

Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat Menggunakan Metode Kohonen SOM dan *K-Means*

Marina Marsudi Putri dan Kartika Fithriasari.
Jurusan Statistika, FMIPA, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia
e-mail: kartika_f@statistics.its.ac.id

Abstrak— Kondisi kesehatan masyarakat di Jawa Timur yang tidak homogen akan menyulitkan bagi Tim Pembina Kota Sehat pada saat melakukan pembinaan dan monitoring. Berdasarkan hal tersebut diperlukan pengelompokan kabupaten dan kota di Jawa Timur yang didasarkan pada kemiripan karakteristik kondisi kesehatan masyarakat, sehingga proses pembinaan dan monitoring akan dilakukan berdasarkan kelompok kabupaten dan kota yang terbentuk. Pengelompokan kabupaten/kota di Jawa Timur akan dilakukan menggunakan metode Kohonen SOM kemudian akan dibandingkan dengan hasil pengelompokan metode klasik *K-Means*, yang merupakan metode pengelompokan yang populer dan sering digunakan, berdasarkan kriteria nilai *icdrate* (*internal cluster dispersion rate*). Pada pengujian Bartlett diperoleh kesimpulan bahwa terdapat hubungan atau korelasi di antara variabel penelitian. Setelah dilakukan analisis faktor untuk mereduksi variabel diperoleh 3 faktor baru yang terbentuk. Berdasarkan nilai *Pseudo Fstatistics* yaitu sebesar 13,819, hasil pengelompokan terbaik adalah menggunakan metode Kohonen SOM dengan jenis topologi *hextop*. Sedangkan pada metode *K-Means* nilai *Pseudo Fstatistics* terbesar yaitu 9,781 ketika digunakan kelompok sebanyak 2 klaster. Perbandingan hasil klaster terbaik berdasarkan nilai *icdrate* diperoleh kesimpulan bahwa nilai *icdrate* metode Kohonen SOM yaitu sebesar 0,962 lebih kecil dibandingkan dengan nilai *icdrate* hasil pengelompokan metode *K-Means* yaitu sebesar 0,988. Pada hasil pengujian *One-way MANOVA* diperoleh kesimpulan bahwa pada masing-masing kelompok yang terbentuk memiliki perbedaan.

Kata Kunci—*icdrate*, Indikator Kesehatan Masyarakat, *K-means*, Kohonen SO.

I. PENDAHULUAN

KONDISI kesehatan masyarakat di Jawa Timur yang tidak homogen akan menyulitkan bagi Tim Pembina Kota Sehat pada saat melakukan pembinaan dan monitoring. Berdasarkan hal tersebut diperlukan pengelompokan kabupaten/kota di Jawa Timur yang didasarkan pada kemiripan karakteristik kondisi kesehatan masyarakat, sehingga proses pembinaan dan monitoring akan dilakukan berdasarkan kelompok-kelompok kabupaten dan kota yang terbentuk.

Indikator kesehatan dapat digunakan untuk mengetahui kondisi kesehatan masyarakat di suatu wilayah, maka dari itu pengelompokan kabupaten/kota di Jawa Timur akan dilakukan berdasarkan pada faktor pembentuk indikator kesehatan di tiap wilayah. Variabel yang mempengaruhi indikator kesehatan

antara lain yaitu variabel ekonomi, pendidikan dan epidemiologi [1].

Pada penelitian ini pengelompokan kabupaten/kota di Jawa Timur akan dilakukan menggunakan metode Kohonen SOM. Kohonen SOM memiliki keunggulan yaitu dapat mengelompokan data yang mengandung *overlapping* dan mengatasi sifat non linier pada klaster yang terbentuk [2]. Hasil pengelompokan kabupaten dan kota menggunakan metode Kohonen SOM akan dibandingkan dengan hasil pengelompokan metode klasik *K-Means*, yang merupakan metode pengelompokan yang populer dan sering digunakan, berdasarkan kriteria nilai *icdrate* (*internal cluster dispersion rate*) yang merupakan tingkat dispersi di dalam klaster.

II. TINJAUAN PUSTAKA

A. Uji Bartlett

Uji Bartlett digunakan untuk mengetahui ada atau tidaknya hubungan antar variabel dalam kasus multivariat. Jika variabel x_1, x_2, \dots, x_p bersifat saling bebas (independen), sehingga matriks korelasi antar variabel sama dengan matriks identitas [3]. Hipotesis pengujian yang digunakan pada Uji Bartlett adalah sebagai berikut:

$H_0: \rho = I$ artinya tidak terdapat hubungan antar variabel

$H_1: \rho \neq I$ artinya terdapat hubungan antar variabel

Nilai statistik Uji Bartlett dapat ditulis dengan persamaan sebagai berikut:

$$\chi^2 = -(n-1 - \frac{2p+5}{6}) \ln|R| \quad (1)$$

Keterangan:

$\ln|R|$: nilai determinan dari matriks korelasi

n : banyaknya observasi

p : banyaknya variabel

Tolak H_0 jika $\chi^2 > \chi^2_{(1/2p)(p-1), \alpha}$ atau jika $p\text{-value} < \alpha$, yang berarti bahwa antara variabel saling berkorelasi. Hal ini menunjukkan adanya hubungan antar variabel.

