

METODE K-MEANS CLUSTER DAN FUZZY C-MEANS CLUSTER

(Studi Kasus: Indeks Pembangunan Manusia di Kawasan Indonesia Timur tahun 2012)

Annas Syaiful Rizal¹ dan R.B Fajriya Hakim²

¹ Mahasiswa Program Studi Statistika, FMIPA Universitas Islam Indonesia Yogyakarta
email: annasipoel@gmail.com

² Dosen Program Studi Statistika, FMIPA Universitas Islam Indonesia Yogyakarta
email: hakimf@fmipa.uui.ac.id

ABSTRAK. Analisis kelompok (cluster analysis) telah digunakan diberbagai bidang ilmu pengetahuan, dengan tujuan mengelompokkan objek/observasi. Hal penting dalam analisis kelompok adalah memperoleh nilai simpangan baku dalam kelompok (S_w) yang minimum dan nilai simpangan baku antar kelompok (S_b) yang maksimum, untuk jumlah kelompok yang optimum didapatkan dengan menggunakan kriteria ukuran validitas, yaitu Indeks Xie dan Beni. Pada penelitian ini dilakukan pengelompokkan variabel pembentuk Indeks Pembangunan Manusia di Kawasan Indonesia Timur pada tahun 2012 dengan menggunakan analisis clustering, yaitu metode k-means dan fuzzy c-means cluster. Variabel-variabel pembentuk IPM, terdiri dari angka harapan hidup; angka melek huruf; rata-rata lamanya sekolah, dan *Purchasing Power Parity* (paritas daya beli). Dari hasil analisis didapatkan fuzzy c means menghasilkan nilai S_w/S_b lebih kecil dibandingkan metode c-means. Pada FCM jumlah kelompok optimum sebanyak enam kelompok dengan nilai Indeks Xie dan Beni 0,0006589. Kelompok 1 terdiri dari 32 kabupaten/kota, kelompok 2 terdiri dari 18 kabupaten/kota, kelompok 3 terdiri dari 26 kabupaten/kota, kelompok 4 terdiri dari 50 kabupaten/kota. Kelompok 5 terdiri dari 49 kabupaten/kota, dan kelompok 6 terdiri dari 44 kabupaten/kota.

Kata kunci: *K-Means cluster; Fuzzy C-Means cluster; Indeks Xie dan Beni; Indeks Pembangunan Manusia*

1. PENDAHULUAN

Tolok ukur keberhasilan pembangunan manusia telah dikembangkan oleh United Nation Development Programme (UNDP) yang dikenal dengan istilah Human Development Index (HDI) atau dikenal dengan Indeks Pembangunan Manusia (IPM) [6]. Indeks Pembangunan Manusia (IPM) / Human Development Index (HDI) adalah pengukuran perbandingan dari harapan hidup, melek huruf, pendidikan dan standar hidup untuk semua negara seluruh dunia. IPM digunakan untuk mengklasifikasikan apakah sebuah negara adalah negara maju, negara berkembang atau negara terbelakang dan juga untuk mengukur pengaruh dari kebijaksanaan ekonomi terhadap kualitas hidup.

Sejak 1 Januari 2001 pemerintah pusat memberikan kewenangan otonomi kepada pemerintah daerah. Masing-masing pemerintah daerah menyusun perencanaan pembangunan dan anggaran keuangan daerahnya, bukan hanya untuk meningkatkan pertumbuhan ekonomi, tetapi juga meningkatkan kesejahteraan masyarakat melalui pembangunan manusia yang meliputi sektor pendidikan, pelayanan kesehatan, dan kebijakan-kebijakan lain yang secara langsung memperbaiki kualitas hidup [18]. Dengan demikian, daerah dapat meningkatkan kesejahteraan masyarakat melalui peningkatan pembangunan manusia yang

tercermin dari indikator Indeks Pembangunan Manusia (IPM) [9] termasuk di Kawasan Timur Indonesia (KTI).

Berdasarkan data Human Development Report (HDR) menunjukkan bahwa **Indeks Pembangunan Indonesia tahun 2011 mengalami kemerosotan secara drastis, yaitu berada di peringkat 124.** Padahal HDR 2010 menunjukkan bahwa Indonesia berada di peringkat ke 108. Perubahan Peringkat menjadi ke 124 ini menunjukkan bahwa pembangunan manusia di Indonesia mengalami perlambatan dibandingkan negara-negara lain. Derajat kesejahteraan masyarakat Indonesia mengalami penurunan secara drastic, hal ini ditunjukkan dari usia harapan hidup (*Life expectancy at birth*). HDR 2010, menunjukkan usia harapan hidup masyarakat Indonesia adalah 71,5 tahun, sedangkan HDR 2011 menunjukkan *Life expectancy* masyarakat Indonesia di usia 69,4 tahun. Dengan menempati peringkat ke-124, pembangunan Indonesia tertinggal dibanding tetangganya. Di ASEAN, peringkat pertama dalam hal kualitas manusia adalah Singapura dengan nilai 0,866.

Kemudian disusul Brunei dengan nilai IPM 0,838, disusul Malaysia (0,761), Thailand (0,682,) dan Filipina (0,644). Sebagai negara besar di kawasan ASEAN, seharusnya Indonesia dapat melampaui setidaknya Malaysia atau Thailand. Apalagi kedepannya, Indonesia akan menghadapi masyarakat ekonomi ASEAN. Bila tidak segera dibenahi maka tak mustahil manusia Indonesia sulit bersaing di regional [2].

