

# A CLUSTERING OF GENERATIVE AND INFECTIOUS DISEASES USING FUZZY C-MEANS

**Novianti Puspitasari**

Teknik Informatika, FKTI  
Universitas Mulawarman, Indonesia  
Jl. Barong Tongkok Kampus Gn. Kelua,  
Samarinda, 75119, Indonesia  
novia.ftik.unmul@gmail.com

**Joan Angelina Widians**

Teknik Informatika, FKTI  
Universitas Mulawarman, Indonesia  
Jl. Barong Tongkok Kampus Gn. Kelua,  
Samarinda, 75119, Indonesia  
angel\_widians@yahoo.com

**Pohny**

Teknik Informatika, FKTI  
Universitas Mulawarman, Indonesia  
Jl. Barong Tongkok Kampus Gn. Kelua,  
Samarinda, 75119, Indonesia  
pohny28@gmail.com

**Abstrack** - Processing information of the most disease suffered by the people in a region, particularly to those receive Health Insurance card (Jaminan Kesehatan Daerah; JAMKESDA) was one of the government focuses. The study used Fuzzy C-Means (FCM), to process the disease data suffered by JAMKESDA users for four years in regional public hospital. The results showed that FCM method could group the disease into two forms namely generative and infectious one. Based on the cluster validity test using Partition Coefficient (PC), the partition value of fuzzy is higher by using two clusters (0.68) rather than three clusters. In conclusion, the forming model of two clusters is more optimal than that three one to process data disease of JAMKESDA users.

## Keywords

Fuzzy C-Means, Diseases, Degenerative, Infection, Clustering.

**Abstrak** - Pengolahan informasi mengenai penyakit yang paling banyak diderita oleh masyarakat disuatu daerah, khususnya bagi masyarakat yang menerima pelayanan berupa kartu Jaminan Kesehatan (JAMKESDA) dari pemerintah merupakan salah satu hal yang menjadi fokus perhatian pemerintah. Penelitian ini menggunakan metode Fuzzy C-Means (FCM), untuk mengolah data penyakit yang diderita oleh masyarakat pengguna JAMKESDA dalam periode 4 tahun di rumah sakit umum daerah. Adapun hasil penelitian menunjukkan bahwa metode FCM mampu mengelompokkan penyakit ke dalam dua jenis kelompok penyakit yaitu degeneratif dan infeksi. Dimana, berdasarkan hasil uji validitas cluster dengan menggunakan metode Partition Coefficient (PC), menghasilkan nilai partisi fuzzy lebih besar menggunakan dua cluster (sebesar 0.68) dibandingkan tiga cluster, sehingga model pembentukan cluster dengan dua cluster lebih optimal dibanding dengan tiga cluster untuk pengolahan data penyakit pada masyarakat pengguna JAMKESDA.

## Kata Kunci

Fuzzy C-Means, Penyakit, Degeneratif, Infeksi, Clustering.

## 1. PENDAHULUAN

Bidang kesehatan saat ini tengah menjadi fokus perhatian yang sangat besar bagi pemerintah. Pemerintah saat ini berupaya untuk selalu meningkatkan taraf kesehatan masyarakat dengan berupaya meningkatkan pelayanan kepada masyarakat. Salah satu bentuk pelayanan kesehatan yang dilakukan yaitu pemberian kartu Jaminan Kesehatan Masyarakat Daerah (JAMKESDA) kepada masyarakat.

RSUD AW Syahrani Samarinda merupakan salah satu rumah sakit pemerintah yang berada di kota Samarinda. Dimana, RSUD

AW Syahrani merupakan rumah sakit rujukan bagi masyarakat dan merupakan salah satu rumah sakit yang menerima pasien pengguna JAMKESDA. Setiap tahunnya, RSUD AW Syahrani menerima banyak pasien baik dari kota Samarinda dan sekitarnya maupun dari berbagai kecamatan dan kabupaten yang ada disekitar kota Samarinda. Jumlah pasien pengguna JAMKESDA yang terus mengalami peningkatan membuat pihak RSUD AW Syahrani mengalami kesulitan dalam hal pengambilan informasi untuk mengetahui penyakit yang paling banyak diderita oleh pasien, dikarenakan jumlah data pasien yang begitu besar. Dimana, hal ini akan berdampak pada upaya pemerintah untuk menanggulangi pencegahan penyakit yang paling banyak diderita oleh masyarakat khususnya masyarakat pengguna JAMKESDA. Oleh karena itu pengolahan informasi mengenai penyakit yang paling banyak diderita oleh pasien pengguna JAMKESDA sangat diperlukan dalam rangka membantu pemerintah meningkatkan taraf kesehatan masyarakat.

Saat ini, pengelompokan data terutama data penyakit dengan memanfaatkan metode kecerdasan buatan (*artificial intelligence*) telah banyak dilakukan, diantaranya menggunakan metode *Naïve Bayes Classifier* [1] untuk mengelompokkan penyakit mata yang paling banyak diderita masyarakat Jawa Barat. Selanjutnya terdapat metode K-Means yang paling sering digunakan untuk mengelompokkan penyakit, diantaranya penyakit karies gigi [2], dan penyakit menular pada manusia [3]. Dari hasil penelitian [3] menunjukkan bahwa metode K-Means mampu mengelompokkan penyakit menular ke dalam 6 (enam) cluster. Namun, metode K-Means memiliki kekurangan yaitu sangat sensitif pada pembangkitan titik pusat awal secara random. Dimana, hasil pengelompokan bersifat tidak unik (selalu berubah-ubah) dan proses pengerjaannya cepat, tetapi keakuratannya tidak dijamin.