B. Analisis Faktor

Analisis faktor merupakan salah satu metode multivariat yang digunakan untuk menganalisis variabel-variabel yang diduga memiliki keterkaitan satu dengan yang lain sehingga keterkaitan tersebut dapat dijelaskan dan dipetakan atau dike-

lompokkan pada faktor yang tepat [4]. Persamaan model analisis faktor dapat ditulis sebagai berikut:

$$\begin{aligned} X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\ X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2 \\ &\vdots \\ X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p \end{aligned} \quad (2)$$

Keterangan:

F_k = common faktor ke- k

ℓ_{hk} = loading faktor ke- k dan variabel ke- h

μ_h = rata-rata variabel ke- h

ε_h = spesifik faktor ke- h

dengan

$i = 1, 2, \dots, n$ adalah banyak observasi

$k = 1, 2, \dots, m$ adalah banyaknya common faktor

$h = 1, 2, \dots, p$ adalah banyaknya variabel

C. Jaringan Syaraf Tiruan

Jaringan syaraf tiruan didefinisikan sebagai suatu sistem pemrosesan informasi yang mempunyai karakteristik menyerupai jaringan saraf manusia [5]. Pada jaringan saraf tiruan terdapat dua metode pelatihan, yaitu pelatihan *supervised* dan *unsupervised*. Metode pelatihan *supervised* mengasumsikan tersedianya target yang mengklasifikasikan contoh-contoh pelatihan ke dalam kelas-kelas. Metode pelatihan *unsupervised* menggunakan contoh yang tidak diklasifikasikan jenisnya, sehingga sistem akan melakukan pemrosesan secara *heuristically*. Seperti jaringan saraf manusia, jaringan saraf tiruan juga terdiri dari beberapa *neuron* dan pada setiap *neuron* terdapat hubungan. *Neuron* akan mentransformasikan informasi yang diterimanya melalui sambungan keluaran menuju *neuron-neuron* yang lain. Masing-masing sambungan memiliki kekuatan hubungan terkait (bobot). Bobot yang diberikan pada setiap *input* didasarkan pada besarnya kekuatan hubungan [6].

D. Kohonen SOM

Jaringan Kohonen SOM banyak digunakan untuk membagi pola masukan ke dalam beberapa kelompok (*cluster*). Dimisalkan masukan berupa vektor yang terdiri dari n komponen yang akan dikelompokkan ke dalam maksimum m buah kelompok yang disebut sebagai vektor contoh. Keluaran jaringan adalah kelompok yang paling dekat atau memiliki kemiripan dengan masukan yang diberikan. Ukuran kedekatan yang dipakai adalah jarak *Euclidean* yang paling minimum. Bobot vektor-vektor contoh berfungsi sebagai penentu kedekatan vektor contoh tersebut dengan masukan yang diberikan [7].

Algoritma yang digunakan untuk melakukan pengelompokan pada jaringan Kohonen SOM adalah sebagai berikut:

0. Inisialisasi yang terdiri dari menetapkan inisialisasi bobot w_{ij} , menetapkan nilai *learning rate* (α) dan menentukan nilai radius R
1. Melakukan langkah 2 sampai 8 ketika kondisi untuk berhenti tidak terpenuhi
2. Melakukan langkah 3 sampai 5 untuk masing-masing vektor *input* x_i

3. Menghitung jarak *Euclidean*, yaitu jarak antara bobot w_{ij} dengan vektor *input* x_i untuk setiap kluster j menggunakan persamaan:

$$D_j = \sum_i (w_{ij} - x_i)^2 \quad (3)$$

Keterangan:

D_j = jarak *Euclidean*

w_{ij} = bobot yang menghubungkan antara vektor *input* x_i menuju ke unit y_j

4. Menentukan nilai J yang meminimumkan nilai D_j
5. Mengubah nilai bobot w_{ij} pada *input* pemenang menggunakan persamaan:

$$w_{ij}(\text{new}) = w_{ij}(\text{old}) + \alpha [x_i - w_{ij}(\text{old})] \quad (4)$$

Keterangan:

$w_{ij}(\text{new})$ = bobot w_{ij} yang baru

$w_{ij}(\text{old})$ = bobot w_{ij} awal

α = nilai *learning rate*

x_i = vektor *input* ke- i

6. Memperbarui nilai *learning rate*
7. Mereduksi nilai radius R
8. Menguji kondisi penghentian iterasi.

E. Metode K-means

K-Means Cluster Analysis merupakan salah satu metode analisis kluster non hirarki yang dapat digunakan untuk mempartisi objek ke dalam kelompok-kelompok berdasarkan kedekatan karakteristik, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu kluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kluster yang lain [4]. Secara umum metode *K-Means Cluster Analysis* menggunakan algoritma sebagai berikut:

1. Menentukan K sebagai jumlah kluster yang dibentuk.
2. Membangkitkan K *centroid* (titik pusat kluster) awal secara *random* kemudian untuk menghitung *centroid* kluster ke- i berikutnya, digunakan rumus sebagai berikut:

$$y_i = \frac{\sum_{i=1}^n x_i}{n} \quad (5)$$

Keterangan:

y = *centroid* pada *cluster*

x_i = obyek pengamatan ke i

n = banyaknya obyek yang menjadi anggota kluster.

3. Menghitung jarak setiap objek ke masing-masing *centroid* pada setiap kluster dengan persamaan:

$$d = \sum_{i=1}^n (x_i - y_i)^2 \quad (6)$$

Keterangan:

x_i = obyek pengamatan ke i

y_i = *centroid* ke i

n = banyaknya obyek yang menjadi anggota kluster

4. Mengalokasikan masing-masing objek ke dalam *centroid* yang paling terdekat.
5. Melakukan iterasi, kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (6)

6. Mengulangi langkah 3 jika posisi *centroid* baru tidak sama.

F. Pemilihan Jumlah Kluster Optimum

Untuk mengetahui jumlah kelompok optimum dapat digunakan kriteria nilai *Pseudo Fstatistics*. Berikut adalah persamaan yang digunakan untuk menghitung nilai *Pseudo Fstatistics*:

$$Pseudo\ Fstatistics = \frac{\left(\frac{R^2}{k-1}\right)}{\left(\frac{1-R^2}{n-k}\right)} \tag{7}$$

Dimana:

$$R^2 = \frac{(SST - SSW)}{SST} \tag{8}$$

$$SST = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (x_{ijk} - \bar{x}_j)^2 \tag{9}$$

$$SSW = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (x_{ijk} - \bar{x}_{jk})^2 \tag{10}$$

Keterangan:

SST (*Sum Square Total*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

SSW (*Sum Square Within*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata kelompoknya

n : banyaknya sampel

c : banyaknya variabel

p : banyaknya kelompok

x_{ijk} : sampel ke-*i* pada variabel ke-*j* kelompok ke-*k*

x̄_j : rata-rata seluruh sampel pada variabel ke-*j*

x̄_{jk} : rata-rata sampel pada variabel ke-*j* dan kelompok ke-*k*

Nilai *Pseudo Fstatistics* tertinggi menunjukkan bahwa jumlah kelompok yang digunakan untuk memartisi data telah optimal [8].

G. Kriteria Perbandingan Kluster Terbaik

Untuk menentukan kriteria kluster terbaik dapat digunakan nilai *icdrate*. *Icdrate* (*internal cluster dispersion*) merupakan tingkat dispersi dalam kluster [9]. Nilai *icdrate* dapat diperoleh dengan menggunakan persamaan:

$$icdrate = 1 - \frac{SSB}{SST} \tag{11}$$

Dimana:

$$SSB = \sum_{j=1}^c \sum_{k=1}^p (\bar{x}_{jk} - \bar{x}_j)^2 \tag{12}$$

Keterangan:

SST (*Sum Square Total*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

x̄_{jk} : rata-rata sampel pada variabel ke-*j* dan kelompok ke-*k*

x̄_j : rata-rata sampel pada variabel ke-*j*

c : banyaknya variabel

p : banyaknya kelompok

H. Uji One-way MANOVA

Pengujian *One-way MANOVA* digunakan untuk membandingkan nilai vektor *mean* antar perlakuan pada data multivariat. Dimana pada pengujian *One-way MANOVA* hanya menggunakan satu faktor atau perlakuan dan tanpa mempertimbangkan interaksi antar perlakuan [4]. Berikut adalah hipotesis yang digunakan dalam pengujian:

$$H_0: \tau_1 = \tau_2 = \tau_3 = \dots = \tau_g = 0$$

$$H_1: \text{minimal ada 1 } \tau_l \neq 0, \text{ dengan } l = 1, 2, \dots, g$$

Statistik uji:

$$F = \frac{SS_{treatment} / (g - 1)}{SS_{residual} / \left(\sum_{l=1}^g n_l - g\right)} \tag{13}$$

Keterangan:

SS_{treatment} = nilai *Sum of Square* perlakuan

SS_{residual} = nilai *Sum of Square* residual

n_l = banyaknya data populasi ke *l*

g = banyaknya populasi

Tolak *H₀* jika $F > F_{\alpha; v_1, v_2}$, dengan $v_1 = g - 1$ dan $v_2 = \sum n_l - g$.