Menurut Kepres Indonesia Nomor 55 Tahun 2001[13] Tentang Dewan Pengembangan Kawasan Indonesia Timur provinsi yang termasuk Kawasan Indonesia Timur adalah Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Nusa Tenggara Barat, Nusa Tenggara Timur, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku Utara, Maluku, Propinsi di Irian Jaya.

Berdasarkan data BPS [7], Indeks Pembangunan Manusia di Kawasan Indonesia Timur menempatkan 5 Provinsi dengan prosentse Indeks Pembangunan Manusia terendah yaitu provinsi Maluku Utara (69,47%), provinsi Papua Barat (69,65%), provinsi Nusa Tenggara Timur (67,75%), provinsi Nusa Tenggara Barat (66,23%), dan provinsi Papua (65,36%).

Tinggi rendahnya IPM Kabupaten/Kota di Indonesia hanya ditunjukkan Indeks Komposit tetapi tidak ditunjukkan variabel mana yang dominan terhadap tinggi/rendahnya peringkat IPM. Pengelompokan wilayah kabupaten/kota di Kawasan Indonesia Timur perlu dilakukan sebagai bahan perencanaan dan evaluasi sasaran program pemerintah untuk meningkatkan angka pembangunan manusia.

Pengelompokan wilayah bertujuan untuk membagi wilayah-wilayah dalam beberapa kelompok dengan karakteristik yang memiliki tingkat keserupaan yang tinggi di dalam setiap kelompok dan memiliki perbedaan antar kelompok [16]. Salah satu metode analisis pengelompokan wilayah adalah analisis cluster. Analisis Cluster (cluster analysis) adalah salah satu analisis peubah ganda (multivariate analysis) yang digunakan untuk mengelompokkan objek-objek sedemikian rupa sehingga objek dalam satu cluster yang sangat mirip dan objek di berbagai cluster cukup berbeda. Analisis cluster dapat dibagi menjadi kelompok berhirarku (hierarchical clustering) dan kelompok tak berhierarki (non hierarchical clustering). Contoh dari teknik kelompok berhierarki adalah single linkage, complete linkage, average linkage, median, dan Ward. Kelompok tak berhierarki adalah K-Means, Adaptive k-means, k-medoids, dan fuzzy clustering [11]. Pada penelitian ini dibahas metode analisis kelompok tak berhierarki (non hierarchical clustering methods). Metode tak berhierarki yang digunakan adalah C-Means Cluster dan Fuzzy C-Means Cluster. Dua metode ini cocok digunakan untuk data berukuran besar dan memiliki tipe peubah kontinu[15].

Fuzzy C-Means adalah sebuah metode *clustering* yang mengizinkan satu data menjadi milik dua atau lebih *cluster*. Metode ini sering digunakan dalam pengenalan pola (*pattern recognition*). Metode Fuzzy C-Means adalah salah satu metode *clustering* yang mengalokasikan kembali data ke dalam masing-masing *cluster* dengan memanfaatkan teori Fuzzy. Dalam metode Fuzzy C-Means dipergunakan variabel membership function i_{ku} , yang merujuk pada seberapa besar kemungkinan suatu data bisa menjadi anggota ke dalam suatu *cluster* [4].

Metode fuzzy c-means cluster sering digunakan dalam melakukan pengelompokan, karena metode ini memberikan hasil yang halus dan cukup efektif [14]. Menurut Sukim [15] Kelebihan dari Fuzzy C-Means Clustering dari metode konvensional adalah metode fuzzy

C-Means clustering lebih sederhana, mudah diimplementasikan, memiliki kemampuan untuk mengelompokkan data yang lebih besar, lebih kokoh terhadap data outlier.

Menurut Bunkers *et al* [5] untuk melihat kinerja kedua metode tersebut digunakan kriteria nilai simpangan baku, yaitu dalam kelompok (S_w) dan antar kelompok (S_b), metode yang akan dipilih nanti adalah metode yang mempunyai nilai S_w yang minimum dan nilai S_b yang maksimum. Kriteria banyak cluster optimum diberikan oleh indeks X_B yang minimum. Rekomendasi untuk menggunakan indeks X_B tertuang dalam penelitian Duo dkk [8] yang menyatakan bahwa indeks X_B memiliki ketepatan dan keandalan yang tinggi baik untuk memberikan banyak kelompok optimum pada metode hard partition seperti K-means cluster maupun pada FCM.

2. METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari situs Portal Data Indonesia (data.go.id). Variabel yang digunakan adalah indikator Indeks Pembangunan Manusia (IPM) 219 kabupaten/kota di Kawasan Indonesia Timur tahun 2012 yang meliputi Angka Harapan Hidup (AHH), Angka Melek Huruf (AMH), Rata-rata Lama Sekolah (RLS), dan Pengeluaran per Kapita yang disesuaikan (PPP)

2.2 K-Means

K-Means merupakan metode klasterisasi yang paling terkenal dan banyak digunakan di berbagai bidang karena sederhana, mudah diimplementasikan, memiliki kemampuan untuk mengklaster data yang besar, mampu menangani data outlier, dan kompleksitas waktunya linear $O(nKT)$ dengan n adalah jumlah dokumen, K adalah jumlah kluster, dan T adalah jumlah iterasi. K-means merupakan metode pengklasteran secara partitioning yang memisahkan data ke dalam kelompok yang berbeda. Dengan partitioning secara iteratif, KMeans mampu meminimalkan rata-rata jarak setiap data ke klasternya. Metode ini dikembangkan oleh Mac Queen pada tahun 1967.

