

Penerapan *Partition Entropy Index*, *Partition Coefficient Index* dan *Xie Beni Index* untuk Penentuan Jumlah Kluster Optimal pada Algoritma Fuzzy C-Means dalam Pemetaan Tingkat Kesejahteraan Penduduk Jawa Tengah

Syaidati Mashfuufah^{1*}, Deden Istiawan²
Akademi Statistika Muhammadiyah Semarang
*Email: syayidaa@gmail.com

Abstrak

Keywords:
Kemiskinan;
Fuzzy C-Means;
Cluster Validity
Index; Kluster
Optimal.

Kemiskinan merupakan salah satu indikator ekonomi untuk melihat tingkat kesejahteraan penduduk di suatu daerah. Menurut data BPS pada September 2016, Jawa Tengah tercatat sebagai provinsi dengan tingkat kemiskinan tertinggi di Jawa. Dalam strategi penanggulangan kemiskinan dibutuhkan data kemiskinan yang akurat sesuai karakteristik tingkat kesejahteraan, sehingga pemerintah dapat menghasilkan kebijakan yang tepat sasaran dan tepat guna. Untuk pengelompokan tingkat kesejahteraan kabupaten/kota di Jawa Tengah diusulkan algoritma Fuzzy C-Means (FCM). FCM memiliki kelebihan dalam ketepatan penempatan pusat kluster dan sangat stabil terhadap outlier, namun jumlah kluster dalam FCM harus ditentukan terlebih dahulu sebelum melakukan penelitian. Untuk penentuan jumlah kluster optimal pada algoritma FCM, pada penelitian ini diusulkan Cluster Validity Index (CVI). CVI merupakan sebuah ukuran validitas untuk menemukan jumlah kluster optimal yang sepenuhnya dapat menjelaskan struktur data dengan cara mengukur derajat kekompakan dalam satu kluster dan separasi antar kluster. Pada penelitian ini diusulkan CVI Partition Entropy Index (PEI), Partition Coefficient Index (PCI) dan Xie Beni Index (XBI). Hasil penelitian menunjukkan jumlah kluster optimal berdasarkan CVI adalah dua kluster. Kluster pertama terdiri dari 22 kabupaten/kota dan kluster kedua terdiri dari 13 kabupaten/kota.

1. PENDAHULUAN

Kemiskinan merupakan masalah klasik yang dihadapi oleh sebagian besar negara berkembang. Kemiskinan juga merupakan salah satu indikator ekonomi untuk melihat tingkat kesejahteraan penduduk di suatu daerah (Hudaya, 2009). Badan Pusat Statistik (BPS) mendefinisikan kemiskinan sebagai keadaan ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran (BPS, 2017). Di Indonesia, kemiskinan merupakan beban berat dan masalah serius jika dilihat dari ketimpangan antara penduduk miskin dan penduduk kaya. Pada September tahun 2016 Jawa Tengah tercatat sebagai provinsi dengan presentase kemiskinan tertinggi di Pulau Jawa yakni 13,19 persen. Jumlah penduduk miskin di Jawa Tengah yakni 4493,75 jiwa merupakan angka tertinggi kedua setelah Provinsi Jawa Timur. Jika dibandingkan dengan tahun 2015, angka kemiskinan Jawa Tengah mengalami penurunan sebesar 0,13 persennamun angka tersebut masih tercatat lebih tinggi apabila dibandingkan dengan angka kemiskinan nasional yaitu 10,70 persen (BPS, 2017).

Pemerintah Provinsi Jawa Tengah memasukkan persoalan kemiskinan ke dalam Rancangan Pembangunan Jangka Menengah Daerah (RPJMD) (Nidyasofa, 2017) yang telah disesuaikan dengan tujuan utama SDGs (*Sustainable Development Goals*) yaitu *no poverty*, tidak ada kemiskinan dalam bentuk apapun di seluruh penjuru dunia (Ishartono and Raharjo, 2015). Untuk mendukung strategi penanggulangan kemiskinan dibutuhkan data yang akurat

dan tepat sasaran (BPS, 2016) yakni dengan melakukan penelitian terkait pemetaan daerah dan karakteristik kemiskinan (Anuraga, 2015). Untuk memahami kemiskinan, tingkat lokalitas masing-masing daerah perlu diperhatikan sehingga kriteria dan pendataan kemiskinan, penentuan sasaran, pemecahan masalah dan upaya penanggulangan kemiskinan dapat dilakukan secara objektif, tepat sasaran dan tepat guna (Jajuli, 2015)(Yulianto and Hidayatullah, 2014).