I. Indikator Kesehatan Masyarakat

Indikator kesehatan dapat digunakan untuk mengetahui kondisi kesehatan masyarakat di suatu wilayah. Dimana terdapat beberapa pendapat tentang variabel yang membentuk Indikator Kesehatan Masyarakat. Menurut Indeks Pembangunan Kesehatan Manusia (IPKM), terdapat 24 indikator kesehatan yang digunakan dalam sebagai ukuran tingkat pembangunan kesehatan manusia disuatu wilayah. Sedangkan menurut *World Health Organization*, Indikator Kesehatan Masyarakat didasarkan pada hal-hal yang berhubungan dengan status kesehatan masyarakat dan berhubungan dengan pelayanan kesehatan. Departemen Kesehatan Republik Indonesia menyebutkan bahwa Indikator Kesehatan Masyarakat menurut Indonesia Sehat 2010 terdiri dari tiga indikator, yaitu Indikator Hasil Akhir, Indikator Hasil Antara serta Indikator Proses dan Masukan [10].

Pada penelitian ini variabel pembentuk Indikator Kesehatan Masyarakat terdiri dari variabel ekonomi, pendidikan dan epidemiologi [1]. Hasil penelitian yang dilakukan oleh Gerring, et al pada tahun 2013 menyebutkan bahwa variabel ekonomi, pendidikan dan epidemiologi dapat digunakan untuk mengetahui kondisi kesehatan masyarakat di suatu wilayah berdasarkan karakteristik wilayah tersebut.

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan pada penelitian ini merupakan data sekunder tentang variabel pembentuk Indikator Kesehatan Masyarakat di 38 kabupaten/kota yang diperoleh dari hasil Riskesdas dan Badan Pusat Statistik Provinsi Jawa Timur pada tahun 2012.

B. Variabel Penelitian

Varibel penelitian yang digunakan merupakan variabel pembentuk Indikator Kesehatan Masyarakat, yang meliputi

variabel ekonomi, pendidikan dan epidemiologi [1] dengan rincian sebagai berikut:

Tabel 1.
Variabel Penelitian

Variabel	Nama Variabel
X ₁	PRDB (Pendapatan Regional Domestik Bruto) per kapita atas dasar harga konstan
X ₂	Angka Harapan Hidup
X ₃	Angka Melek Huruf
X ₄	Rata-rata total lama sekolah penduduk usia 15 tahun
X ₅	Prevalensi TBC (Tuberkulosis)
X ₆	Prevalensi malaria

C. Langkah Analisis

Langkah analisis yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Mendeskripsikan karakteristik masing-masing indikator kesehatan masyarakat berdasarkan kabupaten/kota di Jawa Timur.
2. Mereduksi variabel penelitian
3. Mendapatkan hasil pengelompokan kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM
4. Mendapatkan hasil pengelompokan kabupaten/kota di Jawa Timur menggunakan metode *K-Means*
5. Memperoleh jumlah kelompok optimum pada masing-masing metode Kohonen SOM dan *K-Means*
6. Memperoleh hasil kelompok terbaik yang terbentuk berdasarkan kriteria nilai *icdrate*
7. Membuat kesimpulan.

IV. ANALISIS DAN PEMBAHASAN

Pada bab berikut ini akan dilakukan analisis dan pembahasan tentang indikator kesehatan masyarakat dan variabel yang membentuknya di Jawa Timur pada tahun 2012.

A. Karakteristik Indikator Kesehatan Masyarakat di Jawa Timur

Sebelum menentukan jumlah kelompok optimum pada masing-masing metode Kohonen SOM dan *K-Means*, terlebih dahulu akan dilakukan deskripsi data tentang variabel yang membentuk Indikator Kesehatan Masyarakat di Jawa Timur pada tahun 2012. Rentang nilai maksimum dan minimum pada variabel PDRB per kapita di Jawa Timur yang cukup besar mengindikasikan bahwa pertumbuhan ekonomi di Jawa Timur masih belum merata pada masing-masing kabupaten dan kota. Nilai rata-rata Angka Harapan Hidup di Jawa Timur pada tahun 2012 adalah sebesar 68,76 tahun. Nilai rata-rata tersebut masih berada di bawah rata-rata Angka Harapan Hidup nasional yaitu sebesar 69,87 tahun. Rata-rata Angka Melek Huruf sebesar 90,03 yang masih berada di bawah rata-rata Angka Melek Huruf nasional di tahun 2012 yaitu sebesar 93,25.