Metode *fuzzy clustering*, telah banyak diaplikasikan untuk mengelompokkan suatu data berdasarkan kesamaan atau kemiripan yang dimilikinya. Salah satu metode *fuzzy clustering* yang biasa diterapkan adalah *Fuzzy Clustering Means* (FCM). Dimana, untuk meminimalisasikan fungsi obyektif yang diset dalam proses *clustering* yang ada pada umumnya, FCM berusaha meminimalisasikan variasi didalam suatu cluster dan memaksimalkan variasi antar cluster [4]. Metode FCM telah banyak digunakan sebagai alat untuk memberikan rekomendasi keputusan berbagai keperluan. Diantaranya, dalam bidang geologi dimana FCM diterapkan untuk mengetahui kondisi cuaca dan pergeseran musim yang terjadi [5]. Dalam bidang pengolahan citra digital dan image, FCM dapat digunakan untuk pendeteksian area tumor otak [6, 7], segmentasi kepala janin pada citra USG

[8], dan lain-lain [9-11]. Selanjutnya pada bidang transportasi, FCM digunakan untuk mengelompokkan lalu lintas yang padat di negara perkotaan [12], serta menentukan prioritas suatu perbaikan jalan [13]. Pada bidang pendidikan, FCM digunakan untuk pengelompokan konsentrasi jurusan mahasiswa dengan melihat batas bobot nilai akademis mata kuliah tertentu sesuai dengan hasil pengelompokan yang dilakukan menggunakan metode FCM tanpa mengesampingkan minat personal mahasiswa [14].

Selanjutnya, metode FCM banyak digunakan untuk pembuatan aplikasi yang mendukung pengambilan keputusan, seperti aplikasi penentuan predikat kelulusan mahasiswa berdasarkan data predikat kelulusan seperti jumlah sks, jumlah IPK dan keterangan lulus [15]. Sistem penentuan untuk penerimaan dan penyaluran zakat berdasarkan lima parameter yaitu indeks keluarga, indeks keluarga II, indeks barang, indeks data keluarga dan indeks keimanan [16]. Selain itu, terdapat sistem pendukung keputusan untuk penerimaan Bantuan Langsung Masyarakat (BLM) PNPM-MPd berdasarkan kriteria kelayakan yang digunakan untuk penentuan prioritas usulan yang akan menghasilkan daftar rangking usulan. Dimana, output sistem berupa perangkungan usulan kegiatan desa serta kategori usulan desa, yaitu layak atau tidaknya usulan tersebut dalam menerima bantuan dari PNPM MPd [17].

Metode FCM juga dapat dikombinasikan dengan menggunakan metode kecerdasan buatan yang lainnya. Salah satunya adalah kombinasi antara metode FCM dengan *Radial Basis Function* (RBF). RBF merupakan salah satu jaringan syaraf tiruan yang handal untuk permasalahan regresi dan klasifikasi. Penelitian dengan menggunakan metode FCM dan RBF pada klasifikasi penyakit karies gigi menunjukkan bahwa pada FCM proses pemilihan pusat *cluster* dapat langsung ditemukan sehingga proses penentuan pusat *cluster* menjadi lebih sederhana [18]. Dari beberapa penelitian yang ada, menunjukkan bahwa metode FCM memiliki kemampuan untuk melakukan pengelompokan data yang belum diklasifikasi, kedalam kelas tertentu menurut kesamaan yang dimilikinya, berdasarkan derajat keanggotaan dengan cara minimalisasi nilai fungsi obyektifnya [14]. Sehingga dapat memberikan rekomendasi untuk sebuah pengambilan keputusan. Oleh karena itu, penelitian ini menerapkan metode FCM untuk mengelompokkan data penyakit yang paling banyak diderita oleh masyarakat khususnya masyarakat pengguna JAMKESDA ke dalam dua kategori yaitu infeksi dan degeneratif. Lebih lanjut, hasil penelitian diharapkan mampu memberikan informasi kepada pemerintah mengenai penyakit yang paling banyak diderita oleh masyarakat, sehingga pemerintah dapat melakukan usaha pencegahan dalam bentuk pemeriksaan serta penyuluhan kepada masyarakat.

## 2. TINJAUAN PUSTAKA

Studi literatur ini membahas tentang penyakit degeneratif, penyakit infeksi dan *Fuzzy Clustering Means* (FCM).

