# Penerapan Metode *Fuzzy Subtractive Clustering* (Studi Kasus: PengelompokkanKecamatan di Provinsi Kalimantan Timur Berdasarkan Luas Daerah dan Jumlah Penduduk Tahun 2015)

***The Application of Fuzzy Subtractive Clustering Method (Case Study: Clustering Districts in East BorneoBased on Wide of The District and Total of Population in 2015)***

**Nur Azizah[1], Desi Yuniarti[2], Rito Goejantoro[3]**
[1,3]Laboratorium Statistika Komputasi Jurusan Matematika FMIPA Universitas Mulawarman
[2]Laboratorium Statistika Ekonomi dan Bisnis Jurusan Matematika FMIPA Universitas Mulawarman
E-mail: nurazizah_@outlook.com

***Abstract***

*Cluster analysis is a statistical analysis to classify the objects to be some clusters based on checked variables and similarity of character between the objects. Quality of human living or society has been influenced by many things. In reality, population density is very influential to the quality of human living because high population density will cause many problems that impact on deterioration of quality of human living. Fuzzy Subtractive Cluster (FSC) methods using the data as a candidate of cluster center, so that duty of computation is hanging on the number of data and is not hang at dimension of data. This study aims is to determine the results of FSC at clustering the district in East Borneo based on wide of the district and total of population in 2015. The result shows there is 8 until 24 districts which have high population density. From validity of cluster, it isfounded that the best result for clustering the district in East Borneo based on wide of the district and sum of citizen in 2015 is 2 clusters, there are narrow district with many citizen and wide district with few citizen.*

*Keywords: Clusteranalysis, Fuzzy Subtractive Clustering (FSC), Population density,Validity of cluster.*

## Pendahuluan

Analisis kluster merupakan analisis statistika yang bertujuan untuk mengelompokkan objek-objek amatan menjadi beberapa kelompok berdasarkan variabel-variabel yang diamati.Objek tersebut akan diklasifikasikan ke dalam satu atau lebih kluster (kelompok) sehingga objek-objek yang berada dalam satu kluster akan mempunyai kemiripan antara satu dengan yang lain (Santoso, 2015).

Dalam analisis kluster ada dua metode pengelompokkan, yaitu metode berhirarki (*hierarchical methods*) dan metode tidak berhirarki (*nonhierarchical methods*).Pada proses pengelompokkan (*clustering*) berhirarki atau tidak berhirarki, pembentukan kelompok dilakukan sedemikian rupa sehingga setiap objek berada tepat pada satu kelompok. Akan tetapi, pada suatu saat hal itu tidak dapat dilakukan, karena sebenarnya objek tersebut terletak di antara dua atau lebih kelompok yang lain. Sehingga perlu dilakukan pengelompokkan dengan menggunakan *fuzzy clustering* di mana dalam melakukan pengelompokkan mempertimbangkan tingkat keanggotaan himpunan *fuzzy* sebagai dasar pembobotan (Jang, Sun dan Mizutani, 1997).

Ada beberapa metode (algoritma) yang telah dikembangkan dalam analisis *fuzzy clustering*, salah satunya adalah metode *fuzzy subtractive clustering* (FSC). MetodeFSC ditemukan oleh Chiu pada tahun 1994 dengan suatu pusat kluster pasti merupakan salah satu data yang ikut dikluster, yaitu di mana derajat keanggotaannya pada kluster tersebut sama dengan satu (Kusumadewi dan Purnomo, 2013)

Kualitas hidup manusia atau masyarakat dipengaruhi oleh banyak hal dan kepadatan penduduk merupakan yang sangat berpengaruh, sebab adanya kepadatan penduduk yang tinggi akan banyak menimbulkan berbagai masalah yang berhubungan dengan masalah kependudukan misalnya kemiskinan, perumahan, lapangan pekerjaan, dan lain-lain. Adanya permasalahan yang timbul tersebut akan membawa dampak pada penurunan kualitas hidup masyarakat (Christiani, Tedjo dan Martono, 2014).

Sebagaimana pertumbuhan penduduk, persebaran penduduk di Kalimantan Timur juga tidak merata.Pola persebaran penduduk Kalimantan Timur menurut luas daerah sangat timpang, sehingga menyebabkan terjadinya perbedaan tingkat kepadatan penduduk yang mencolok antar daerah (Badan Pusat Statistik Provinsi Kalimantan Timur, 2016).

Penelitian ini dibatasi pada penerapan metode *fuzzy subtractive clustering* dengan objek pada luas daerah dan jumlah penduduk tiap kecamatan di Provinsi Kalimantan Timur tahun 2015. Indikator yang digunakan adalah luas daerah dan jumlah penduduk serta jari-jari atribut data yang dipilih adalah 0,25; 0,27; 0,3; 0,5; dan 0,7 agar menghasilkan kluster berturut-turut berjumlah 6, 5, 4, 3 dan 2.

Sehingga rumusannya adalah bagaimana hasil penerapan metode *fuzzy subtractive clustering* pada pengelompokkan kecamatan di Provinsi Kalimantan Timur berdasarkan luas daerah dan jumlah penduduk tahun 2015?

Dengan demikian tujuan dari penelitian ini yaitu mengetahui hasil penerapan metode *fuzzy subtractive clustering* pada pengelompokkan kecamatan di Provinsi Kalimantan Timur berdasarkan luas daerah dan jumlah penduduk di tahun 2015.

**Analisis Kluster**

Analisis kluster merupakan suatu teknik yang dipergunakan untuk mengelompokkan objek atau kasus ke dalam kelompok yang relatif homogen yang disebut kluster. Objek dalam setiap kelompok cenderung mirip atau sama dan tidak berbeda jauh (tidak sama) dengan objek dari kluster lainnya (Supranto, 2004).

Objek yang dikluster bisa berupa produk (barang dan jasa), makhluk hidup (tumbuhan dan hewan), atau orang (disebut responden, konsumen, partisipan, atau yang lain). Objek tersebut akan diklasifikasikan ke dalam satu atau lebih kluster sehingga objek-objek yang berada dalam satu kluster akan mempunyai kemiripan satu dengan yang lain (Johnson dan Wichern, 2002).