- 1) Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk.
- 2) Bangkitkan k centroid (titik pusat klaster) awal secara random.
- 3) Hitung jarak setiap data ke masing-masing centroid menggunakan rumus korelasi antar dua objek yaitu Euclidean Distance dan kesamaan Cosine.
- 4) Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
- 5) Tentukan posisi centroid baru ($k C$) dengan cara menghitung nilai rata-rata dari data-data yang ada pada centroid yang sama.

$$C_k = \left(\frac{1}{n_k} \right) \sum d_i$$

Dimana $k n$ adalah jumlah dokumen dalam cluster k dan $i d$ adalah dokumen dalam cluster k .

- 6) Kembali ke langkah 3 jika posisi centroid baru dengan centroid lama tidak sama.

2.3 Fuzzy c-means (FCM)

Logika fuzzy adalah suatu cara yang tepat untuk memetakan suatu ruang input kedalam suatu ruang output, mempunyai nilai kontinyu. Fuzzy dinyatakan dalam derajat dari suatu keanggotaan dan derajat dari kebenaran. Oleh sebab itu sesuatu dapat dikatakan sebagian benar dan sebagian salah pada waktu yang sama [10].

Logika Fuzzy memungkinkan nilai keanggotaan antara 0 dan 1, tingkat keabuan dan juga hitam dan putih, dan dalam bentuk linguistik, konsep tidak pasti seperti "sedikit", "lumayan" dan "sangat" [19]. Kelebihan dari teori logika fuzzy adalah kemampuan dalam proses penalaran secara bahasa (linguistic reasoning). Sehingga dalam perancangannya tidak memerlukan persamaan matematik dari objek yang akan dikendalikan. Beberapa jurnal yang membahas tentang fuzzy c-means dapat dilihat dalam Ananthi Sheshasayee dan P. Sharmila [1], Babankumar S. Bansod dan O. P. Pandey [3], Z. Hilal Inana dan Mehmet Kuntalp [20].

Fuzzy C-Means clustering adalah suatu teknik clustering yang banyak digunakan dalam aplikasi clustering. Fuzzy C-Means menerapkan pengelompokan fuzzy, dimana setiap data dapat menjadi anggota dari beberapa cluster dengan derajat keanggotaan yang berbeda-beda pada setiap cluster. Fuzzy C-Means merupakan algoritma iteratif, yang menerapkan iterasi pada proses clustering data. Tujuan dari Fuzzy C-Means adalah untuk mendapatkan pusat cluster yang nantinya akan digunakan untuk mengetahui data yang masuk ke dalam sebuah cluster. Algoritma Fuzzy C-Means adalah sebagai berikut [3]:

- 1) Menentukan data yang akan dicluster X, berupa matriks berukuran $n \times m$ (n =jumlah sampel data, m =atribut setiap data). X_{ij} =data sampel ke- i ($i=1,2,\dots,n$), atribut ke- j ($j=1,2,\dots,m$).
- 2) Tentukan jumlah cluster (c), pangkat (w), maksimum iterasi (MaxIter), error terkecil yang diharapkan (ζ), fungsi obyektif awal ($P_0=0$), iterasi awal ($t=1$).
- 3) Bangkitkan bilangan random μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$; sebagai elemen-elemen matriks partisi awal U. Matriks partisi (U) pada pengelompokan fuzzy memenuhi kondisi sebagai berikut [9]:

$$\mu_{ik} \in [0,1]; \quad 1 \leq i \leq n; \quad 1 \leq k \leq c$$

μ_{ik} adalah derajat keanggotaan yang merujuk pada seberapa besar kemungkinan suatu data bisa

menjadi anggota ke dalam suatu cluster. Hitung jumlah setiap kolom (atribut):

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

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

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

- 5) Hitung fungsi obyektif pada iterasi ke- t , P_t :
Fungsi obyektif digunakan sebagai syarat perulangan untuk mendapatkan pusat cluster yang tepat. Sehingga diperoleh kecenderungan data untuk masuk ke cluster mana pada step akhir. Untuk iterasi awal nilai $t=1$.

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

- 6) Hitung perubahan matriks partisi:

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

- 7) Cek kondisi berhenti:
- $|Pt - Pt-1| < \zeta$) atau ($t > \text{MaxIter}$) maka berhenti;
 - Jika tidak, iterasi dinaikkan $t=t+1$, ulangi langkah ke-4.

2.4 Evaluasi Hasil Pengelompokkan

Tingkat keberhasilan usaha ditentukan berdasarkan penilaian kinerja kelima metode tersebut. Penilaian dapat dilakukan dengan membandingkan hasil pengelompokan oleh masing-masing metode dengan menggunakan kriteria dua nilai simpangan baku, yaitu rata-rata simpangan baku dalam kelompok (S_w) dan simpangan baku antar kelompok (S_b) [5].

Rumus rata-rata simpangan baku dalam kelompok:

$$S_w = K^{(-1)} \sum_{k=1}^k S_k$$

keterangan, K = banyaknya kelompok yang terbentuk dan S_k = Simpangan baku kelompok ke-k.

Rumus rata-rata simpangan baku antar kelompok:

$$S_b = \left[(K-1)^{-1} \sum_{k=1}^K (\bar{X}_k - \bar{X})^2 \right]^{1/2}$$

Keterangan, \bar{X}_k adalah rata-rata kelompok dan \bar{X} adalah rata-rata keseluruhan kelompok.