Nilai rata-rata lama sekolah terbesar di Jawa Timur dimiliki oleh Kota Malang yaitu sebesar 10,87 tahun sedangkan nilai rata-rata terendah dimiliki oleh Kabupaten Sampang yaitu sebesar 4,22 tahun. Nilai rata-rata prevalensi TBC di Jawa Timur adalah sebesar 0,2263%. Nilai tersebut berarti bahwa terdapat 0,2263% dari seluruh penduduk di Jawa Timur yang

menderita penyakit TBC. Rata-rata prevalensi malaria di Jawa Timur mencapai 0,4895%. Nilai rata-rata tersebut masih berada di bawah rata-rata prevalensi malaria nasional yaitu sebesar 0,6%.

B. Mereduksi Variabel Penelitian

Reduksi variabel bertujuan untuk menggabungkan variabel penelitian yang memiliki korelasi atau saling berhubungan. Untuk melakukan reduksi variabel terlebih dahulu dilakukan pengujian Bartlett untuk mengetahui ada atau tidaknya hubungan diantara variabel penelitian. Hasil pengujian Bartlett diperoleh nilai χ^2 hitung sebesar 111,572 dengan derajat bebas 15. Nilai tersebut lebih besar dibandingkan dengan nilai χ^2 tabel yaitu sebesar 24,996. Hasil tersebut dapat digunakan untuk mengambil keputusan Tolak H₀, yang berarti bahwa terdapat hubungan atau korelasi di antara variabel penelitian. Untuk mengatasi hal tersebut dapat dilakukan reduksi variabel dengan menggunakan analisis faktor.

Berikut ini adalah *eigen value* yang dapat digunakan untuk menentukan banyaknya faktor baru yang terbentuk:

Tabel 2.
Nilai *Eigen Value* Analisis Faktor

Banyak Komponen	Total	<i>Eigen Value</i>	
		% Variasi	% Kumulatif
1	2,796	46,605	46,605
2	1,317	21,951	68,556
3	0,897	14,946	83,502
4	0,691	11,511	95,013
5	0,225	3,746	98,759
6	0,074	1,241	100

Jumlah faktor baru yang terbentuk dapat didasarkan pada banyaknya komponen yang memiliki nilai *eigen value* atau λ yang lebih besar dari nilai 1. Variabel akan direduksi menjadi 3 faktor yang sehingga mampu menjelaskan sebesar 83,502% dari keragaman total. Nilai *loading factor* dapat digunakan untuk menentukan pengelompokan variabel menjadi beberapa faktor yang baru. Untuk lebih memudahkan reduksi variabel dan penamaan faktor baru maka dilakukan rotasi pada nilai *loading factor* dengan hasil sebagai berikut:

Tabel 3.
Rotasi Nilai *Loading Factor* untuk Reduksi Variabel

Variabel	Komponen		
	1	2	3
PDRB per Kapita (X ₁)	0,397	0,680	-0,034
Angka Harapan Hidup (X ₂)	0,856	-0,107	0,303
Angka Melek Huruf (X ₃)	0,949	0,046	0,055
Rata-rata Lama Sekolah (X ₄)	0,948	0,076	-0,126
Prevalensi TBC (X ₅)	-0,247	0,824	-0,069
Prevalensi Malaria (X ₆)	0,064	-0,073	0,986

Setelah dilakukan rotasi pada nilai *loading factor* diperoleh hasil bahwa variabel Angka Harapan Hidup (X₂), Angka Melek Huruf (X₃) dan Rata-rata Lama Sekolah (X₄) dikelompokkan menjadi 1 faktor dan nama faktor pembentuk IPM (Indeks Pembangunan Manusia). Variabel prevalensi malaria (X₆)

diletakkan ke dalam faktor 3 dengan nama prevalensi malaria sebab hanya terdapat satu variabel pada faktor tersebut. Variabel PDRB per Kapita (X_1) dan prevalensi TBC (X_5) diletakkan ke dalam faktor 2 dengan nama variabel ekonomi.

Pada analisis faktor akan diperoleh nilai *score factor* yang merupakan kombinasi antara variabel yang direduksi ke dalam beberapa faktor. Pada penelitian ini enam variabel direduksi menjadi tiga faktor, sehingga terdapat tiga *score factor* yang akan digunakan sebagai nilai pengganti enam variabel.