Pemetaan dan pengelompokan daerah berdasarkan karakteristik kemiskinan dan indikator kesejahteraan dapat dilakukan menggunakan analisis kluster. Analisis kluster atau klustering merupakan suatu teknik dalam *data mining* yang bersifat *unsupervised learning*. Klustering bertujuan untuk mengelompokkan data tidak berlabel ke dalam beberapa kelompok berdasarkan karakteristik masing-masing. Data yang memiliki kemiripan karakteristik terletak pada kelompok yang sama sedangkan yang memiliki karakteristik berbeda terletak pada kelompok yang berbeda (Han, Kamber and Pei, 2012). Salah satu teknik dalam klustering yang paling banyak dikenal dan sering digunakan adalah Fuzzy C-Means (Simhachalam and Ganesan, 2015).

Penelitian mengenai pengelompokan kemiskinan di Jawa Tengah menggunakan metode Fuzzy C-Means pernah dilakukan oleh Nidyasofa (Nidyasofa, 2017). Penelitian tersebut mengelompokkan kabupaten/kota di Jawa Tengah menjadi tiga kelompok berdasarkan variabel indikator kemiskinan yaitu jumlah penduduk, kepala rumah tangga (KRT) perempuan, anak tidak sekolah, individu yang cacat, individu yang memiliki penyakit kronis, pengangguran, sumber air minum tidak terlindungi, sumber penerangan tidak listrik, bahan bakar memasak menggunakan kayu bakar/arang/minyak tanah dan fasilitas tempat buang air besar (BAB) tidak tersedia. Penelitian lain dilakukan oleh Purnamasari dengan menerapkan algoritma Fuzzy C-Means untuk mengelompokkan kabupaten/kota di Jawa Tengah berdasarkan indikator Indeks Pembangunan Manusia (IPM) yaitu Angka Harapan Hidup (AHH), Angka Melek Huruf (AMH), Rata-rata Lama Sekolah (RLS), dan Kemampuan Daya Beli Penduduk atau *Purchasing Power Parity* (PPP) (Purnamasari, Yasin and Wuryandari, 2014).

Algoritma Fuzzy C-Means merupakan algoritma terbaik dibandingkan dengan algoritma K-Means, *Self Organizing Map* (SOM) *neural network* dan *clustering hierarchy* (Mingoti and Lima, 2006) karena sangat stabil terhadap *outlier* dan *overlapping* serta akurasi dan kualitas kluster yang komparatif (Sivarathri and Govardhan, 2014). Algoritma Fuzzy C-Means memiliki kemampuan yang baik dalam mendeteksi kluster (Pimentel and de Souza, 2016) dan memiliki ketepatan pada penempatan pusat kluster (Scitovski, Vidovi and Bajer, 2016). Namun algoritma Fuzzy C-Means memiliki kelemahan pada penentuan jumlah kluster optimal (Yang and Nataliani, 2017)(Rezaee, 2010) yang harus ditentukan oleh peneliti sebelum melakukan penelitian. Penentuan jumlah kluster yang tidak sesuai akan berdampak pada hasil partisi yang tidak optimal sehingga dapat mengakibatkan kesalahan ketika pengambilan keputusan (Wang and Zhang, 2007). Penentuan jumlah kluster pada suatu data merupakan teori penting dan masalah praktis (Kolesnikov, Trichina and Kauranne, 2014) karena memiliki peran penting untuk mendapatkan hasil dugaan yang baik pada analisis kluster (Erilli *et al.*, 2011)(Wu *et al.*, 2014). Untuk memvalidasi apakah jumlah kluster yang ditentukan dapat menjelaskan keseluruhan data atau tidak, dapat menggunakan *Cluster Validity Index* (CVI) (Wang and Zhang, 2007). CVI merupakan ukuran validitas untuk menemukan jumlah kluster optimal yang sepenuhnya dapat menjelaskan struktur data (Wu and Yang, 2005)(Zhao and Fränti, 2014). CVI mengukur derajat kekompakan dan separasi struktur data pada seluruh kluster dan menemukan jumlah kluster optimal yang kompak dan terpisah dari kluster yang lain (Wu and Yang, 2005).