2.5 Indeks Xie dan Beni

Indeks XB ditemukan oleh Xie dan Beni yang pertama kali dikemukakan pada tahun 1991. Ukuran kevalidan cluster merupakan proses evaluasi hasil clustering untuk menentukan cluster mana yang terbaik. Ada dua kriteria dalam mengukur kevalidan suatu cluster, yaitu [17]:

- Compactness, yaitu ukuran kedekatan antar anggota pada tiap cluster.
- Separation, yaitu ukuran keterpisahan antar cluster satu dengan cluster yang lainnya.

Rumus kevalidan suatu cluster atau indeks Xie-Beni (XB) yaitu:

$$XB = \frac{\sum_{i=1}^c \sum_{j=1}^n \mu_{ik}^{w*} \|V_i - X_j\|^2}{n * \min_{i,j} \|V_i - V_j\|^2}$$

3. HASIL DAN PEMBAHASAN

3.1 Penentuan Metode Terbaik

Pengolahan data pada penelitian ini menggunakan bantuan program yaitu R i386 3.1.2 dengan *library e1071*. Ada dua pendekatan metode yang dilakukan dalam penelitian ini, yaitu dengan metode nonhierarchical clustering yaitu K Means dan Fuzzy C Means. Untuk penerapan metode fuzzy c means diawali dengan menentukan kelompok yang akan dilakukan dari banyak cluster, $k=2, \dots, 10$. Menentukan pangkat atau pembobot ke-fuzzy-an dalam penelitian ini nilai m yang digunakan adalah 2 karena berdasarkan penelitian Klawonn dan Keller [9] nilai dari m yang paling optimal dan sering dipakai adalah 2. Pemilihan jarak yang

menghasilkan kualitas pengelompokan terbaik dilakukan dengan memperhatikan nilai rasio rata-rata simpangan baku dalam kelompok dan simpangan baku antar kelompok yang minimum.

Tabel 1. Hasil Rasio Simpangan Baku

Metode	S_w/S_B				
	AHH	AMH	RLS	PPP	Average
	Jumlah cluster = 2				
K Means	17,8332	5,8239	6,5709	6,4687	9,1742
Fuzzy C Means	17,7853	5,5523	6,5437	6,4952	9,0869
	Jumlah cluster = 3				
K Means	11,833	4,3337	5,4301	4,1243	6,4303
Fuzzy C Means	11,3647	4,3907	3,3625	3,3458	5,6161
	Jumlah cluster = 4				
K Means	10,5053	7,2512	7,12558	4,9402	7,4555
Fuzzy C Means	10,2960	7,3139	7,12551	4,8992	7,4087
	Jumlah cluster = 5				
K Means	12,4229	6,8783	6,0683	4,7076	7,5193
Fuzzy C Means	11,0032	7,3139	7,9734	4,2069	7,7459
	Jumlah cluster = 6				
K Means	10,2120	6,4177	8,6608	4,2037	7,3486
Fuzzy C Means	12,8757	8,3955	8,9081	4,7252	8,7261
	Jumlah cluster = 7				
K Means	10,7950	6,7180	8,9830	4,1537	7,6624
Fuzzy C Means	12,2491	11,665	10,741	5,1241	9,9450
	Jumlah cluster = 8				
K Means	11,8465	11,4456	10,1275	5,3142	9,6834
Fuzzy C Means	11,7042	11,2759	10,9591	4,6129	9,6380
	Jumlah cluster = 9				

K Means	11,0597	13,1884	9,8525	6,8229	10,2309
Fuzzy C Means	11,1788	10,6098	10,7357	5,0017	9,38115
Jumlah cluster = 10					
K Means	10,9852	11,6214	9,9417	5,5986	9,5367
Fuzzy C Means	10,9118	10,0386	9,7480	4,6616	8,8400

Dari tabel 1 diatas, terlihat bahwa rata-rata simpangan baku menggunakan metode Fuzzy C Means memberikan nilai rasio simpangan baku yang lebih kecil dibandingkan metode K Means. Hal ini menjelaskan bahwa pengelompokkan dengan metode Fuzzy C Means memberikan hasil pengelompokkan yang lebih baik.

3.2 Penentuan Jumlah Cluster Optimum

Jumlah cluster optimum diberikan ketika nilai Xie dan Beni minimum pada lembah pertama.

Tabel 2. Nilai Indeks Xie dan Beni

Jumlah Kelompok	Banyaknya Iterasi	Fungsi Objektif	Indeks Xie dan Beni
2	37	2808,22	0,001682
3	77	1741,35	0,001665
4	40	1210,29	0,000865
5	86	931,54	0,000791
6	123	737,84	0,000658
7	107	599,03	0,000723
8	80	496,42	0,000666
9	70	426,025	0,001035
10	63	369,022	0,001005

Dari tabel 2, *Clustering* yang dilakukan diperoleh hasil indeks Xie dan Beni, nilai fungsi obyektif selama iterasi. Dalam penelitian ini, didapatkan indeks xie dan beni optimum adalah 0,000658 dengan jumlah kelompok yaitu 6, proses iterasinya berhenti pada iterasi ke-123 karena nilai $|Pt - Pt-1| < \xi$. Nilai fungsi obyektif pada iterasi terakhir yang diperoleh adalah 737,84. Banyaknya kabupaten/kota anggota kelompok pertama sampai dengan kelompok keenam berturut-turut adalah 32, 18, 26, 50, 49, 44.