**Logika *Fuzzy***

Dalam kamus *Oxford*, istilah *fuzzy* didefinisikan sebagai *blurred* (kabur atau remang-remang), *indistinct* (tidak jelas), *imprecisely defined* (didefinisikan secara tidak presisi), *confused* (membingungkan), dan *vague* (tidak jelas). Bagi orang-orang yang belum pernah mendengar sistem *fuzzy,* bisa saja menjadi salah mengerti ketika membaca definisi-definisi tersebut. Namun menurut Naba (2009), penggunaan istilah *fuzzy* tidak dimaksudkan untuk mengacu pada sebuah sistem yang tidak jelas/remang-remang/kabur definisinya, cara kerjanya, atau deskripsinya. Sebaliknya, yang dimaksud dengan sistem *fuzzy* adalah suatu sistem yang dibangun dengan definisi, cara kerja, dan deskripsi yang jelas berdasarkan pada teori logika *fuzzy*.

Menurut Cox (1994), ada beberapa alasan mengapa orang menggunakan logika *fuzzy*, antara lain:

1. Konsep logika *fuzzy* mudah dimengerti. Karena logika *fuzzy* menggunakan dasar teori himpunan, maka konsep matematis yang mendasari penalaran logika *fuzzy* sangat sederhana dan mudah dimengerti.
2. Logika *fuzzy* sangat fleksibel, artinya mampu beradaptasi dengan perubahan-perubahan, dan ketidakpastian yang menyertai permasalahan.
3. Logika *fuzzy* memiliki toleransi terhadap data-data yang tidak tepat. Jika diberikan sekelompok data yang cukup homogen, dan kemudian ada beberapa data yang "eksklusif", maka logika *fuzzy* memiliki kemampuan untuk menangani data ekslusif tersebut.
4. Logika *fuzzy* mampu memodelkan fungsi-fungsi nonlinier yang kompleks.
5. Logika *fuzzy* dapat membangun dan mengaplikasikan pengalaman-pengalaman para pakar secara langsung tanpa harus melalui proses pelatihan.
6. Logika *fuzzy* dapat bekerja sama dengan teknik-teknik kendali secara konvensional. Hal ini pada umumnya terjadi pada aplikasi di bidang teknik mesin maupun teknik elektro.
7. Logika *fuzzy* didasarkan pada bahasa alami. Logika *fuzzy* menggunakan bahasa sehari-hari sehingga mudah dimengerti.

**Teori Himpunan *Fuzzy***

Pada teori himpunan *fuzzy*, komponen utama yang sangat berpengaruh adalah fungsi keanggotaan.Fungsi keanggotaan mempresentasikan derajat kedekatan suatu obyek terhadap atribut tertentu.Teori himpunan *fuzzy* merupakan kerangka matematis yang digunakan utnuk mempresentasikan ketidakpastian, ketidakjelasan, ketidaktepatan, kekurangan informasi, dan kebenaran parsial.Pembahasan tentang ketidakjelasan (*vagueness*) telah dimulai semenjak tahun 1937, ketika seorang filosof bernama Max Black mengemukakan pendapatnya tentang ketidakjelasan (Naba, 2009).

Black mendefinisikan suatu proporsi tentang ketidakjelasan sebagai suatu proporsi di mana status kemungkinan dari proporsi tersebut tidak didefinisikan dengan jelas. Sebagai contoh, untuk menyatakan seseorang termasuk dalam kategori muda, pernyataan 'muda' dapat memberikan interpretasi yang berbeda oleh setiap individu, dan kita tidak dapat memberikan umur tertentu untuk menyatakan seseorang masih muda atau tidak (Naba, 2009).

**Konsep Dasar Himpunan *Fuzzy***

Menurut Yan, Ryan dan Power (1994), himpunan *fuzzy* didasarkan pada gagasan untuk memperluas jangkauan fungsi keanggotaan pada himpunan *crisp* sedemikian sehingga fungsi tersebut mencakup bilangan *real* pada interval [0,1]. Derajat keanggotaannya menunjukkan bahwa suatu elemen dalam semesta pembicaraan tidak hanya berada pada 0 dan 1, namun juga nilai yang terletak di antaranya. Dengan kata lain, nilai kebenaran suatu pernyataan tidak hanya bernilai benar atau salah. Nilai 1 menunjukkan benar, nilai 0 menunjukkan salah dan masih ada nilai-nilai yang terletak antara benar dan salah (Kusumadewi dan Purnomo, 2013).

**Keanggotaan Himpunan *Fuzzy***

Himpunan *crisp* dapat didefinisikan dengan fungsi karakteristik.Misalkan, $U$ adalah semesta pembicaraan. Fungsi karakteristik $\mu_A(x)$ dari himpunan *crisp* $A$ pada $U$ memberi nilai dalam [0,1] dan didefinisikan sebagai $\mu_A(x) = 1$ jika $x$ adalah suatu anggota dari $A$ dan $\mu_A(x) = 0$ jika $x$ bukan anggota dari $A$, atau:

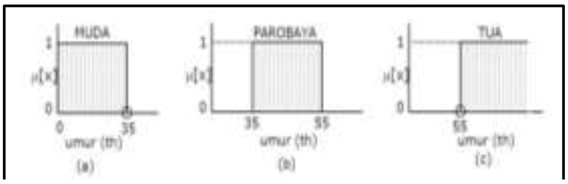
$$\mu_A(x) = \begin{cases} 1 & \text{jika dan hanya jika } x \in A \\ 0 & \text{jika dan hanya jika } x \notin A \end{cases} \quad (1)$$

Perhatikan bahwa batas himpunan $A$ adalah jelas dan melibatkan dua kelas dikotomisasi (yaitu $x \in A$ atau $x \notin A$) dan semesta pembicaraan $U$ adalah himpunan *crisp*(Robandi, 2006).

Untuk lebih memahami himpunan *fuzzy*, diberikan contoh berikut. Misalkan variabel umur dibagi menjadi 3 kategori, yaitu (Kusumadewi dan Purnomo, 2013):

MUDA : umur < 35 tahun
PAROBAYA : 35 ≤ umur ≤ 55 tahun
TUA : umur > 55 tahun

Nilai keanggotaan secara grafis, himpunan MUDA, PAROBAYA dan TUA ini dapat dilihat pada Gambar 1.