C. Pemilihan Klaster Terbaik

Pengelompokan kabupaten/kota menggunakan metode Kohonen SOM dan *K-Means* masing-masing akan dibentuk menjadi 2 klaster dan 3 klaster. pengelompokan menggunakan metode Kohonen SOM akan digunakan tiga jenis topologi, Berikut adalah nilai *Pseudo Fstatistics* pada masing-masing kelompok dan jenis topologi metode Kohonen SOM:

Tabel 4.

Nilai <i>Pseudo Fstatistics</i> untuk Metode Kohonen SOM			
Jumlah Neuron	Jenis Topologi		
	<i>Gridtop</i>	<i>Hextop</i>	<i>Randtop</i>
2 klaster	11,054	11,967	11,928
3 klaster	13,814	13,819	13,819

Pada tabel 4 terlihat bahwa nilai *Pseudo Fstatistics* pada metode Kohonen SOM tertinggi adalah ketika digunakan jumlah kelompok sebanyak 3 klaster dengan jenis topologi *hextop*. Selanjutnya dilakukan pengelompokan menggunakan metode *K-Means*. Nilai *Pseudo Fstatistics* ketika digunakan jumlah kelompok 2 klaster lebih besar dibandingkan dengan ketika digunakan jumlah kelompok sebanyak 3 klaster. Berdasarkan hal tersebut maka pengelompokan menggunakan metode *K-Means* akan dilakukan dengan jumlah kelompok sebanyak 2 klaster.

Langkah selanjutnya adalah menentukan hasil kelompok terbaik berdasarkan kriteria nilai *icdrate*. Berikut ini adalah nilai *icdrate*, SSW (*Sum Square Within*) dan SSB (*Sum Square Between*) hasil pengelompokan menggunakan metode Kohonen SOM dan *K-Means*.

Tabel 5.

Nilai *Icdrate*, SSW dan SSB untuk Metode Kohonen SOM dan *K-Means*

Metode	Jumlah Kelompok	<i>icdrate</i>	SSW	SSB
Kohonen SOM	3 klaster	0,962	62,023	4,261
<i>K-Means</i>	2 klaster	0,988	87,283	1,312

Pengelompokan kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM dengan membagi data menjadi 3 klaster memiliki nilai *icdrate* yang lebih kecil dibandingkan dengan mengelompokan kabupaten/kota di Jawa Timur menjadi 2 klaster menggunakan metode *K-Means*. Berdasarkan hal tersebut dalam disimpulkan bahwa hasil klaster terbaik adalah menggunakan metode Kohonen SOM dengan jumlah kelompok sebanyak 3 klaster.

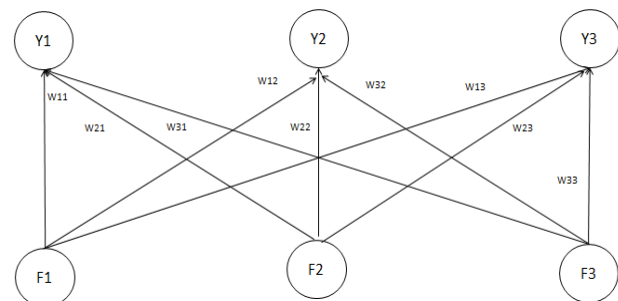
Berdasarkan tabel 5 diketahui bahwa mengelompokan kabupaten/kota di Jawa Timur menjadi 3 klaster menggunakan Kohonen SOM memiliki nilai SSW yang lebih kecil dibandingkan dengan metode *K-Means*. Selain itu nilai SSB metode Kohonen SOM lebih besar dibandingkan dengan nilai SSB

metode *K-Means*. Kedua hal tersebut juga menunjukkan bahwa hasil klaster terbaik adalah menggunakan metode Kohonen SOM dengan jumlah kelompok sebanyak 3 klaster.

D. Pengelompokan Menggunakan Metode Terbaik

Berdasarkan kriteria nilai *Pseudo Fstatistics* dan nilai *icdrate* diketahui bahwa hasil klaster terbaik merupakan pengelompokan menggunakan metode Kohonen SOM dengan topologi *hextop* dan jumlah neuron sebanyak 3. Jumlah neuron pada jaringan Kohonen SOM menunjukkan banyaknya kelompok yang terbentuk.