3.3 Nilai pusat cluster atau center

Pada iterasi terakhir (iterasi ke-123), pusat *cluster* V_{kj} yang dihasilkan dengan $k=1,2,3,4$ dan $j=1,2,3,4,5,6$ adalah:

$$V_{kj} = \begin{pmatrix} 96.44 & 96.02 & 139.59 & 627.04 \\ 168.01 & 97.35 & 16.47 & 642.89 \\ 62.24 & 44.94 & 30.13 & 592.38 \\ 101.87 & 90.89 & 72.85 & 619.01 \\ 30.78 & 92.99 & 102.86 & 622.68 \\ 163.34 & 94.86 & 124.82 & 634.55 \end{pmatrix}$$

Nilai ini merupakan nilai dari koordinat keenam titik pusat *cluster* dan memberikan garis besar tiap *cluster* yaitu:

- Untuk pusat *cluster* 1 terdiri dari provinsi dengan angka harapan hidup (AHH) yaitu 96,44, Angka Melek Huruf yaitu 96,02, Rata-rata Lama Sekolah (LS) yaitu 139,59, Paritas Daya Beli (PPP) yaitu 627,04.
 - Untuk pusat *cluster* 2 terdiri dari provinsi dengan angka harapan hidup (AHH) yaitu 168,01, Angka Melek Huruf yaitu 97,35, Rata-rata Lama Sekolah (LS) yaitu 16,47, Paritas Daya Beli (PPP) yaitu 642,89%.
 - Untuk pusat *cluster* 3 terdiri dari provinsi dengan angka harapan hidup (AHH) yaitu 62,24, Angka Melek Huruf yaitu 44,94, Rata-rata Lama Sekolah (LS) yaitu 30,13, Paritas Daya Beli (PPP) yaitu 592,38.
 - Untuk pusat *cluster* 4 terdiri dari provinsi dengan angka harapan hidup (AHH) yaitu 101,87, Angka Melek Huruf yaitu 90,89, Rata-rata Lama Sekolah (LS) yaitu 72,85, Paritas Daya Beli (PPP) yaitu 619,01.
 - Untuk pusat *cluster* 5 terdiri dari provinsi dengan angka harapan hidup (AHH) yaitu 30,78, Angka Melek Huruf yaitu 92,99, Rata-rata Lama Sekolah (LS) yaitu 102,86, Paritas Daya Beli (PPP) yaitu 622,68%.
 - Untuk pusat *cluster* 6 terdiri dari provinsi dengan angka harapan hidup (AHH) yaitu 163,34, Angka Melek Huruf yaitu 94,86, Rata-rata Lama Sekolah (LS) yaitu 124,82, Paritas Daya Beli (PPP) yaitu 634,55
- 3.4 Derajat keanggotaan kabupaten/kota untuk setiap *cluster* pada iterasi terakhir (iterasi ke-123) dapat dilihat pada tabel 3 dibawah ini:

Tabel 3. Derajat Keanggotaan Kabupaten/Kota Untuk Setiap Cluster Pada Iterasi Terakhir

No	Provinsi	Derajat Keanggotaan (μ) Provinsi Untuk Setiap Cluster Pada Iterasi Terakhir					
		μ_{i1}	μ_{i2}	μ_{i3}	μ_{i4}	μ_{i5}	μ_{i6}
1	Kab. Sambas	0.094	0.060	0.234	0.160	0.396	0.052
2	Kab. Bengkayang	0.072	0.173	0.093	0.515	0.053	0.092
3	Kab. Landak	0.028	0.005	0.016	0.031	0.908	0.008

4	Kab. Pontianak	0.077	0.062	0.135	0.547	0.127	0.048
5	Kab. Sanggau	0.065	0.181	0.091	0.531	0.053	0.077
6	Kab. Ketapang	0.048	0.053	0.085	0.720	0.055	0.036
7	Kab. Sintang	0.044	0.056	0.043	0.776	0.031	0.046
8	Kab. Kapuas Hulu	0.133	0.042	0.092	0.340	0.338	0.053

217	Kab. Intan Jaya	0.019	0.024	0.863	0.047	0.030	0.015
218	Kab. Deiyai	0.010	0.011	0.924	0.026	0.018	0.008
219	Kota Jayapura	0.064	0.442	0.115	0.237	0.065	0.074

Dari derajat keanggotaan setiap kabupaten/kota pada iterasi terakhir dapat diperoleh informasi mengenai kecenderungan kabupaten/kota untuk masuk ke cluster mana. Derajat keanggotaan terbesar menunjukkan bahwa kecenderungan tertinggi kabupaten/kota untuk masuk menjadi anggota cluster tersebut. Misalnya untuk Kab. Sambas, dapat menjadi:

- a) Anggota cluster ke-1 dengan derajat keanggotaan 0.094
- b) Anggota cluster ke-2 dengan derajat keanggotaan 0.060
- c) Anggota cluster ke-3 dengan derajat keanggotaan 0.234
- d) Anggota cluster ke-4 dengan derajat keanggotaan 0.160
- e) Anggota cluster ke-5 dengan derajat keanggotaan 0.396
- f) Anggota cluster ke-6 dengan derajat keanggotaan 0.052

Derajat keanggotaan terbesarnya terletak di cluster kelima, maka kabupaten Sambas akan dimasukkan kedalam cluster ke-5. Hasil selengkapnya pengelompokan 219 kabupaten/kota kedalam 6 cluster dapat dilihat pada tabel dibawah ini:

Tabel 4. Hasil Pengelompokan Kabupaten/Kota di Kawasan Indonesia Timur menggunakan FCM

Kelompok	Anggota Kelompok
1	Kota Pontianak, Kab. Barito Selatan, Kab. Lamandau, Kab. Seruyan, Kab. Katingan, Kab. Pulang Pisau, Kab. Gunung Mas, Kab. Barito Timur, Kota Banjarmasin, Kab. Kutai Kartanegara, Kab. Kutai Timur, Kab. Malinau, Siau Tagulandang Biaro, Banggai, Konawe, Kolaka, Buton Utara, Konawe Utara, Kota Mataram, Kab. Maluku Tenggara, Kab. Maluku Tengah, Kab. Buru, Kab.