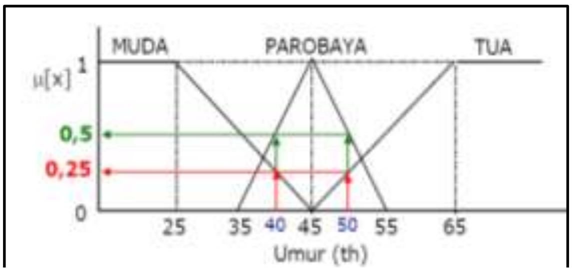



Gambar 1. Himpunan: (a) MUDA (b) PAROBAYA (c) TUA

Pada Gambar 2.1 dapat dilihat bahwa:

1. Apabila seseorang berusia 34 tahun, maka ia dikatakan MUDA ($\mu_{MUDA}[34]=1$);
2. Apabila seseorang berusia 35 tahun, maka ia dikatakan TIDAK MUDA ($\mu_{MUDA}[35]=0$);
3. Apabila seseorang berusia 35 tahun kurang 1 hari, maka ia dikatakan TIDAK MUDA ($\mu_{MUDA}[35 \text{ th } -1 \text{ hr}]=0$);
4. Apabila seseorang berusia 35 tahun, maka ia dikatakan PAROBAYA ($\mu_{PAROBAYA}[35]=1$);
5. Apabila seseorang berusia 34 tahun, maka ia dikatakan TIDAK PAROBAYA ($\mu_{PAROBAYA}[34]=0$);
6. Apabila seseorang berusia 35 tahun kurang 1 hari, maka ia dikatakan TIDAK PAROBAYA ($\mu_{TIDAK\ PAROBAYA}[35 \text{ th}-1 \text{ hr}]=0$);

Dari permisalan tersebut dapat dikatakan bahwa pemakaian himpunan *crisp* untuk menyatakan umur sangat tidak adil, adanya perubahan kecil saja pada suatu nilai mengakibatkan perbedaan kategori yang cukup signifikan.Himpunan *fuzzy* digunakan untuk mengantisipasi hal tersebut.Seseorang dapat masuk dalam 2 himpunan yang berbeda, MUDA dan PAROBAYA, PAROBAYA dan TUA.Seberapa besar eksistensinya dalam himpunan tersebut dapat dilihat pada nilai keanggotaannya.




Gambar 2. Himpunan *fuzzy* untuk variabel umur

Pada Gambar 2 dapat dilihat bahwa (Kusumadewi dan Purnomo, 2013):

1. Seseorang yang berumur 40 tahun, termasuk dalam himpunan MUDA dengan $\mu_{MUDA}[40] = 0{,}25$; namun dia juga termasuk dalam himpunan PAROBAYA dengan $\mu_{PAROBAYA}[40] = 0{,}5$.
2. Seseorang yang berumur 50 tahun, termasuk dalam himpunan MUDA dengan $\mu_{MUDA}[50] = 0{,}25$; namun dia juga termasuk dalam himpunan PAROBAYA dengan $\mu_{PAROBAYA}[50] = 0{,}5$.

***Fuzzy Clustering***

Menurut Jang, Sun dan Mizutami (1997), *fuzzy clustering* adalah metode pengelompokan berdasarkan derajat keanggotaan yang mencakup himpunan *fuzzy* sebagai dasar pembobotan bagi pengelompokan. Masing-masing data diberikan nilai kemungkinan untuk bisa bergabung ke setiap kelompok yang ada, yang berarti data tidak mutlak atau tegas menjadi anggota satu kelompok saja, tetapi juga mempunyai nilai kemungkinan untuk menjadi anggota kelompok lain dengan derajat keanggotaan terbesar menunjukkan kecenderungan yang tinggi suatu data untuk menjadi anggota kelompok tertentu.

***Fuzzy Subtractive Clustering***

Metode *Fuzzy Subtractive Clustering* (FSC) menggunakan data sebagai kandidat dari pusat kluster, sehingga beban komputasi tergantung dari jumlah data dan tidak bergantung dari dimensi data. Jumlah pusat kluster ditentukan melalui proses iterasi untuk mencari titik-titik dengan jumlah tetangga terbanyak (Laksono dan Hafis, 2013).

Adapun algoritma *fuzzy subtractive clustering* adalah sebagai berikut (Kusumadewi dan Purnomo, 2013):

1. Menginput data yang akan dikluster berupa matriks ($x_{ij}$) berukuran $n \times m$, dengan $n$ adalah banyaknya jumlah sampel data dan $m$ merupakan banyaknya atribut setiap data, serta $x_{ij}$ = data sampel ke-$i$ ($i = 1, 2, \ldots, n$) dan atribut ke-$j$ ($j = 1, 2, \ldots, m$).
2. Menetapkan nilai:
   a. Jari-jari setiap atribut data ($r_j$; $j = 1, 2, \ldots, m$). Jari-jari merupakan vektor yang akan menentukan seberapa besar pengaruh pusat kluster pada tiap-tiap data. Dengan demikian, suatu titik data akan memiliki potensi yang besar jika dia memiliki banyak tetangga dekat.
   b. *Squash factor* (q). *Squash factor* merupakan suatu konstanta untuk menetapkan besar radius titik-titik data sekitar pusat kluster yang akan diukur penurunan potensi datanya.
   c. *Accept ratio*. *Accept ratio* merupakan nilai batas bawah di mana suatu titik data yang menjadi kandidat (calon) pusat kluster diperbolehkan untuk menjadi pusat kluster.
   d. *Reject ratio*. *Reject ratio* merupakan batas atas di mana suatu data yang menjadi kandidat (calon) pusat kluster tidak diperbolehkan untuk menjadi pusat kluster.
   e. Nilai minimum dan maksimum data.
3. Menormalisasi setiap data dengan rumus:

$$x_{ij_{norm}} = \frac{x_{ij} - x_{j_{min}}}{x_{j_{max}} - x_{j_{min}}} \quad (2)$$

4. Menentukan potensi awal tiap titik data ($D_k$ ; $k$ = 1, 2, …, $n$). Adapun langkah-langkahnya sebagai berikut.
   a. Menghitung jarak setiap data terhadap $T_j$ dengan rumus:

$$Dist_{ij}(x_k) = \frac{T_j - x_{ij_{norm}}}{r_j}; \quad (3)$$

di mana $T_j = x_{ij_{norm}}$ ; $j$ = 1, 2, ..., $m$

Untuk menghitung $Dist_{ij}(x_1)$, nilai $T_j$ yang digunakan adalah $x_{1j_{norm}}$ dengan nilai $T_j$ tidak berubah hingga $i$=$n$. Selanjutnya untuk $Dist_{ij}(x_2)$ digunakan $x_{2j_{norm}}$ sebagai nilai $T_j$. Demikian seterusnya hingga $Dist_{ij}(x_n)$.

   b. Menentukan potensi awal tiap data dengan rumus:

Jika $m = 1$, maka $D_k = \sum_{i=1}^{n} e^{-4(Dist_{ij}(x_k)^2)}$ (4)

Jika $m > 1$, maka $D_k = \sum_{i=1}^{n} e^{-4(\sum_{j=1}^{m} Dist_{ij}(x_k)^2)}$ (5)

5. Mencari titik dengan potensi tertinggi:
   a. $M = \max[D_k | k = 1, 2, \ldots, n]$; untuk iterasi ke-1 (6)
   b. $Z = \max[D_k | k = 1, 2, \ldots, n]$; untuk iterasi ke-2, 3, dan seterusnya (7)
6. Menentukan pusat kluster dan kurangi potensinya terhadap titik-titik di sekitarnya dengan cara menetapkan vektor V sebagai calon pusat kluster dan menghitung nilai rasio

$$R = \frac{Z}{M}; \quad (8)$$

khusus pada iterasi pertama nilai Z=M.