### 2.1 Penyakit Degeneratif dan Infeksi

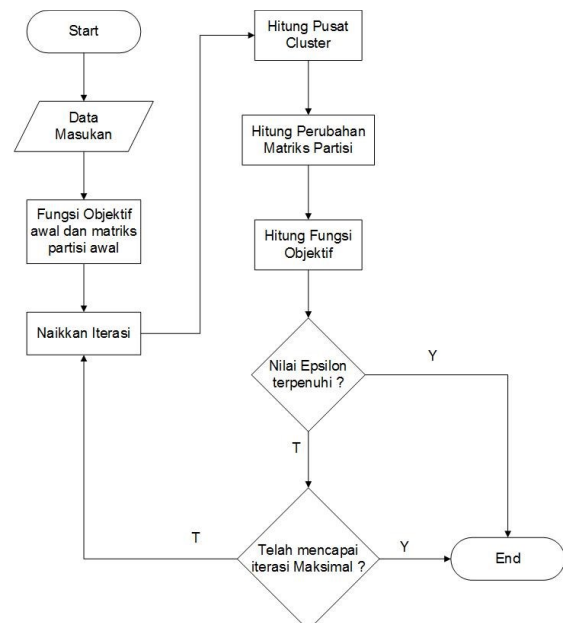
Penyakit infeksi merupakan suatu penyakit yang disebabkan karena adanya mikroba patogen [19]. Penyakit ini menular dari satu orang ke orang lain. Orang yang sehat harus dihindarkan dari orang-orang yang menderita penyakit dari golongan ini. Salah satu penyebab penyakit infeksi adalah bakteri [20]. Bakteri yang dapat menyebabkan terjadi infeksi contohnya *Escherichia coli* dan *Bacillus subtilis*. Sedangkan, penyakit degeneratif adalah penyakit yang menyebabkan terjadinya kerusakan atau

penghacuran terhadap jaringan atau organ tubuh. Proses dari kerusakan ini disebabkan oleh proses kemunduran fungsi sel tubuh yaitu dari keadaan normal menjadi lebih buruk seiring dengan bertambahnya usia maupun karena gaya hidup yang tidak sehat (Pusdatin Kemenkes RI, 2010).

Ada berbagai macam teori yang menjelaskan penyebab penyakit degeneratif, salah satunya adalah teori radikal bebas. Berdasarkan teori ini, penyebab dari penyakit degeneratif adalah adanya proses oksidasi radikal bebas dalam mekanisme biokimia yang terjadi dalam tubuh manusia [21].

### 2.2 Fuzzy C-Means (FCM)

*Fuzzy clustering* merupakan bagian dari *pattern recognition* atau pengenalan pola yang memainkan peran yang sangat penting dalam pencarian struktur dalam data untuk menentukan *cluster* optimal dalam suatu ruang vektor yang didasarkan pada bentuk normal Euclidian untuk jarak antar vektor. *Fuzzy Clustering Means* (FCM) adalah suatu teknik peng-*cluster-an* data. Dimana, keberadaan tiap-tiap data dalam suatu *cluster* ditentukan oleh derajat keanggotaan [22]. Konsep dasar FCM yang pertama kali adalah menentukan pusat *cluster*, yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum akurat. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan derajat keanggotaan tiap-tiap titik data secara berulang, maka akan dapat dilihat bahwa pusat *cluster* akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada minimasi fungsi obyektif yang menggambarkan jarak dari titik data yang diberikan ke pusat *cluster* yang terbobot oleh derajat keanggotaan titik data tersebut. Output dari FCM bukan merupakan *fuzzy inference system*, namun merupakan deretan pusat *cluster* dan beberapa derajat keanggotaan untuk tiap-tiap titik data. Informasi ini dapat digunakan untuk membangun suatu *fuzzy inference system* [17]. Berikut ini flowchart dari metode FCM seperti terlihat pada Gambar 1.



Gambar 1 Flowchart Fuzzy C-Means

Sedangkan, algoritma FCM yang digunakan adalah sebagai berikut [14]:

1. Input data yang akan di-cluster X, berupa matriks berukuran  $n \times m$  ( $n$  = jumlah sampel data,  $m$  = atribut data).  $X_{ij}$  = data sampel ke- $i$  ( $i = 1, 2, \dots, n$ ), atribut ke- $j$  ( $j = 1, 2, \dots, m$ )
2. Selanjutnya tentukan nilai-nilai awal perhitungan seperti:
 

Jumlah cluster	= $c$ ;
Pangkat	= $w$ ;
Maksimum Iterasi	= $MaksIter$ ;
Error terkecil yang diharapkan	= $\xi$
Fungsi obyektif awal	= $P_0 = 0$ ;
Iterasi awal	= $t = 1$ ;
3. Bangkitkan bilangan random dalam bentuk elemen matriks partisi awal  $U$   $\mu_{ik} = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, c$ ; kemudian hitunglah jumlah setiap kolom dengan persamaan (1).

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

Dimana :

$\mu_{ik}$  : derajat keanggotaan

$Q_i$  : Jumlah nilai derajat keanggotaan perkolom = 1 dengan  $i = 1, 2, \dots, n$ ;

Selanjutnya tentukan nilai matriks partisi awal, dengan persamaan (2).

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

4. Hitung pusat cluster ke- $k$  :  $V_{kj}$ , dimana  $k = 1, 2, \dots, c$ ; dan  $j = 1, 2, \dots, m$ ; dengan persamaan (3)

$$V_{kj} = \frac{\sum_{i=0}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=0}^n (\mu_{ik})^w} \quad (3)$$

Dimana :

$V$  : pusat cluster

$X_i$  : parameter ke- $i$

5. Selanjutnya, lakukan perhitungan nilai fungsi obyektif pada iterasi ke- $t$  ( $P_t$ ), dengan persamaan (4).

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

Dimana  $P_t$ : nilai fungsi obyektif iterasi ke- $t$

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

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

Dimana  $i = 1, 2, \dots, n$ ;  $k = 1, 2, \dots, c$ ;

7. Langkah terakhir adalah dengan mengecek kondisi berhenti, dengan ketentuan sebagai berikut:

1. Jika :  $(|P_t - P_{t-1}| < \xi)$  atau  $t > MaksIter$  maka berhenti;
2. Jika tidak :  $t = t + 1$ , ulangi langkah ke-4