Terdapat 3 vektor *input*, dimana 3 faktor tersebut merupakan hasil reduksi variabel yang mewakili 6 variabel awal. Masing-masing vektor *input i* yang menuju neuron ke *j* dihubungkan dengan bobot. Bobot tersebut merupakan nilai *centroid* atau pusat klaster. Berikut adalah arsitektur jaringan Kohonen SOM menggunakan 3 neuron:



Gambar 1. Arsitektur Jaringan Kohonen SOM Menggunakan 3 Neuron

Dengan menggunakan jumlah kelompok sebanyak 3 klaster hasil pengelompokan kabupaten/kota di Jawa Timur berdasarkan Indikator Kesehatan Masyarakat adalah sebagai berikut klaster 1 terdiri dari 12 kabupaten yang meliputi Kabupaten Jember, Bondowoso, Situbondo, Probolinggo, Pasuruan, Bojonegoro, Tuban, Lamongan, Bangkalan, Sampang, Pamekasan, Sumenep. Klaster 2 terdiri dari 17 kabupaten/kota yang meliputi Ponorogo, Malang, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Batu. Serta klaster 3 yang terdiri dari 5 kabupaten/kota yaitu Pacitan, Trenggalek, Tulungagung, Blitar, Kediri, Lumajang, Banyuwangi, Gresik, Kota Surabaya.

Berikut ini adalah nilai rata-rata untuk setiap variabel pada masing-masing kelompok yang terbentuk:

Tabel 6.

Nilai Rata-rata Hasil Pengelompokan Klaster Terbaik

Variabel	Kelompok 1	Kelompok 2	Kelompok 3	Rata-rata Populasi
PDRB	6315,70	8356,6	19538,5	10360
Angka Harapan Hidup	64,88	70,5	70,7	68,763
Angka Melek Huruf	82,40	94,1	92,6	90,03
Lama Sekolah	6,21	8,7	7,8	7,695
Prevalensi TBC	0,28	0,17	0,27	0,2263
Prevalensi Malaria	0,38	0,39	0,81	0,4895

Nilai rata-rata populasi merupakan nilai rata-rata setiap variabel secara keseluruhan. Untuk mengetahui karakteristik dari klaster yang terbentuk dapat digunakan nilai rata-rata dari masing-masing variabel pada setiap klaster yang kemudian dibandingkan dengan nilai rata-rata setiap variabel secara keseluruhan. Klaster 1 merupakan klaster yang memiliki kondisi ekonomi dan pendidikan yang kurang baik, namun memiliki kondisi epidemiologi yang cukup baik. Klaster 2 memiliki kondisi ekonomi yang cukup baik namun memiliki kondisi pendidikan dan epidemiologi yang baik. Klaster 3 merupakan kelompok yang memiliki kondisi ekonomi dan pendidikan yang baik namun memiliki kondisi epidemiologi yang kurang baik.

Hasil kelompok kabupaten/kota di Jawa Timur menggunakan metode Kohonen SOM diharapkan memiliki perbedaan karakteristik antar kelompok yang terbentuk. Untuk membuktikan hal tersebut dapat dilakukan pengujian MANOVA (*Multivariate Analysis of Variance*) untuk mengetahui ada atau tidaknya perbedaan pada setiap klaster yang terbentuk. Pengujian MANOVA yang akan dilakukan merupakan pengujian *One-way* MANOVA dimana hanya terdapat satu faktor atau perlakuan dan tidak terdapat interaksi antar perlakuan yang diduga mempengaruhi variabel respon, yaitu kelompok yang terbentuk. Dimana satu faktor tersebut terdiri dari tiga kategori, yaitu kelompok 1, kelompok 2 dan kelompok 3. Hasil pengujian *One-way* MANOVA disajikan pada Tabel 7 sebagai berikut:

Tabel 7. Hasil Pengujian *One-way* MANOVA pada Klaster

Variabel	<i>P-value</i>
PDRB	0,194
Angka Harapan Hidup	0,000
Angka Melek Huruf	0,000
Lama Sekolah	0,000
Prevalensi TBC	0,092
Prevalensi Malaria	0,000

Pada hasil pengujian *One-way* MANOVA diperoleh nilai *p-value* untuk variabel Angka Harapan Hidup, Angka Melek Huruf, lama sekolah dan prevalensi malaria masing-masing sebesar 0,000. Nilai tersebut lebih kecil dibandingkan dengan nilai $\alpha = 0,05$ sehingga Tolak H_0 yang berarti bahwa pada masing-masing klaster 1, klaster 2 dan klaster 3 yang terbentuk memiliki perbedaan. Sedangkan variabel PDRB per kapita dan prevalensi TBC memiliki nilai *p-value* sebesar 0,194 dan 0,092. Kedua nilai tersebut lebih besar dibandingkan dengan nilai $\alpha = 0,05$ sehingga Gagal Tolak H_0 , yang artinya variabel PDRB per kapita dan prevalensi TBC pada masing-masing klaster yang terbentuk adalah tidak memiliki perbedaan. Pada hasil pengujian *One-way* MANOVA diperoleh kesimpulan bahwa hasil pengelompokan memiliki perbedaan nilai pada masing-masing variabel respon yang membentuknya. Meskipun terdapat dua variabel yang tidak memiliki perbedaan pada setiap klaster yang terbentuk, yaitu PDRB per kapita dan prevalensi TBC. Namun ketika digunakan nilai $\alpha = 0,1$ maka variabel prevalensi TBC sebab nilai *p-value* pada variabel TBC sebesar 0,092 lebih kecil dibandingkan dengan nilai α . Sehingga dapat disimpulkan bahwa variabel prevalensi TBC memiliki perbedaan pada masing-masing klaster yang terbentuk.