Setelah nilai rasio diperoleh, ada 3 keadaan yang dapat terjadi:

1) Jika nilai *ratio* >*accept ratio*, calon pusat kluster dapat diterima sebagai pusat kluster baru dan ditulis $C_l$. Prosedur kerja yang harus dilakukan selanjutnya setelah mendapat kluster baru adalah mengurangi potensi titik-titik data yang lain dengan cara:

a) $$S_{ij} = \frac{C_{l_j} - x_{ij_{norm}}}{r_j * q} \quad (9)$$

dengan:

$S_{ij}$ = pengurangan potensi data sampel ke-$i$, atribut ke-$j$

$C_{l_j}$ = pusat kluster ke-$l$ dengan atribut ke-$j$

$x_{ij_{norm}}$ = data sampel ke-$i$ dan atribut ke-$j$ setelah normalisasi

$r_j$ = jari-jari setiap atribut data

$q$ = *squash factor*

b) $$D_{C_{l_i}} = \mathrm{M} \times e^{-4\left[\sum_{j=1}^{m}(S_{ij})^2\right]} \quad (10)$$

dengan:

$D_{C_{l_i}}$ = potensi data kluster ke-$l$ pada sampel ke-$i$

M = potensi data tertinggi pada iterasi pertama

$S_{ij}$ = pengurangan potensi data sampel ke-$i$, atribut ke-$j$

c) $$D_i^{\,t} = D_i^{\,t-1} - D_{C_{l_i}}; \quad (11)$$

dengan:

$D_i^{\,t}$ = potensi data baru sampel ke-$i$ dengan $t$ menunjukkan banyak iterasi

$D_i^{\,t-1}$ =potensi data baru sampel ke-$i$ pada iterasi sebelumnya

$D_{C_{l_i}}$ =potensi data kluster ke-$l$ pada sampel ke-$i$

2) Jika nilai *ratio* < *accept ratio* dan nilai *ratio* > *reject ratio*, calon baru akan diterima jika keberadaannya cukup jauh dari pusat kluster yang telah ada.

Prosedur dalam keadaan ini:

a) Misalkan $Md$ = -1;

b) Kerjakan untuk $l$ = 1 sampai $l$ = $p$; $p$ = banyaknya kluster

$$Sd_l = \sum_{j=1}^{m}\left(\frac{V_j - C_{l_j}}{r_j}\right)^2 \quad (12)$$

dengan:

$V_j$ = calon pusat kluster

$C_{l_j}$ = pusat kluster ke-$l$ pada atribut ke-$j$

$r_j$ = jari-jari setiap atribut data

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

Jika ($Sd_l$>Md), maka $Md$ tidak berubah

$Mds = \sqrt{Md}$;

dengan $Mds$ adalah jarak terdekat data calon pusat kluster dengan pusat kluster.

Jika (*ratio* + *Mds*) ≥ 1, calon pusat kluster diterima sebagai pusat kluster baru.
Jika (*ratio* + *Mds*) < 1, calon pusat kluster tidak diterima dan tidak akan dipertimbangkan kembali sebagai pusat kluster dan potensi data tersebut diset menjadi 0.

3) Jika nilai *ratio* < *accept ratio* dan nilai *ratio* < *reject ratio*, sudah tidak ada lagi calon pusat kluster baru dan iterasi dihentikan.

7. Mengembalikan pusat kluster dari bentuk ternormalisasi ke bentuk semula (denormalisasi).

$$C_{lj_{denorm}} = C_{lj} * (x_{max\,j} - x_{min\,j}) + x_{min\,j} \quad (13)$$

dengan:
$C_{lj_{denorm}}$ = pusat kluster ke-*l* pada atribut ke-*j* setelah didenormalisasi
$C_{lj}$ = pusat kluster ke-*l* pada atribut ke-*j* (bentuk ternormalisasi)
$x_{max\,j}$ = data maksimum pada atribut ke-*j*
$x_{min\,j}$ = data minimum pada atribut ke-*j*

8. Menghitung nilai simpangan baku kluster

$$\sigma_j = r_j * \frac{X_{j\,max} - X_{j\,min}}{\sqrt{8}} \quad (14)$$

dengan:
$\sigma_j$ = nilai simpangan baku kluster
$r_j$ = jari-jari setiap atribut data
$x_{max\,j}$ = data maksimum pada atribut ke-*j*
$x_{min\,j}$ = data minimum pada atribut ke-*j*

Hasil dari algoritma *subtractive clustering* ini berupa matriks pusat kluster (C) dan simpangan baku ($\sigma$) akan digunakan untuk menentukan nilai parameter fungsi keanggotaan Gauss. Derajat keanggotaan suatu titik data $x_i$ pada kluster ke-*l* adalah (Kusumadewi dan Purnomo, 2013):

$$\mu_{li} = e^{-\sum_{j=i}^{m} \frac{\left(x_{ij} - C_{lj}\right)^2}{2\sigma_j^2}} \quad (15)$$

dengan:
$\mu_{li}$ = derajat keanggotaan kluster ke-*l* pada sampel ke-*i*
$x_{ij}$ = data sampel ke-*i* dan atribut ke-*j*
$C_{lj}$ = pusat kluster ke-*l* pada atribut ke-*j* (bentuk ternormalisasi)
$\sigma_j^2$ = kuadrat nilai simpangan baku / variansi kluster

9. Menentukan letak kluster

Dalam menentukan letak kluster suatu data, digunakan nilai derajat keanggotaan yang telah didapat sebelumnya.Di mana kluster dengan nilai derajat keanggotaan yang paling besar pada suatu data merupakan letak kluster data tersebut.

