu mirip dengan anggota yang lainnya.

V. KESIMPULAN

Untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan jumlah fasilitas pelayanan kesehatannya dapat digunakan metode *Fuzzy C-Means*. Yang menghasilkan 3 *cluster*, *cluster* pertama yaitu tingkat jumlah fasilitas kesehatan tinggi terdapat 4 kabupaten/kota yaitu Kab. Malang, Kab. Jember, Kab. Banyuwangi, dan Kota Surabaya. *Cluster* kedua yaitu tingkat jumlah fasilitas kesehatan sedang terdapat 26 kabupaten/kotadan *cluster* ketiga yaitu tingkat jumlah fasilitas kesehatan rendah terdapat 8 kabupaten/kota yaitu Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, dan Kota Batu. Untuk menganalisis hasil *cluster* digunakan *Silhouette Coefficient* dan *Purity*. Nilai *Silhouette Coefficient* pada hasil *cluster* adalah 0,695, sehingga *cluster* dikatakan sudah baik dalam kedekatan data di dalam satu *cluster*. Untuk nilai *Purity*

dari hasil *cluster* adalah 1, jadi di setiap *cluster* selalu dihasilkan anggota kelompok yang selalu mirip dengan anggota yang lainnya. Saran yang dapat digunakan agar mendapat tingkat akurasi yang lebih baik dalam penelitian ini serta untuk memberikan kontribusi yang lebih besar di dunia penelitian dengan membuat pemetaan berbasis web-GIS untuk mengetahui sebaran penyakit sebagai langkah mitigasi dalam penyediaan pelayanan kesehatan [10], selanjutnya hasil penelitian ini dapat dikembangkan dengan cara memodifikasi algoritma *Fuzzy C-Means* yang digunakan saat ini, atau dengan menggabungkan algoritma *Fuzzy C-Means* dengan algoritma yang lainnya.

REFERENSI

- [1] Undang Undang Republik Indonesia No. 36, 2009.
- [2] A. Ramadhan, Z. Efendi and Mustakim, "Perbandingan K-Means dan Fuzzy C-Means untuk Pengelompokan Data User Knowledge Modeling," *Seminar Nasional Teknologi Informasi, Komunikasi, dan Industri (SNTKI)*, 2017.
- [3] T. A. Risqiyani and A. Kesumawati, "Pengelompokan Kabupaten Kota di Provinsi Jawa Tengah dengan Fuzzy C-Means Clustering," *Seminar Nasional Matematika dan Pendidikan Matematika UNY*, 2016.
- [4] R. Handoyo, R. R. M and S. M. Nasution, "Perbandingan Metode Clustering Menggunakan Metode Single Linkage dan K-Means Pada Pengelompokan Dokumen," *JSM STMIK Mikroskil*, p. 77, 2014.
- [5] Suhartoyo, H., & Wijaya, T. A. (2016). Rancangan Estimasi Biaya dengan Teknik COCOMO II dan Neuro Fuzzy Studi Kasus : Sistem Informasi Rumah Sakit. *Inform*, 1(1). Retrieved from <http://ejournal.unitomo.ac.id/index.php/inform/article/view/215/118>
- [6] E. D. Cahyati, D. Herawatie and E. Wuryanto, "Implementasi K-Means Clustering Untuk Pemetaan Desa dan Kelurahan di Kabupaten Bangkalan Berdasarkan Contraceptive Prevalance Rate dan Tingkat Pendidikan," *Seminar Nasional Matematika dan Aplikasinya, Surabaya Universitas Airlangga*, 2017.
- [7] F. Febrianti, M. Hafiyusholeh and A. H. Asyhar, "Perbandingan Pengklusteran Data Iris Menggunakan Metode K-Means dan Fuzzy C-Means," *Jurnal Matematika "MANTIK"*, 2016.
- [8] S. Kusumadewi and H. Purnomo, *Aplikasi Logika Fuzzy untuk Pendukung Keputusan Edisi 2*, Yogyakarta: Graha Ilmu, 2013.
- [9] M. N. Sutoyo and A. T. Sumpala, "Penerapan Fuzzy C-Means untuk Deteksi Dini Kemampuan Penalaran Matematis," *Scientific Journal of Informatics*, 2015.
- [10] Vitianingsih, A. V., Cahyono, D., & Choiron, A. (2017). Analysis and design of web-geographic information system for tropical diseases-prone areas: A case study of East Java Province, Indonesia. In *2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)* (pp. 255–260). IEEE Xplore Digital Library. <https://doi.org/10.1109/ICITACEE.2017.8257713>