V. KESIMPULAN DAN SARAN

Untuk meningkatkan kondisi kesehatan masyarakat di Jawa Timur, monitoring dan evaluasi dapat dilakukan oleh Pemerintah Jawa Timur berdasarkan klaster yang terbentuk. Pada klaster 1 yang dapat disebut sebagai klaster yang memiliki kondisi kesehatan masyarakat yang paling rendah, pemerintah perlu memperbaiki kondisi ekonomi, pendidikan dan epidemiologi. Sedangkan pada klaster 3 yang dapat disebut sebagai klaster dengan kondisi ekonomi yang sedang, pemerintah perlu memperbaiki kondisi epidemiologi yang bertujuan untuk mengurangi angka prevalensi penyakit TBC dan malaria. Selain itu pemerintah juga perlu mempertahankan kondisi ekonomi dan pendidikan yang berada pada klaster 3 sebab klaster 3 telah memiliki kondisi ekonomi dan pendidikan yang baik. Klaster 2 dapat disebut sebagai kelompok yang memiliki kondisi kesehatan masyarakat yang tinggi. Dimana hal tersebut tercermin pada kondisi ekonomi, pendidikan dan epidemiologi pada klaster 2 yang telah baik. Meskipun demikian, pemerintah perlu menjaga kondisi kesehatan masyarakat pada klaster 2 dan juga meningkatkan nilai PDRB per kapita pada klaster 2.

Saran yang dapat diberikan kepada Pemerintah Provinsi Jawa Timur dan Dinas Kesehatan Provinsi Jawa Timur adalah hasil penelitian ini dapat digunakan sebagai informasi untuk meningkatkan kondisi kesehatan masyarakat di Jawa Timur dalam rangka mencapai Program Kota Sehat Nasional Pemerintah perlu memperhatikan variabel ekonomi dan pendidikan. Kondisi ekonomi dan pendidikan di Jawa Timur yang semakin baik juga akan menyebabkan kondisi kesehatan masyarakat yang lebih baik. Saran untuk penelitian selanjutnya adalah menggunakan variabel penelitian yang mempresentasikan kondisi kesehatan masyarakat di Jawa Timur secara menyeluruh sehingga hasil pengelompokan yang dihasilkan juga optimal.

DAFTAR PUSTAKA

- [1] Gerring, J., Thacker, Strom. C., Enikolopov, R., Julian, A., and Maguire, M. (2013). Assessing Health System Performance: A Model-based Approach. *Social Science & Medicine* 93, 21-28.
- [2] Ettaouil, M., Abdelatifi, E., Belhabib, F., Moutaouakil, Karim El. (2012). Learning Algorithm of Kohonen Network with Selection Phase. *WSEAS Transactions on Computers Issue 11, Volume 11, November 2012. E-ISSN: 2224-2872*.
- [3] Morrison, Donald. F. (1967). *Multivariate Statistical Methods Third Edition*. USA: McGRAW-Hill Book Company.
- [4] Johnson. R. and Wichern. D. 2007. *Applied Multivariate Statistical Analysis. 6th Edition*. New Jersey: Prentice Hall.
- [5] Siang, Jong Jek. (2009). *Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan MATLAB*. Yogyakarta: ANDI.
- [6] Kusumadewi, S. (2004). *Membangun Jaringan Syaraf Tiruan (menggunakan MATLAB & Excel Link)*. Yogyakarta: ANDI.
- [7] Fausett, Laurence. (1994). *Fundamental Neural Network: Architectures, Algorithm, and Applications*. New Jersey: Prentice Hall Inc.
- [8] Orpin, A. R., & Kostylev, V. E. (2006). *Toward a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats. Marine Geology*, 209-222.
- [9] Mingoti, S.A. & Lima, J.O. (2006). *Comparing SOM Neural Network with Fuzzy C-Means, K-Means and Traditional Hierarchical Clustering Algorithm*: 1742-1759.
- [10] Badan Pusat Statistik. (2011). *Kajian Indikator Kesehatan (Laporan Sosial 2010)*. Jakarta: Badan Pusat Statistik.