**Validitas *Fuzzy Clustering***

Dalam metode pengelompokkan yang menggunakan konsep *fuzzy*, sebuah data bisa menjadi anggota di semua kluster dengan nilai derajat keanggotaan yang dimilikinya.Semakin tinggi nilai derajat keanggotaan pada sebuah kluster maka semakin besar kecenderungannya menjadi anggota kluster tersebut.Bezdek (1981) mengusulkan validitas dengan menghitung koefisien partisi atau *Partition Coefficient* (PC) sebagai evaluasi nilai keanggotaan data pada setiap kluster.Nilai PC Index (PCI) hanya mengevaluasi nilai derajat keanggotaan tanpa memandang nilai data. Nilainya dalam rentang [0,1], nilai yang semakin besar (mendekati 1) mempunyai arti bahwa kualitas kluster yang didapat semakin baik. Berikut formula untuk menghitung PCI:

$$\text{PCI} = \frac{1}{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{K} \mu_{ij}{}^2\right) \quad (16)$$

*N* merupakan jumlah data dalam set data, *K* merupakan jumlah kluster, sedangkan $u_{ij}$ menyatakan nilai keanggotaan data ke-*i* pada kluster ke-*j* (Prasetyo, 2014).

Bezdek (1974) juga mengusulkan validitas dengan menghitung entropi partisi atau *Partition Entropy* (PE).Nilai PE *Index* (PEI) mengevaluasi keteracakan data dalam kluster. Nilainya berada dalam rentang [0,1], nilai yang semakin kecil (mendekati 0) mempunyai arti bahwa kluster yang didapat semakin baik. Berikut formula untuk menghitung PEI (Prasetyo, 2014):

$$\text{PEI} = -\frac{1}{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{K} \mu_{ij} \times {}^2log\, \mu_{ij}\right) \quad (17)$$

PCI dan PEI memiliki kecenderungan monotonik terhadap K. Modifikasi nilai PCI (MPCI) dilakukan oleh Dave (1996) terhadap kecenderungan monotonik tersebut. Formula yang digunakan seperti berikut:

$$\text{MPCI} = 1 - \frac{K}{K-1}(1 - \text{PCI}) \quad (18)$$

Nilai MPCI yang didapat adalah 0 ≤ MPCI ≤ 1, nilai yang semakin besar (mendekati 1) mempunyai arti bahwa kualitas kluster yang didapat semakin baik (Prasetyo, 2014).

Fukuyama dan Sugeo (1989) dalam Prasetyo (2014) mengusulkan validitas dengan formula seperti persamaan berikut:

$$\begin{aligned}\text{FSI} &= \sum_{j=1}^{K}\sum_{i=1}^{N} u_{ij}^{m} \times d\left(x_i, c_j\right)^2 - \sum_{j=1}^{K}\sum_{i=1}^{N} u_{ij}^{m} \times d\left(c_j, \bar{x}\right)^2 \\ &= J_m(u,c) - K_m(u,c)\end{aligned} \quad (19)$$

*m* merupakan bobot pangkat (*weighting exponent*), nilainya *m* > 1, $d\left(x_i, c_j\right)$merupakan jarak antara data ke-*i* terhadap pusat kluster ke-*j*,$c_j$ adalah pusat kluster ke-*j*, $d\left(c_j, \bar{x}\right)$ merupakan jarak antara pusat hasil klustering terhadap rata-

rata semua data, $J_m(u,c)$ adalah nilai fungsi objektif yang mengukur kohesi, sedangkan $K_m(u,c)$adalah nilai fungsi objektif yang mengukur nilai separasi. Secara umum, nilai *Fukuyama Sugeno Index* (FSI) yang semakin kecil mempunyai arti bahwa kualitas kluster yang didapatkan semakin baik.

Xie dan Beni (1991) juga mengusulkan validitas untuk mengevaluasi kluster yang didapat dengan modifikasi oleh Pal dan Bezdek (1995). Formula yang digunakan seperti berikut:

$$\text{XBI} = \sum_{j=1}^{K}\sum_{i=1}^{N}\frac{u_{ij}^{m} \times d\left(x_i, c_j\right)^2}{N \times \min\limits_{\text{i, j}}\left(d\left(c_i, c_j\right)^2\right)} = \frac{J_m(u,c)/N}{Sep(c)} \quad (20)$$

$J_m(u,c)$adalah ukuran kohesi, sedangkan *Sep(c)* adalah ukuran separasi. Secara umum, nilai yang terbaik untuk *Xie Beni Index* (XBI) adalah nilai indeks yang semakin kecil.Nilai XBI yang semakin kecil mempunyai arti kualitas hasil pengelompokkan yang semakin baik (Wu danYang, 2005) (Prasetyo, 2014).

**Analisis Deskriptif**

Hasil analisis statistika deskriptif yang dilakukan dapat dilihat pada Tabel 1.

**Tabel 1. Statistika Deskriptif**

| | Luas Kecamatan | Jumlah Penduduk |
|---|---|---|
| Minimum | 10,22 | 2.286 |
| Maksimum | 7.764,5 | 135.814 |
| Jumlah | 149.147,72 | 3.433.485 |
| Rata-Rata | 1.448,0361 | 33.381,6311 |
| Variansi | 2.554.025,085 | 1.195.247.581,62 |

Berdasarkan Tabel 1 dapat dilihat bahwa nilai minimum pada data luas kecamatan adalah sebesar 10,22 yang berarti bahwa luas kecamatan paling kecil di Provinsi Kalimantan Timur adalah sebesar 10,22 km$^2$. Luas kecamatan paling besar di Provinsi Kalimantan Timur adalah sebesar 7.764,5 km$^2$.Jumlah luas kecamatan Provinsi Kalimantan Timur adalah sebesar 149.147,72 km$^2$.Rata-rata luas kecamatan di Provinsi Kalimantan Timur adalah sebesar 1.448,0361 km$^2$. Dan nilai variansi pada data luas kecamatan sebesar 2.554.025,085.

Berdasarkan Tabel 1 juga dapat diketahui bahwa nilai minimum pada data jumlah penduduk adalah sebesar 2.286 yang berarti bahwa jumlah penduduk paling sedikit di Provinsi Kalimantan Timur adalah sebesar 2.286 jiwa. Jumlah penduduk paling banyak di Provinsi Kalimantan Timur adalah sebesar 135.814 jiwa. Total jumlah penduduk Provinsi Kalimantan Timur adalah sebesar 3.433.485 jiwa. Rata-rata jumlah penduduk di Provinsi Kalimantan Timur adalah sebesar 33.382 jiwa. Dan nilai variansi pada data luas daerah sebesar 1.195.247.581,62.