	Kepulauan Aru, Kab. Seram Bagian Barat, Kab. Halmahera Tengah, Kab. Halmahera Utara, Kab. Manokwari, Kab. Sorong Selatan, Kab. Sorong, Kab. Maybrat, Kab. Jayapura, Kab. Biak Numfor
2	Kota Palangka Raya, Kota Banjar Baru, Kota Balikpapan, Kota Samarinda, Kota Bontang, Manado, Tomohon, Palu, Bantaeng, Wajo, Kota Makassar, Palopo, Kendari, Kota Kupang, Kota Ambon, Kota Ternate, Kota Sorong, Kota Jayapura
3	Kab. Kayong Utara, Kab. Lombok Utara, Kota Bima, Kab. Sumba Barat, Kab. Sumba Tengah, Kab. Sumba Barat Daya, Kab. Sabu Raijua, Kota Gorontalo, Kab. Tandrauw, Kab. Jayawijaya, Kab. Boven Digoel, Kab. Mappi, Kab. Asmat, Kab. Yahukimo, Kab. Pegunungan Bintang, Kab. Tolikara, Kab. Waropen, Kab. Mamberamo Raya, Kab. Nduga, Kab. Lanny Jaya, Kab. Mamberamo Tengah, Kab. Yalimo, Kab. Puncak, Kab. Dogiyai, Kab. Intan Jaya, Kab. Deiyai
4	Kab. Bengkayang, Kab. Pontianak, Kab. Sanggau, Kab. Ketapang, Kab. Sintang, Kab. Kapuas Hulu, Kab. Sekadau, Kab. Melawi, Kota Singkawang, Kab. Sukamara, Kab. Murung Raya, Kab. Tanah Laut, Kab. Tapin, Bolang Mongondow Selatan, Kepulauan Selayar, Takalar, Pangkajene dan Kepulauan, Bone, Buton, Konawe Selatan, Bombana, Wakatobi, Mamuju, Mamuju Utara, Kab. Timor Tengah Selatan, Kab. Timor Tengah Utara, Kab. Belu, Kab. Alor, Kab. Lembata, Kab. Flores Timur, Kab. Sikka, Kab. Ngada, Kab. Manggarai, Kab. Rote Ndao, Kab. Manggarai Barat, Kab. Manggarai Timur, Kab. Boalemo, Kab. Gorontalo, Kab. Pohuwato, Kab. Gorontalo Utara, Kab. Buru Selatan, Kab. Teluk Wondama, Kab. Teluk Bintuni, Kab. Nabire, Kab. Yapen Waropen, Kab. Paniai, Kab. Puncak Jaya, Kab. Mimika, Kab. Sarmi, Kab. Keerom
5	Kab. Sambas, Kab. Landak, Kab. Kubu Raya, Kota Baru, Kab. Banjar, Kab. Barito Kuala, Kab. Hulu Sungai Selatan, Kab. Hulu Sungai Tengah, Kab. Hulu Sungai Utara, Kab. Tabalong, Kab. Tanah Bumbu, Kab. Balangan, Banggai Kepulauan, Morowali, Poso, Donggala, Tolitoli, Buol, Parigi Moutong, Tojo Unauna, Sigi, Jeneponto, Muna, Kolaka Utara, Majene, Polewali Mandar, Kab. Lombok Barat, Kab. Lombok Tengah , Kab. Lombok Timur, Kab. Sumbawa, Kab. Dompu, Kab. Bima, Kab. Sumbawa Barat, Kab. Sumba Timur, Kab. Kupang, Kab. Ende, Kab. Nagekeo, Kab. Maluku Tenggara Barat, Kab. Seram Bagian Timur, Kab. Maluku Barat Daya, Kab. Halmahera Barat, Kab. Kepulauan Sula, Kab. Halmahera Selatan, Kab. Halmahera Timur, Kab. Pulau Morotai, Kota Tidore Kepulauan, Kab. Raja Ampat, Kab. Merauke, Kab. Supiori
6	Kab. Kotawaringin Barat, Kab. Kotawaringin Timur, Kab. Kapuas, Kab. Barito Utara, Kab. Pasir, Kab. Kutai Barat, Kab. Berau, Kab. Bulungan, Kab. Nunukan, Kab. Penajam Paser Utara, Kab. Tana Tidung, Kota Tarakan, Bolang

	Mongondow, Minahasa, Kepulauan Sangihe, Kepulauan Talaud, Minahasa Selatan, Minahasa Utara, Bolang Mongondow Utara, Minahasa Tenggara, Bolang Mongondow Timur, Bitung, Kotamobagu, Bulukumba, Gowa, Sinjai, Maros, Barru, Soppeng, Rappang, Pinrang, Enrekang, Luwu, Tana Toraja, Luwu Utara, Luwu Timur, Toraja Utara, Kota Parepare, Baubau, Mamasa, Kab. Bone Bolango, Kota Tual, Kab. Fakfak, Kab. Kaimana
--	--

Setelah kelompok terbentuk, terhadap seluruh obyek penelitian sebanyak 219 kabupaten/kota diambil rata-rata dari masing-masing variabel pembentuk IPM yaitu AHH, AMH, LS, dan PPP (\bar{X}). Selanjutnya masing-masing kelompok diambil rata-rata untuk variabel AHH, AMH, MYS, dan PPP (\bar{X}_c). Pada setiap variabel didalam kelompok diberi tanda, jika ($\bar{X} \geq \bar{X}_c$) maka diberi tanda positif (+), sedangkan jika ($\bar{X} \leq \bar{X}_c$) maka diberi tanda negatif (-).