### 2.3 Uji Validitas Clustering

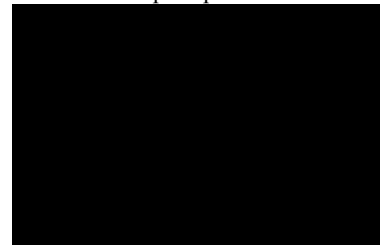
Di dalam FCM, pengelompokan data dilakukan dengan cara mengelompokkan tiap-tiap data ke dalam beberapa cluster berdasarkan derajat keanggotaan [22]. Selanjutnya, hasil clustering yang telah terbentuk akan diuji validitasnya menggunakan perhitungan koefisien partisi atau *Partition Coefficient* (PC). *Partition Coefficient* (PC) digunakan sebagai evaluasi data pada setiap cluster, dimana nilai *Partition Coefficient* (PC) hanya mengevaluasi nilai derajat keanggotaan tanpa memandang nilai data yang mengandung informasi [23]. Nilai hasil uji validitas PC berada dalam rentang [0,1]. Jika nilai PC semakin besar (mendekati 1) maka kualitas cluster yang didapat semakin baik. Berikut adalah formula untuk menghitung *Partition Coefficient* (PC).

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

Dimana,  $n$  adalah jumlah data;  $c$  adalah jumlah cluster, dan  $\mu_{ik}$  menyatakan nilai keanggotaan dari data ke- $k$  pada cluster ke- $i$ .

## 3. METODE PENELITIAN

Penelitian ini menggunakan *datasets* pengujian berupa data penyakit pasien pengguna JAMKESDA di RSUD AW Syahrani Kota Samarinda dari tahun 2012-2016 sebanyak 1000 data. Sebelum dataset tersebut diujikan, terlebih dahulu dilakukan proses normalisasi data menggunakan fungsi keanggotaan linier menaik pada *logika fuzzy*. Adapun fungsi keanggotaan untuk representasi linier menaik seperti pada Gambar 2 berikut.



Gambar 2. Fungsi keanggotaan representasi linier menaik.

Selanjutnya, untuk menghitung fungsi keanggotaan representasi linier menaik menggunakan persamaan (7).

$$\mu[x] = \begin{cases} 0; & \rightarrow x \leq a \\ \frac{x-a}{b-a}; & \rightarrow a \leq x \leq b \\ 1; & \rightarrow x \geq b \end{cases} \quad (7)$$

Dimana,  $a$  = nilai domain yang mempunyai derajat keanggotaan nol,  $b$  = nilai domain yang mempunyai derajat keanggotaan satu sedangkan  $x$  = nilai input yang akan di ubah ke dalam bilangan *fuzzy*.

Sebagai contoh, pada data umur dilakukan normalisasi ke dalam bilangan *fuzzy* dengan menggunakan fungsi keanggotaan representasi linier naik dengan nilai  $a = 20$  tahun dan  $b = 90$  tahun. Selanjutnya, apabila umur pasien adalah 45 tahun maka umur tersebut berada diantara  $a \leq x \leq b$  dengan bilangan *fuzzy*-nya adalah 0,357 dengan menggunakan persamaan (7) maka perhitungannya adalah

$$\mu_{\text{umur}} [45] = \frac{45-20}{90-20} = 0,357$$

Lebih lanjut, data kode penyakit dinormalisasi dengan menggunakan fungsi keanggotaan representasi linier naik dengan nilai  $a = 1$  dan  $b = 12$ . Sementara untuk data lama idap menggunakan nilai  $a = 0,5$  bulan dan  $b = 5$  bulan.

Data hasil normalisasi dari data penyakit pasien pengguna JAMKESDA sebanyak 1000 data dapat dilihat pada Tabel 1.

**Tabel 1. Data Penyakit**

NO	Jenis Kelamin	Kode Penyakit	Umur	Lama Idap
1	Perempuan	0,364	0,357	0,111
2	Perempuan	0,909	0,257	0,067
3	Perempuan	0,273	0,000	0,778
4	Perempuan	1,000	0,186	0,333
5	Perempuan	0,818	0,514	0,067
6	Laki-laki	0,273	0,000	0,556
7	Laki-laki	0,273	0,671	1,000
8	Laki-laki	0,273	0,000	0,556
9	Laki-laki	0,727	0,300	1,000
10	Laki-laki	0,636	0,371	1,000
11	Perempuan	0,364	0,271	0,111
12	Laki-laki	0,273	0,000	0,000
13	Perempuan	0,273	0,286	0,000
14	Perempuan	0,727	0,571	0,000
15	Laki-laki	0,273	0,000	1,000
16	Perempuan	0,273	0,014	1,000
17	Laki-laki	0,273	0,314	0,111
18	Laki-laki	0,727	0,629	0,111
19	Perempuan	0,727	0,429	0,000
20	Perempuan	0,727	0,214	0,556
21	Perempuan	0,182	0,000	0,089
22	Laki-laki	0,182	0,671	0,333
23	Laki-laki	0,182	0,600	0,111
24	Perempuan	0,273	0,000	1,000
25	Perempuan	0,182	0,229	0,556
26	Laki-laki	0,727	0,729	0,333
27	Laki-laki	0,273	0,000	1,000
28	Laki-laki	0,727	0,529	1,000
29	Perempuan	1,000	0,329	0,556
30	Laki-laki	0,727	0,629	0,022
31	Laki-laki	0,273	0,386	0,111
32	Perempuan	0,182	0,343	0,089
33	Laki-laki	0,727	0,771	0,000
34	Laki-laki	0,727	0,386	0,333
35	Laki-laki	0,818	0,600	0,556
36	Perempuan	0,273	0,000	1,000
37	Perempuan	0,727	0,600	0,000
38	Laki-laki	0,364	0,557	0,044
39	Laki-laki	0,545	0,100	0,000
40	Laki-laki	0,727	0,629	0,000
41	Laki-laki	0,364	0,414	1,000
42	Laki-laki	0,727	0,529	0,778
43	Perempuan	0,727	0,271	0,556
44	Laki-laki	0,727	0,714	1,000
45	Laki-laki	0,727	0,443	1,000
46	Perempuan	0,000	0,000	0,000
47	Perempuan	0,364	0,371	0,111
48	Perempuan	0,182	0,586	0,111