**Menginput Data**

Data dibuat dalam bentuk matriks berukuran 103×2.Ukuran matriks ini menunjukkan bahwa jumlah sampel sebanyak 103 kecamatan dan variable yang digunakan dalam analisis ada dua yaitu luas kecamatan dan jumlah penduduk. Adapun data dalam bentuk matriks adalah sebagai berikut:

$$x = \begin{pmatrix} 221{,}29 & 54.819 \\ 17{,}18 & 74.012 \\ 11{,}12 & 36.919 \\ \vdots & \vdots \\ 901{,}8 & 2.873 \end{pmatrix}$$

Kolom pertama memuat luas daerah tiap kecamatan dengan satuan kilometer persegi (km$^2$) dan selanjutnya disebut sebagai atribut pertama serta kolom kedua memuat jumlah penduduk tiap kecamatan dengan satuan jiwa dan selanjutnya disebut sebagai atribut kedua.Baris pertama memuat data sampel pertama, hingga baris terakhir adalah data sampel ke-103.

**Menentukan Nilai-Nilai Parameter**

a. Jari-jari: $r_j$ = 0,25; 0,27; 0,3; 0,5; 0,7 *j*=1, 2
   Jari-jari yang digunakan dalam penelitian ini ada sebanyak 5 yakni 0,25; 0,27; 0,3; 0,5 dan 0,7.
b. *Squash factor*: $q = 1{,}25$
   *Squash factor* sebesar 1,25 menunjukkan bahwa besar radius titik-titik data sekitar pusat kluster yang akan diukur penurunan potensi datanya adalah sebesar 1,25
c. *Accept ratio* sebesar 0,5 menunjukkan bahwa nilai batas bawah dari titik data yang menjadi calon pusat kluster agar diperbolehkan untuk menjadi pusat kluster adalah sebesar 0,5.
d. *Reject ratio* : 0,15
   *Reject ratio* sebesar 0,15 menunjukkan bahwa nilai batas atas dari titik data yang menjadi calon pusat kluster tidak diperbolehkan untuk menjadi pusat kluster adalah sebesar 0,15.
e. Data minimum :$x_{1_{min}} = 10{,}22$; $x_{2_{min}} = 2.286$. Data minimum pada atribut pertama sebesar 10,22 menunjukkan luas kecamatan paling kecil dan data minimum pada atribut kedua sebesar 2.286 menunjukkan jumlah penduduk paling sedikit.
f. Data maksimum: $x_{1_{max}} = 7.764{,}5$; $x_{2_{max}} = 135.184$.Data maksimum pada atribut pertama sebesar 7764,5 menunjukkan luas kecamatan paling besar dan data maksimum pada atribut kedua sebesar 135.184 menunjukkan jumlah penduduk paling banyak.

**Menormalisasi Data**

Dalam menormalisasi data, digunakan persamaan (2) dan diperoleh hasil seperti pada Tabel 2.

**Tabel 2. Data Setelah Dinormalisasi**

| i | Kecamatan | Luas Kecamatan | Jumlah Penduduk |
|---|---|---|---|
| 1 | Palaran | 0,027219806 | 0,393423102 |
| 2 | Samarinda Ilir | 0,000897569 | 0,537160745 |
| 3 | Samarinda Kota | 0,000116065 | 0,259368822 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 101 | Long Bagun | 0.639773132 | 0.052618177 |
| 102 | Long Hubung | 0.067147433 | 0 |
| 103 | Laham | 0.114979083 | 0.004396082 |

Tabel 2 menunjukkan hasil normalisasi dari tiap data. Nilai 0,027219806 merupakan data normalisasi dari daerah Palaran atau sampel pertama dan dari variabel luas daerah atau atribut pertama. Nilai 0,393423102 merupakan data normalisasi dari daerah Palaran atau sampel pertama dan dari variabel jumlah penduduk atau atribut kedua, dan seterusnya.

**Mencari Nilai Pusat Kluster**

Analisis dengan r = 0,25 menghasilkan 6 kluster, r = 0,27 menghasilkan 5 kluster, r = 0,3 menghasilkan 4 kluster, r = 0,5 menghasilkan 3 kluster dan r = 0,7 menghasilkan 2 kluster. Adapun nilai pusat kluster yang dihasilkan adalah sebagai berikut.

$$C_{r=0,25} = \begin{bmatrix} 0,105849 & 0,097111 \\ 0,041959 & 0,491994 \\ 0,393865 & 0,068338 \\ 0,000111 & 0,800753 \\ 0,230657 & 0,287692 \\ 0,664946 & 0,0607663 \end{bmatrix}$$

$$C_{r=0,27} = \begin{bmatrix} 0,105849 & 0,097111 \\ 0,041959 & 0,491994 \\ 0,393865 & 0,068338 \\ 0,000111 & 0,800753 \\ 0,736090 & 0,089846 \end{bmatrix}$$

$$C_{r=0,3} = \begin{bmatrix} 0,105849 & 0,097111 \\ 0,041959 & 0,491994 \\ 0,393865 & 0,068338 \\ 0003559 & 0,909862 \end{bmatrix}$$

$$C_{r=0,5} = \begin{bmatrix} 0,118435 & 0,120200 \\ 0 & 0,639828 \\ 0,639773 & 0,052618 \end{bmatrix}$$

$$C_{r=0,7} = \begin{bmatrix} 0,149013 & 0,162970 \\ 0,000111 & 0,800753 \end{bmatrix}$$

Pada $C_{r=0,25}$, nilai 0,118435 menunjukkan nilai pusat kluster pada atribut pertama di kluster pertama, nilai 0,120200 menunjukkan nilai pusat kluster pada atribut kedua di kluster pertama, nilai 0 menunjukkan nilai pusat kluster pada atribut pertama di kluster kedua, nilai 0,639828 menunjukkan nilai pusat kluster pada atribut kedua di kluster kedua, nilai 0,639773 menunjukkan nilai pusat kluster pada atribut pertama di kluster ketiga, dan nilai 0,052618 menunjukkan nilai pusat kluster pada atribut kedua di kluster ketiga. Begitu seterusnya hingga $C_{r=0,7}$.

**Mengembalikan Pusat Kluster dari Bentuk Normalisasi ke Bentuk Semula**

Dalam mengembalikan pusat kluster dari bentuk normalisasi ke bentuk semula (*denormalisasi*), digunakan persamaan (13), nilai pusat kluster yang diperolehdapat dilihat pada Tabel 3 – Tabel 7.