Tabel 5. Karakteristik Kelompok Berdasarkan Rata-Rata

Kelompok	AHH	AMH	LS	PPP
1	-	+	+	+
2	+	+	+	+
3	-	-	-	-
4	+	+	+	-
5	-	+	+	+
6	+	+	+	+

Berdasarkan Tabel 5 terlihat bahwa kelompok 1 variabel yang memiliki nilai rata-rata diatas nilai rata-rata di Kawasan Indonesia Timur adalah angka melek huruf, rata-rata lama sekolah, dan paritas daya beli. Sehingga pada kelompok 1 variabel yang perlu dapat perhatian lebih adalah variabel angka harapan hidup.

Kelompok 2 memiliki nilai rata-rata pembentuk Indeks Pembangunan Manusia (IPM) diatas nilai rata-rata di Kawasan Indonesia Timur. Sehingga anggota yang berada di kelompok 2 merupakan kelompok yang paling berhasil dari kelompok lainnya.

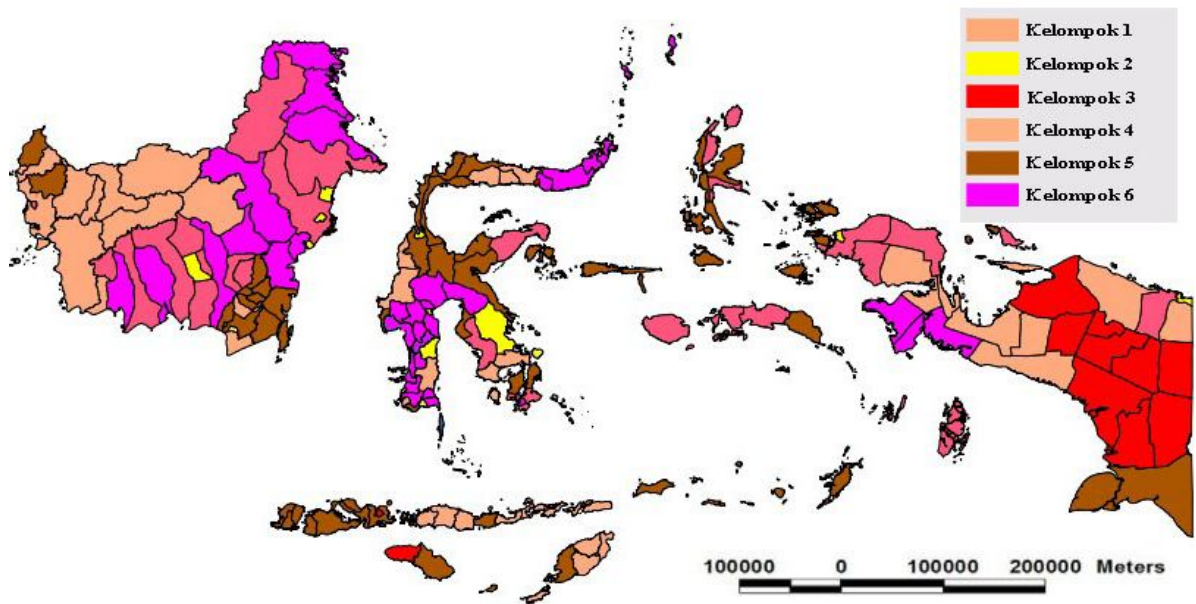
Kelompok 3 memiliki nilai rata-rata pembentuk Indeks Pembangunan Manusia (IPM) dibawah nilai rata-rata di Kawasan Indonesia Timur. Sehingga anggota yang berada di kelompok 3 merupakan kelompok yang dengan IPM terendah dari kelompok lainnya.

Kelompok 4 variabel yang memiliki nilai rata-rata diatas nilai rata-rata di Kawasan Indonesia Timur adalah angka harapan hidup, angka melek huruf, rata-rata lama sekolah. Sehingga pada kelompok 4 variabel yang perlu dapat perhatian lebih adalah variabel paritas daya beli.

Kelompok 5 memiliki karakteristik yang hampir sama dibandingkan kelompok 1. Namun kelompok 5 memiliki rata-rata nilai pembangun IPM lebih kecil dibandingkan

kelompok 1. Variabel yang memiliki nilai rata-rata diatas nilai rata-rata di Kawasan Indonesia Timur adalah angka melek huruf, rata-rata lama sekolah, dan paritas daya beli. Sehingga pada kelompok 1 variabel yang perlu dapat perhatian lebih adalah variabel angka harapan hidup.

Kelompok 6 memiliki karakteristik yang hampir sama dengan kelompok 2. Namun kelompok 6 memiliki rata-rata nilai pembangun IPM lebih kecil dibandingkan kelompok 2 . Kelompok 6 memiliki nilai rata-rata pembentuk Indeks Pembangunan Manusia (IPM) diatas nilai rata-rata di Kawasan Indonesia Timur.