NO	Jenis Kelamin	Kode Penyakit	Umur	Lama Idap
49	Perempuan	0,727	0,600	1,000
50	Laki-laki	0,727	0,514	0,111
:	:	:	:	:
:	:	:	:	:
1000	Perempuan	0,909	0,571	0,044

Dari Tabel 1 diperoleh informasi bahwa terdapat tiga parameter yang menjadi dasar dalam pengelompokan penyakit yaitu kode penyakit, umur, dan lama mengidap penyakit. Adapun, parameter umur menjelaskan usia pasien yang mengidap suatu penyakit. Sedangkan, parameter kode penyakit menjelaskan tentang hasil diagnosis penyakit yang diterjemahkan ke dalam kode internasional kesehatan yaitu ICD10. Dimana, kode tersebut merupakan acuan bagi tenaga kesehatan dalam mendiagnosa penyakit. Dan, parameter lama mengidap penyakit menjelaskan mengenai rentang waktu pasien mengidap suatu penyakit dalam hitungan bulan. Di dalam penelitian ini, pengolahan data penyakit menggunakan bantuan program Matlab.

#### 4. HASIL DAN PEMBAHASAN

Pada sub bagian ini, akan diuraikan hasil pengujian metode FCM yang diujikan pada data hasil diagnosa penyakit pada pasien pengguna JAMKESDA.

##### 4.1 Hasil Analisa FCM

Langkah pertama yang dilakukan yaitu menetapkan jumlah *cluster*. Dalam penelitian ini data penyakit akan dikelompokkan ke dalam 2 (dua) *cluster* dengan maksimum iterasi sebanyak 100 kali. Lebih lanjut, dalam proses *clustering*, parameter kode penyakit dijadikan sebagai  $X_{i1}$ , parameter umur dijadikan sebagai  $X_{i2}$  dan parameter lama mengidap dijadikan sebagai  $X_{i3}$ . Dari *cluster*-ing yang dilakukan akan diperoleh hasil berupa nilai fungsi obyektif selama iterasi, pusat *cluster* serta derajat keanggotaan penyakit untuk setiap *cluster* pada iterasi terakhir. Di dalam penelitian ini, proses iterasinya berhenti pada iterasi ke-35 karena nilai fungsi obyektif ( $|P_t - P_{t-1}| < \xi$ ), sehingga nilai fungsi obyektif yang diperoleh pada iterasi terakhir (iterasi ke-35) adalah 127,458104. Hasil *cluster*-ing yang dilakukan adalah sebagai berikut.

1. Nilai fungsi obyektif selama 35 iterasi terlihat pada Tabel 2.

**Tabel 2. Fungsi Obyektif selama 35 iterasi**

Iterasi Ke-	Fungsi Obyektif
1	172,932812
2	142,339417
3	142,321687
4	142,272343
5	142,131887
:	:
:	:
:	:
35	127,458104

2. Derajat keanggotaan penyakit untuk setiap *cluster* pada iterasi terakhir (iterasi ke-35) terlihat pada Tabel 3.

**Tabel 3. Derajat Keanggotaan iterasi terakhir**

Data Ke-	Derajat Keanggotaan	
	C1	C2
1	0,965346952	0,034653048
2	0,791650629	0,208349371

Data Ke-	Derajat Keanggotaan	
	C1	C2
3	0,196970611	0,803029389
4	0,620689421	0,379310579
5	0,834848641	0,165151359
6	0,349643764	0,650356236
7	0,166063151	0,833936849
8	0,349643764	0,650356236
9	0,151511797	0,848488203
10	0,110759056	0,889240944
11	0,946067489	0,053932511
12	0,78736884	0,21263116
13	0,898186614	0,101813386
14	0,864461429	0,135538571
15	0,174863384	0,825136616
16	0,16919478	0,83080522
17	0,903354867	0,096645133
18	0,829703133	0,170296867
19	0,898178488	0,101821512
20	0,377449836	0,622550164
21	0,736904413	0,263095587
22	0,614068946	0,385931054
23	0,793084862	0,206915138
24	0,174863384	0,825136616
25	0,294781437	0,705218563
26	0,654576383	0,345423617
27	0,174863384	0,825136616
28	0,172658873	0,827341127
29	0,460903131	0,539096869
30	0,842983744	0,157016256
31	0,908611849	0,091388151
32	0,84636106	0,15363894
33	0,785932194	0,214067806
34	0,745342001	0,254657999
35	0,436827817	0,563172183
36	0,174863384	0,825136616
37	0,854464265	0,145535735
38	0,909062852	0,090937148
39	0,877249914	0,122750086
40	0,843793689	0,156206311
41	0,071209598	0,928790402
42	0,197844116	0,802155884
43	0,371792476	0,628207524
44	0,228337328	0,771662672
45	0,156466152	0,843533848
46	0,696973665	0,303026335
47	0,965774772	0,034225228
48	0,798285106	0,201714894
49	0,191775712	0,808224288
50	0,876047786	0,123952214
:	:	:
:	:	:
:	:	:
1000	0,782571342	0,217428658