**Tabel 3. Nilai pusat kluster dengan $r_j$ = 0,25**

| Kluster | Luas Kecamatan | Jumlah Penduduk | Data ke- | Keterangan Kluster |
|---|---|---|---|---|
| 1 | 831 | 15.253 | 73 | Luas kecamatan kecil, jumlah penduduk sedikit. |
| 2 | 335,58 | 67.981 | 29 | Luas kecamatansangat kecil, jumlah penduduk banyak. |
| 3 | 3.064,36 | 11.411 | 80 | Luas kecamatansangat besar, jumlah penduduk sangat sedikit. |
| 4 | 11,08 | 109.209 | 14 | Luas kecamatansangat kecil sekali,jumlah penduduk sangat banyak sekali. |
| 5 | 1.798,8 | 40.701 | 92 | Luas kecamatan besar, jumlah penduduk sangat banyak. |
| 6 | 5.166,4 | 10.400 | 43 | Luas kecamatansangat besar sekali, jumlah penduduk sangat sedikit sekali. |

**Tabel 4. Nilai pusat kluster dengan $r_j$ = 0,27**

| Kluster | Luas Kecamatan | Jumlah Penduduk | Data ke- | Keterangan Kluster |
|---|---|---|---|---|
| 1 | 831 | 15.253 | 73 | Luas kecamatan sedang, jumlah penduduk sedang. |
| 2 | 335,58 | 67.981 | 29 | Luas kecamatan kecil, jumlah penduduk banyak |
| 3 | 3.064,36 | 11.411 | 80 | Luas kecamatan besar, jumlah penduduk sangat sedikit. |
| 4 | 11,08 | 109.209 | 14 | Luas kecamatan sangat kecil, jumlah penduduk sangat banyak. |
| 5 | 5.718,07 | 14.283 | 57 | Luas kecamatan sangat besar, jumlah penduduk sedikit. |

**Tabel 5. Nilai pusat kluster dengan $r_j$ = 0,3**

| Kluster | Luas Kecamatan | Jumlah Penduduk | Data ke- | Keterangan Kluster |
|---|---|---|---|---|
| 1 | 831,00 | 15.253 | 73 | Luas kecamatan besar, jumlah penduduk sedikit. |
| 2 | 335,58 | 67.981 | 29 | Luas kecamatan kecil, jumlah penduduk banyak. |
| 3 | 3.064,36 | 11.411 | 80 | Luas kecamatan sangat besar, jumlah penduduk sangat sedikit. |
| 4 | 37,82 | 123.778 | 11 | Luas kecamatan sangat kecil, jumlah penduduk sangat banyak. |

**Tabel 6. Nilai pusat kluster dengan $r_j$ = 0,5**

| Kluster | Luas Kecamatan | Jumlah Penduduk | Data ke- | Keterangan Kluster |
|---|---|---|---|---|
| 1 | 928,60 | 18.336 | 86 | Luas kecamatan sedang, jumlah penduduk sedang |
| 2 | 10,22 | 87.721 | 16 | Luas kecamatan kecil, jumlah penduduk banyak |
| 3 | 4.971,20 | 9.312 | 101 | Luas kecamatan besar, jumlah penduduk sedikit |

**Tabel 7. Nilai pusat kluster dengan $r_j$ = 0,7**

| Kluster | Luas Kecamatan | Jumlah Penduduk | Data ke- | Keterangan Kluster |
|---|---|---|---|---|
| 1 | 1.165,71 | 24.047 | 94 | Luas kecamatan besar, Jumlah penduduk sedikit |
| 2 | 11,08 | 109.209 | 14 | Luas kecamatan kecil, Jumlah penduduk banyak |

**Menghitung Nilai Simpangan Baku**

Dalam menghitung nilai simpangan baku kluster, digunakan persamaan (14) dan diperoleh hasil sebagai berikut.

**Tabel 8. Nilai Simpangan Baku**

| r | Simpangan baku | |
|---|---|---|
| | Atribut pertama | Atribut kedua |
| 0,25 | 685,387997 | 11.802,319285 |
| 0,27 | 740,219036 | 12.746,504828 |
| 0,3 | 822,4655958 | 14.162,7831417 |
| 05 | 1.370,775993 | 23.604,638570 |
| 0,7 | 1.919,0863902 | 33.046,4939973 |

**Menghitung Derajat Keanggotaan**

Dalam menghitung derajat keanggotaan, digunakan persamaan (15) dan diperoleh nilai derajat keanggotaan tiap-tiap kluster dan sampel.

**Menentukan Letak Kluster**

Letak kluster masing-masing daerah ditentukan berdasarkan nilai derajat keanggotaannya pada tiap kluster.Di mana nilai derajat keanggotaan tertinggi dari semua kluster yang ada menunjukkan letak kluster daerah tersebut. Hasil menunjukkan bahwa dengan r = 0,25, data pertama terletak pada kluster kedua, data kedua memiliki nilai derajat keanggotaan tertinggi pada kluster kedua, hal ini berarti bahwa data kedua terletak pada kluster kedua, dan demikian seterusnya.

Pada saat r = 0,25 atau terbentuk 6 kluster,Kecamatan Sungai Kunjang, Samarinda Ulu, Samarinda Utara, Sungai Pinang, Balikpapan Selatan, Balikpapan Utara, Balikpapan Tengah, dan Balikpapan Barat perlu mendapat perhatian lebih karena termasuk padakluster keempat di mana kluster tersebut merupakan kluster dengan luas daerah sangat kecil sekali tetapi jumlah penduduk sangat banyak sekali.

Pada saat r = 0,27 atau terbentuk 5 kluster dapat dilihat bahwa Kecamatan Samarinda Seberang, Loa Janan Ilir, Sungai Kunjang, Samarinda Ulu, Samarinda Utara, Sungai Pinang, Balikappan Selatan, Balikpapan Utara, Balikpapan Tengah, Balikpapan Barat, dan Tenggarong perlu mendapat perhatian lebih karena termasuk pada kluster keempat di mana kluster tersebut merupakan kluster dengan luas daerah sangat kecil tetapi jumlah penduduk sangat banyak.

Pada saat r = 0,3 atau terbentuk 4 kluster dapat dilihat bahwa Kecamatan Sungai Kunjang, Samarinda Ulu, Samarinda Utara, Sungai Pinang, Balikpapan Selatan, Balikpapan Utara, Balikpapan Tengah, dan Tenggarong perlu mendapat perhatian lebih karena termasuk pada kluster keempat di mana kluster tersebut merupakan kluster dengan luas daerah sangat kecil tetapi jumlah penduduk sangat banyak.

Pada saat r = 0,5 atau terbentuk 3 kluster dapat dilihat bahwa Kecamatan Palaran Ilir, Samarinda Utara, Sungai Pinang,Balikpapan Selatan, seluruh Kota Balikpapan, seluruh Kota Bontang, Penajam, Tanah Grogot, Tanjung Redeb, Sangatta Utara, Samboja, Loa Janan, Tenggarong, dan Tenggarong Seberang perlu mendapat perhatian lebih karena termasuk pada kluster kedua di mana kluster tersebut merupakan kluster dengan luas daerah kecil tetapi jumlah penduduk banyak.