Gambar 1. Peta Kawasan Indonesia Timur Berdasarkan Pengelompokan Pembentuk Indeks Pembangunan Manusia 2012

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang dilakukan untuk mengelompokkan kabupaten/kota di Kawasan Indonesia Timur berdasarkan Indeks Pembangunan Manusia (IPM) diambil keputusan:

1. Berdasarkan evaluasi hasil pengelompokkan untuk mendapatkan nilai kualitas ketepatan pengelompokkan menggunakan simpangan baku dalam kelompok dan antar kelompok. Pengelompokkan menggunakan metode Fuzzy C Means karena menghasilkan rasio S_w/S_b yang lebih kecil dibandingkan metode K Means.
2. Penentuan jumlah cluster optimum menggunakan Indeks Xie dan Beni menghasilkan 6 kelompok dengan nilai 0,000658.
3. Berdasarkan hasil karakteristik setiap kelompok yang dibandingkan dengan nilai rata-rata IPM kabupaten/kota di Kawasan Indonesia Timur. Kelompok 2 dan Kelompok 6 memiliki nilai rata-rata pembentuk Indeks Pembangunan Manusia (IPM) diatas nilai rata-rata di Kawasan Indonesia Timur. Sehingga anggota yang

berada di kelompok 2 dan kelompok merupakan kelompok yang berhasil dari kelompok lainnya. Namun kelompok 2 memiliki nilai pembentuk Indeks Pembangunan Manusia yang lebih tinggi dibandingkan kelompok 6.

4. Sementara, Kelompok 3 memiliki nilai rata-rata pembentuk Indeks Pembangunan Manusia (IPM) dibawah nilai rata-rata di Kawasan Indonesia Timur. Sehingga anggota yang berada di kelompok 3 merupakan kelompok yang dengan IPM terendah dari kelompok lainnya.

DAFTAR PUSTAKA

- [1] Ananthi Sheshasayee dan P. Sharmila. 2004. *Comparative Study of Fuzzy C Means and K Means Algorithm for Requirements Clustering*. Vol 7(6), 853-857.
- [2] Arika, Yovita. 2012. Indeks Pembangunan Manusia Indonesia Sangat Rendah. KOMPAS, 17 April 2012.
- [3] Babankumar S. Bansod and O. P. Pandey. 2013. *An Application of PCA and Fuzzy C-Means to Delineate Management Zones and Variability Analysis of Soil*. ISSN 1064-2293. Vol.46 No.5, pp.556-564
- [4] Bezdek, J. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithm*. Plenum Press., New York
- [5] Bunkers, M.J dan Miller, J.R. 1996. *Definition of Climate Regions in the Northern Plains Using an Objective Cluster Modification Technique*. Journal of Climate. Vol.9, pp. 130-146.
- [6] BPS. 2012. *Indeks Pembangunan Manusia 2010-2011*.
- [7] BPS. 2011. *Indeks Pembangunan Manusia 2009-2010*.
- [8] Duo, C., Xue, L. dan Du-Wu, C. 2007. *An Adaptive Cluster Validity Index for the Fuzzy C-Means*. International Journal of Computer Science and Network Security, Vol.7 No.2, pp.146-156.
- [9] Klawonn, F. dan Keller, A. 1997. *Fuzzy Clustering and Fuzzy Rules*. Science Journal
- [10] Kusumadewi, Sri dan Hari Purnomo. 2004. *Aplikasi Logika Fuzzy Untuk Mendukung Keputusan*. Yogyakarta: Graha Ilmu.
- [11] Oyelade, O. J., Oladipupo, O. O., and Obagbuwa, O. O. 2010. *Application of k-means clustering algorithm for prediction of Students' Academic performance*. International Journal of Computer Science and Information Security. Vol 7(1). 292-295, (IF: 0.423).
- [12] Purnamasari, S.B. 2014. *Pemilihan Cluster Optimum Pada Fuzzy C-Means (Studi Kasus: Pengelompokan Kabupaten/Kota di Jawa Tengah berdasarkan Indikator Indeks Pembangunan Manusia)*. Universitas Diponegoro. Semarang
- [13] Republik Indonesia. 2001. Menurut Keputusan Presiden Republik Indonesia Nomor 55 Tahun 2001 Tentang Dewan Pengembangan Kawasan Indonesia Timur. Sekretariat Kabinet RI. Jakarta
- [14] Shihab, A. I., 2000, *Fuzzy Clustering Algorithm and Their Application to Medical Image Analysis*, Dissertation, University of London, London.

- [15] Sukim. 2011. *Studi Tentang Metode C-Means Cluster dan Fuzzy C-Means Cluster Serta Aplikasinya Pada Kasus Pengelompokan Desa/Kelurahan Berdasarkan Status Ketertinggalan*. Institut Teknologi Sepuluh Nopember, Surabaya.
- [16] Widodo, Agus., Purhadi. 2012. *Perbandingan Metode Fuzzy C-Means dan Fuzzy C-Shell Clustering (Studi Kasus Kabupaten/Kota di Pulau Jawa Berdasarkan Variabel Pembentuk Indeks Pembangunan Manusia)*. Institut Teknologi Sepuluh Nopember, Surabaya.
- [17] Xie, X.I dan Beni, G. 1991, *A validity measure for fuzzy clustering*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13. hal. 841-847.
- [18] Yunitasari, M. 2007. *Analisis Hubungan Antara Pertumbuhan Ekonomi Dengan Pembangunan Manusia*. Jawa Timur
- [19] Zadeh, L.A. 1965. *Fuzzy Sets, Information and Control Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley, California*.
- [20] Z. Hilal Inana and Mehmet Kuntalp 2007. *A Study on Fuzzy C-Means Clustering-Based Systems in Automatic Spike Detection*. 1160-1166