Berdasarkan data pada tabel 3 yang menunjukkan derajat keanggotaan setiap data penyakit ke dalam setiap *cluster* diperoleh informasi mengenai kecenderungan data untuk masuk ke *cluster* mana. Derajat keanggotaan terbesar menunjukkan

kecenderungan tertinggi sebuah penyakit untuk masuk menjadi anggota *cluster*. Sebagai contoh untuk data penyakit ke-6, dapat menjadi:

- Anggota *cluster* pertama (C1) dengan derajat keanggotaan sebesar 0,349643764.
- Anggota *cluster* kedua (C2) dengan derajat keanggotaan sebesar 0,650356236.

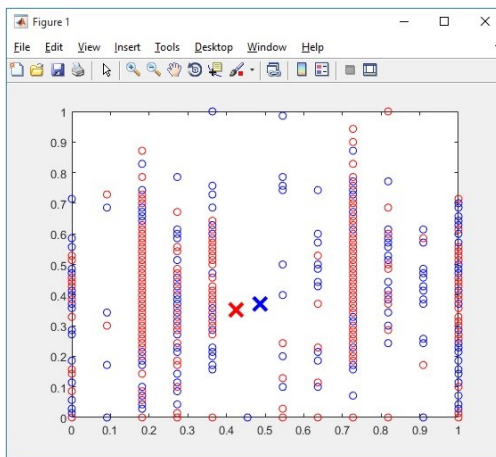
Berdasarkan data diatas, dapat dilihat bahwa nilai derajat keanggotaan terbesarnya terletak di *cluster* kedua, sehingga data penyakit ke-6 akan dimasukkan ke dalam *cluster* kedua (C2). Hasil selengkapnya dari pengelompokan data penyakit ke dalam 2 *cluster* dapat dilihat pada Tabel 4.

Tabel 4. Hasil *Clustering*

Data Ke-	Derajat Keanggotaan		Data Masuk Cluster
	C1	C2	
1	0,965346952	0,034653048	1
2	0,791650629	0,208349371	1
3	0,196970611	0,803029389	2
4	0,620689421	0,379310579	1
5	0,834848641	0,165151359	1
6	0,349643764	0,650356236	2
7	0,166063151	0,833936849	2
8	0,349643764	0,650356236	2
9	0,151511797	0,848488203	2
10	0,110759056	0,889240944	2
11	0,946067489	0,053932511	1
12	0,78736884	0,21263116	1
13	0,898186614	0,101813386	1
14	0,864461429	0,135538571	1
15	0,174863384	0,825136616	2
16	0,16919478	0,83080522	2
17	0,903354867	0,096645133	1
18	0,829703133	0,170296867	1
19	0,898178488	0,101821512	1
20	0,377449836	0,622550164	2
21	0,736904413	0,263095587	1
22	0,614068946	0,385931054	1
23	0,793084862	0,206915138	1
24	0,174863384	0,825136616	2
25	0,294781437	0,705218563	2
26	0,654576383	0,345423617	1
27	0,174863384	0,825136616	2
28	0,172658873	0,827341127	2
29	0,460903131	0,539096869	2
30	0,842983744	0,157016256	1
31	0,908611849	0,091388151	1
32	0,84636106	0,15363894	1
33	0,785932194	0,214067806	1
34	0,745342001	0,254657999	1
35	0,436827817	0,563172183	2
36	0,174863384	0,825136616	2
37	0,854464265	0,145535735	1
38	0,909062852	0,090937148	1
39	0,877249914	0,122750086	1
40	0,843793689	0,156206311	1
41	0,071209598	0,928790402	2
42	0,197844116	0,802155884	2
43	0,371792476	0,628207524	2

Data Ke-	Derajat Keanggotaan		Data Masuk Cluster
	C1	C2	
44	0,228337328	0,771662672	2
45	0,156466152	0,843533848	2
46	0,696973665	0,303026335	1
47	0,965774772	0,034225228	1
48	0,798285106	0,201714894	1
49	0,191775712	0,808224288	2
50	0,876047786	0,123952214	1
:	:	:	:
:	:	:	:
:	:	:	:
1000	0,782571342	0,217428658	1

Dari hasil percobaan menggunakan Matlab, perhitungan akhir dengan metode FCM terhadap seluruh data penyakit sebanyak 1000 data diperoleh *Cluster* 1 (C1) beranggotakan 663 data dan *Cluster* 2 (C2) beranggotakan 337 data. Plot hasil *cluster* dapat dilihat pada Gambar 3.



Gambar 3. Hasil Clustering menggunakan aplikasi Matlab

3. Nilai pusat *cluster* pada iterasi terakhir pusat *cluster*  $V_{kj}$  yang dihasilkan  $k = 1,2$  dengan  $j = 1,2,3$  ditampilkan oleh Tabel 5.