Pada saat r = 0,7 atau terbentuk 2 kluster dapat dilihat bahwa Kecamatan Samarinda Ilir, Samarinda Seberang, Loa Janan Ilir, Sungai Kunjang, Samarinda Ulu, Samarinda Utara, Sungai Pinang, Balikpapan Selatan, Balikpapan Timur, Balikpapan Utara, Balikpapan Tengah, Balikpapan Barat, Balikpapan Kota, Bontang Selatan, Bontang Utara, Penajam, Tanah Grogot, Tanjung Redeb, Sangatta Utara, Loa Janan, Tenggarong, dan Tenggarong Seberang perlu mendapat perhatian lebih karena termasuk pada kluster kedua di mana kluster tersebut merupakan kluster dengan luas daerah kecil tetapi jumlah penduduk banyak.

**Menghitung Validitas Kluster**

PCI dihitung menggunakan persamaan (16) dan menghasilkan nilai sebagai berikut.

**Tabel 9. Nilai PCI**

| Jumlah Kluster | Nilai PCI |
|---|---|
| 6 | 0,428611545 |
| 5 | 0,426009184 |
| 4 | 0,44216901 |
| 3 | 0,551756595 |
| 2 | **0,610935194** |

PEI dihitung menggunakan persamaan (17) dan menghasilkan nilai sebagai berikut.

**Tabel 10. Nilai PEI**

| Jumlah Kluster | Nilai PEI |
|---|---|
| 6 | 0,604275623 |
| 5 | 0,481245685 |
| 4 | 0,484954004 |
| 3 | 0,524676027 |
| 2 | **0,461963998** |

MPCI dihitung menggunakan persamaan (18) dan menghasilkan nilai sebagai berikut.

**Tabel 11. Nilai MPCI**

| Jumlah Kluster | Nilai MPCI |
|---|---|
| 6 | 0,314333854 |
| 5 | 0,28251148 |
| 4 | 0,256225347 |
| 3 | **0,327634892** |
| 2 | 0,221870388 |

FSI dihitung menggunakan persamaan (19) dan menghasilkan nilai sebagai berikut.

**Tabel 12. Nilai FSI**

| Jumlah Kluster | Nilai FSI |
|---|---|
| 6 | -43.746.534.592,5131 |
| 5 | -43.646.112.295,1199 |
| 4 | **-54.592.304.890,3321** |
| 3 | -34.922.585.696,7829 |
| 2 | -42.573.009.736,5859 |

XBI dihitung menggunakan persamaan (20) dan menghasilkan nilai sebagai berikut.

**Tabel 13. Nilai XBI**

| Jumlah Kluster | Nilai XBI |
|---|---|
| 6 | 3,795505693 |
| 5 | 1,505139383 |
| 4 | 1,156510043 |
| 3 | 0,859357405 |
| 2 | **0,033935859** |

Berdasarkan Tabel 9 – Tabel 13 dapat dilihat bahwa pembentukan menjadi 2 kluster lebih baik untuk mengelompokkan kecamatan di Provinsi Kalimantan Timur berdasarkan luas daerah dan jumlah penduduk tahun 2015 jika dibandingkan dengan pembentukan menjadi 6, 5, 4, atau 3 kluster. Dari indeks validitas dapat dilihat bahwa untuk menjadi 2 kluster mempunyai 3 nilai tingkat validitas yang lebih baik dibandingkan kluster lain dengan nilai PCI terbesar yakni sebesar 0,610935194, serta nilai PEI dan XBI terkecil yakni berturut-turut sebesar 0,461963998 dan 0,033935859. Jumlah kluster sebanyak 4 dan 3 hanya mempunyai satu indeks validitas yang baik yakni berturut-turut pada nilai FSI sebesar -4.592.304.890,3321 dan nilai MPCI sebesar 0,327634892.

**Kesimpulan**

Kesimpulan dari hasil analisis penerapan metode *fuzzy subtractive clustering*adalah jumlah yang baik untuk mengelompokkan kecamatan di Provinsi Kalimantan Timur berdasarkan luas daerah dan jumlah penduduk ada sebanyak 2 kluster atau dengannilai r = 0,7.Sebanyak 22 daerah yang termasuk dalam kluster luas daerah kecil tetapi jumlah penduduk banyak dan 81 daerah lainnya termasuk dalam kluster luas kecamatan besar tetapi jumlah penduduk sedikit.

**Daftar Pustaka**

Badan Pusat Statistik Provinsi Kalimantan Timur.(2015). Kalimantan Timur dalam Angka (*Kalimantan Timur in Figures*) 2015 Nomor Katalog: 1102001.64.

Christiani, C., Tedjo, P & Martono, B. (2014). Analisis Dampak Kesehatan Penduduk Terhadap Kualitas Hidup Masyarakat Provinsi Jawa Tengah. *Serat Acitya – Jurnal Ilmiah UNTAG Semarang.*

Cox, E. (1994).*The Fuzzy Systems Handbook (A Prsctitioner's Guide to Building, Using, and Maintaining Fuzzy systems)*. Massachusetts: Academic Press, Inc.

Jang, J-S.R., Sun, C-T & Mizutani, E. (1997).*Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. New Jersey: Prentice-Hall Inc.

Johnson, R.A & Wichern, D.W. (2002).*Applied Multivariate Statistical Analysis Fifth Edition*. New Jersey: Prentice Hall.

Kusumadewi, S & Purnomo, H. (2013).*Aplikasi Logika Fuzzy untuk Pendukung Keputusan Edisi 2*. Yogyakarta: Graha Ilmu.

Laksono, H & Hafis, M. (2013).*Aplikasi Fuzzy Clustering dengan Menggunakan Algoritma Subtractive Clustering untuk Perkiraan Kebutuhan Energi Listrik Jangka Panjang di Provinsi Sumatera Barat dari Tahun 2012-2021*.Jurnal Teknologi Informasi & Pendidikan, 6.

Naba, A. (2009). *Belajar Cepat Fuzzy Logic Menggunakan Matlab.* Yogyakarta: ANDI.

Prasetyo.(2014). *Data Mining Mengolah Data Menjadi Informasi Menggunakan Matlab*. Yogyakarta: ANDI.

Santoso, S. (2015).*Menguasai Statistik Multivariat*. Jakarta: PT Elex Media Komputindo

Supranto, J. (2004). *Analisis Multivariat Arti & Interpretasi*. Jakarta: Rineka Cipta.

Yan, J., Ryan, M & Power, J. (1994).*Using Fuzzy Logic Towards intelligent Systems.* London: Prentice-Hall International.