Tabel 5. Pusat *cluster* iterasi terakhir

$V_{kj}$	$V_{11}$	$V_{12}$	$V_{13}$
Cluster 1	0,486944888	0,370950018	0,1335154
Cluster 2	0,424911939	0,35162871	0,773743461

Berdasarkan data pada Tabel 5, dapat diperoleh nilai pusat *cluster*. Nilai pusat *cluster* merupakan nilai dari koordinat ke-2 titik pusat *cluster* dan memberikan informasi garis besar tiap *cluster* yaitu:

- a. *Cluster* 1 berisi kelompok penyakit dengan parameter kode penyakit 0,486944888; parameter umur sekitar 0,370950018; dan parameter lama idap penyakit 0,1335154.
- b. *Cluster* 2 berisi kelompok penyakit dengan parameter kode penyakit 0,424911939; parameter umur sekitar 0,35162871; dan parameter lama idap penyakit 0,773743461.

Dari hasil pusat *cluster* iterasi terakhir dapat diperoleh informasi bahwa *Cluster* 1 (C1) berisi kode penyakit 0,486944888 yang

setelah dilakukan normalisasi, merupakan kode ICD untuk penyakit *degeneratif* dengan rata-rata umur penderita berkisar antara 18 - 86 tahun (setelah dilakukan normalisasi terhadap parameter umur sebesar 0,370950018). Dimana, lama waktu pasien mengidap penyakit tersebut antara 1 - 4 bulan (berdasarkan hasil normalisasi lama idap sebesar 0,1335154). Sementara untuk *Cluster* 2 (C2) berisi kode penyakit 0,424911939 yang setelah dilakukan normalisasi merupakan kode penyakit infeksi dengan rata-rata umur penderita berkisar antara 1 - 86 tahun (setelah dilakukan normalisasi terhadap parameter umur sebesar 0,35162871) dengan lama waktu pasien mengidap penyakit tersebut lebih dari 4 bulan (berdasarkan hasil normalisasi lama idap sebesar 0,773743461).

Berdasarkan hasil perhitungan pusat kedua *cluster* dan hasil *clustering* pada Tabel 4 menunjukkan bahwa kelompok penyakit *degeneratif* paling banyak diderita oleh masyarakat pengguna JAMKESDA di kota Samarinda. Beberapa jenis penyakit yang termasuk kelompok penyakit degeneratif adalah hipertensi, jantung, diabetes, stroke, gagal ginjal, hepatitis, asam urat, rematik, *osteoporosis* dan kanker.

#### 4.2 Partition Coefficient (PC)

Tahap selanjutnya adalah pengujian validasi menggunakan metode *Partition Coefficient (PC)* yang memberikan informasi kevaliditan atas pembentukan *cluster* yang optimal. Dalam penelitian ini, metode PC digunakan untuk menguji hasil klastering yang dibuat sebanyak dua *cluster* dan hasil klastering sebanyak tiga *cluster*.

Tabel 6. Hasil *Partition Coefficient (PC)*

Cluster	Hasil PC
K=2	0,685426108
K=3	0,643846403

Dari tabel 6 diperoleh informasi bahwa nilai PC dua *cluster* menunjukkan nilai yang lebih besar (mendekati 1) dibandingkan dengan nilai PC tiga *cluster*. Hal ini menunjukkan bahwa hasil klastering dengan jumlah dua *cluster* merupakan jumlah *cluster* yang ideal untuk mendapatkan informasi tentang pengelompokan penyakit yang diderita oleh masyarakat pengguna JAMKESDA di kota Samarinda.

## 5. KESIMPULAN

Berdasarkan hasil perhitungan menggunakan metode FCM dalam mengelompokkan penyakit yang diderita oleh masyarakat pengguna JAMKESDA di kota Samarinda, maka dapat disimpulkan bahwa metode FCM mampu melakukan pengelompokan data penyakit ke dalam dua kelompok yaitu penyakit *degeneratif* dan *infeksi*. Selanjutnya dari analisis uji hasil perhitungan validasi *cluster* menggunakan metode *Partition Coefficient (PC)* yang dilakukan, menunjukkan bahwa pengelompokan penyakit dengan dua *cluster* merupakan jumlah *cluster* yang ideal dalam mengelompokkan data penyakit. Perbandingan metode dalam *clustering*, optimalisasi metode FCM menggunakan metode pengambilan keputusan seperti *Simple Additive Weight*, algoritma genetika merupakan rencana penelitian ke depan.

**DAFTAR PUSTAKA**

- [1] Z. T. Budiman, W. Witanti, and D. Nursantika, "Klasifikasi Kecenderungan Penyakit Mata Di Jawa Barat dengan Association Rule dan NAÏVE BAYES CLASSIFIER," *SEMNASTEKNOMEDIA ONLINE*, vol. 5, no. 1, pp. 2-2-7, 2017.
- [2] N. Meisida, O. Soesanto, and H. K. Candra, "K-MEANS untuk Klasifikasi Penyakit Karies Gigi," *KLIK-KUMPULAN JURNAL ILMU KOMPUTER*, vol. 1, no. 1, pp. 12-22, 2016.
- [3] A. Bastian, "Penerapan Algoritma K-Means Clustering Analysis Pada Penyakit Menular Manusia (Studi Kasus Kabupaten Majalengka)," *Jurnal Sistem Informasi*, vol. 14, no. 1, pp. 28-34, 2018.
- [4] H. L. Sari and D. Suranti, "Perbandingan Algoritma Fuzzy C-Means (FCM) Dan Algoritma Mixture Dalam Penclusteraan Data Curah Hujan Kota Bengkulu," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2016.
- [5] A. Ramadhani, A. Farmadi, and I. Budiman, "Clustering Data Cuaca Untuk Pengenalan Pola Perioditas Iklim Wilayah Pelaihari Dengan Metode Fuzzy C-Means," *Jurnal Teknologi dan Industri (Diskontinu)*, vol. 3, no. 1, pp. 57-64, 2015.
- [6] N. Menon and R. Ramakrishnan, "Brain Tumor Segmentation in MRI images using unsupervised Artificial Bee Colony algorithm and FCM clustering," in *Communications and Signal Processing (ICCSP), 2015 International Conference on*, 2015, pp. 0006-0009: IEEE.
- [7] R. Preetha and G. Suresh, "Performance analysis of fuzzy c means algorithm in automated detection of brain tumor," in *Computing and Communication Technologies (WCCCT), 2014 World Congress on*, 2014, pp. 30-33: IEEE.
- [8] P. D. W. Ayu, "Segmentasi Kepala janin Pada Citra USG Dalam Ruang Warna RGB dengan Metode Fuzzy C-Means," *PROSIDING CSGTEIS 2013*, 2013.
- [9] Z. Zhang and B. Gu, "Intrusion Detection Network Based on Fuzzy C-Means and Particle Swarm Optimization," in *Proceedings of the 6th International Asia Conference on Industrial Engineering and Management Innovation*, 2016, pp. 111-119: Springer.
- [10] A. Ansari and A. Riasi, "Customer clustering using a combination of fuzzy c-means and genetic algorithms," *International Journal of Business and Management*, vol. 11, no. 7, p. 59, 2016.
- [11] Y. Zheng, B. Jeon, D. Xu, Q. Wu, and H. Zhang, "Image segmentation by generalized hierarchical fuzzy C-means algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 28, no. 2, pp. 961-973, 2015.
- [12] G. Zhu, J. Chen, and P. Zhang, "Fuzzy c-means clustering identification method of urban road traffic state," in *Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on*, 2015, pp. 302-307: IEEE.
- [13] N. Puspitasari, R. Rosmasari, and S. Stefanie, "Penentuan Prioritas Perbaikan Jalan Menggunakan Fuzzy C-Means: Studi Kasus Perbaikan Jalan Di Kota Samarinda," *Jurnal Teknologi dan Sistem Komputer*, vol. 5, no. 1, pp. 7-14, 2017.
- [14] T. A. Munandar and W. O. Widyarto, "Clustering Data Nilai Mahasiswa untuk Pengelompokan Konsentrasi Jurusan Menggunakan Fuzzy Cluster Means," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2013.
- [15] L. Rusdiana, "Aplikasi Berbasis Fuzzy C-Means Dalam Penentuan Predikat Kelulusan Mahasiswa," *Jurnal Ilmu Komputer*, vol. 2, no. 2, pp. 1-9, 2016.
- [16] R. J. E. Putra, N. Nasution, and Y. Yummastian, "Aplikasi E-Zakat Penerimaan dan Penyaluran Menggunakan Fuzzy C-Means (Studi Kasus: LAZISMU Pekanbaru)," *DIGITAL ZONE: JURNAL TEKNOLOGI INFORMASI DAN KOMUNIKASI*, vol. 6, no. 2, 2015.
- [17] A. Ahmadi and S. Hartati, "Penerapan Fuzzy C-Means dalam Sistem Pendukung Keputusan untuk Penentuan Penerima Bantuan Langsung Masyarakat (BLM) PNPM-MPd (Studi Kasus PNPM-MPd Kec. Ngadirojo Kab. Pacitan)," *Berkala Ilmiah MIPA*, vol. 23, no. 3, 2015.
- [18] N. Astuti, "Algoritma Fuzzy C-Means (FCM) Untuk Penentuan Nilai Center Radial Basis Function (RBF) Pada Klasifikasi Data Penyakit Karies Gigi," *Jurnal Elektronik Nasional Teknologi dan Ilmu Komputer*, vol. 1, no. 01, 2017.
- [19] D. Mariati, "Potensi Isolat Actinomyces dari Rizosfer Padi (*Oryza sativa* L.) Sebagai Penghasil Antibiotik," Universitas Muhammadiyah Surakarta, 2013.
- [20] M. Radji, "Buku ajar mikrobiologi panduan mahasiswa farmasi dan kedokteran," *Jakarta: Penerbit Buku Kedokteran EGC*, pp. 11-19, 2011.
- [21] B. S. Santoso, "PERBANDINGAN AKTIVITAS ANTIOKSIDAN ANTARA JUS BUAH MENGKUDU (*Morinda citrifolia*) DAN JUS RIMPANG TEMULAWAK (*Curcuma xanthorrhiza*)," *PHARMACON*, vol. 6, no. 3, 2017.
- [22] I. Irsalina, E. Supriyati, and T. Khotimah, "Clustering Gender Berdasarkan Nilai Maksimum Minimum Amplitudo Suara Berbasis Fuzzy C-Means (FCM)," *Prosiding SNATIF*, vol. 1, pp. 419-424, 2014.
- [23] M. N. Sutoyo and A. T. Sumpala, "Penerapan Fuzzy C-Means untuk Deteksi Dini Kemampuan Penalaran Matematis," *Scientific Journal of Informatics*, vol. 2, no. 2, pp. 129-135, 2016.