Pada penelitian ini digunakan CVI yakni *Partition Entropy Index* (PEI), *Partition Coefficient Index* (PCI) dan *Xie Beni Index* (XBI) untuk memvalidasi dan mengetahui apakah jumlah kluster optimal dapat didapatkan diantara jumlah kluster yang telah ditentukan pada pengelompokan dengan menggunakan algoritma Fuzzy C-Means. Data yang digunakan yaitu data indikator status kesejahteraan penduduk Jawa Tengah dari Tim

Nasional Percepatan Penanggulangan Kemiskinan (TNP2K) tahun 2015 dengan sepuluh variable yaitu jumlah penduduk, kepala rumah tangga (KRT) perempuan, anak tidak sekolah, individu yang cacat, individu yang memiliki penyakit kronis, pengangguran, sumber air minum tidak terlindungi, sumber penerangan tidak listrik, bahan bakar memasak menggunakan kayu bakar/arang/minyak tanah dan fasilitas tempat buang air besar (BAB) tidak tersedia. *Partition Entropy Index* (PEI) diusulkan oleh James C. Bezdek pada tahun 1974 (Bezdek, 1974) untuk mengevaluasi keteracakan data dalam kluster. *Partition Coefficient Index* (PCI) juga diusulkan oleh James C. Bezdek pada tahun 1981 (Bezdek, 1981) dengan mengevaluasi nilai derajat keanggotaan tanpa memandang nilai vektor (data). Xie dan Beni juga mengusulkan indeks validitas *Xie Beni Index* (XBI) untuk mengevaluasi kluster yang diperoleh namun telah dimodifikasi oleh (Pal and Bezdek, 1995).

2. METODE

Klustering menggunakan algoritma Fuzzy C-Means pertama kali diperkenalkan oleh Dunn pada tahun 1973 dan diperbaiki oleh James C. Bezdek (Bezdek, Ehrlich and Full, 1984). Algoritma Fuzzy C-Means merupakan algoritma klustering yang paling banyak dikenal dan memiliki ketepatan pada penempatan pusat kluster, namun memiliki kelemahan pada penentuan jumlah kluster optimal. Untuk itu penelitian ini menerapkan CVI yaitu *Partition Entropy Index* (PEI), *Partition Coefficient Index* (PCI) dan *Xie Beni Index* (XBI) pada algoritma Fuzzy C-Means dalam pengelompokan kabupaten/kota di Jawa Tengah berdasarkan data indikator status kesejahteraan dari Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K). Berikut ini adalah langkah-langkah algoritma Fuzzy C-Means:

1. Menentukan jumlah kluster ($k \geq 2$)
2. Menentukan pangkat untuk matriks partisi w ($w > 1$)
3. Menginisialisasi matriks partisi awal $U_i = [u_{ij}]$ untuk $i = 1, 2, \dots, k$ seperti persamaan

$$\sum_{j=1}^k u_{ij} = 1$$
4. Menghitung nilai *centroid* c_{ij} untuk $i = 1, 2, \dots, k$ menggunakan persamaan $c_{ij} = \frac{\sum_{i=1}^N u_{ij}^w x_{ij}}{\sum_{i=1}^N u_{ij}^w}$
5. Menghitung nilai derajat keanggotaan menggunakan persamaan $u_{ij} = \frac{D(x_i, c_j)^{-2}}{\sum_{l=1}^k D(x_i, c_l)^{-2}}$ untuk $U_i = [u_{ij}]$ untuk $i = 1, 2, \dots, k$
6. Menghitung nilai fungsi objektif menggunakan persamaan $J = \sum_{i=1}^N \sum_{j=1}^k (u_{ij})^w D(x_i, c_j)^2$
Jika perubahan nilai fungsi objektif sudah berada di bawah nilai ambang batas, maka berhenti. Jika masih di atas nilai ambang batas yang ditentukan, ulangi langkah keempat dan kelima.

Setelah jumlah kluster ditentukan pada algoritma Fuzzy C-Means, pertanyaan selanjutnya adalah apakah kluster tersebut secara akurat dapat menyajikan struktur data atau tidak. Karena pada algoritma Fuzzy C-Means membutuhkan penentuan jumlah kluster sejumlah k kluster maka perlu dilakukan validasi menggunakan indeks validitas (Wu and Yang, 2005). Berikut ini adalah penjelasan CVI yang digunakan pada penelitian ini.

1. *Partition Entropy Index* (PEI)
Indeks validitas yang pertama kali berhubungan dengan algoritma Fuzzy C-Means adalah *Partition Entropy Index* (Bezdek, 1974) didefinisikan dengan persamaan berikut

$$PEI = -\frac{1}{N} \left(\sum_{i=1}^N \sum_{j=1}^K u_{ij} \times \log_2 u_{ij} \right) \quad 1.1$$

Pada umumnya kluster optimal didapatkan jika nilai yang diperoleh mendekati kecil (mendekati 0)

2. *Partition Coefficient Index (PCI)*

Partition Coefficient Index didefinisikan dengan persamaan berikut

$$PCI = \frac{1}{N} (\sum_{i=1}^N \sum_{j=1}^K u_{ij}^2) \tag{1.2}$$

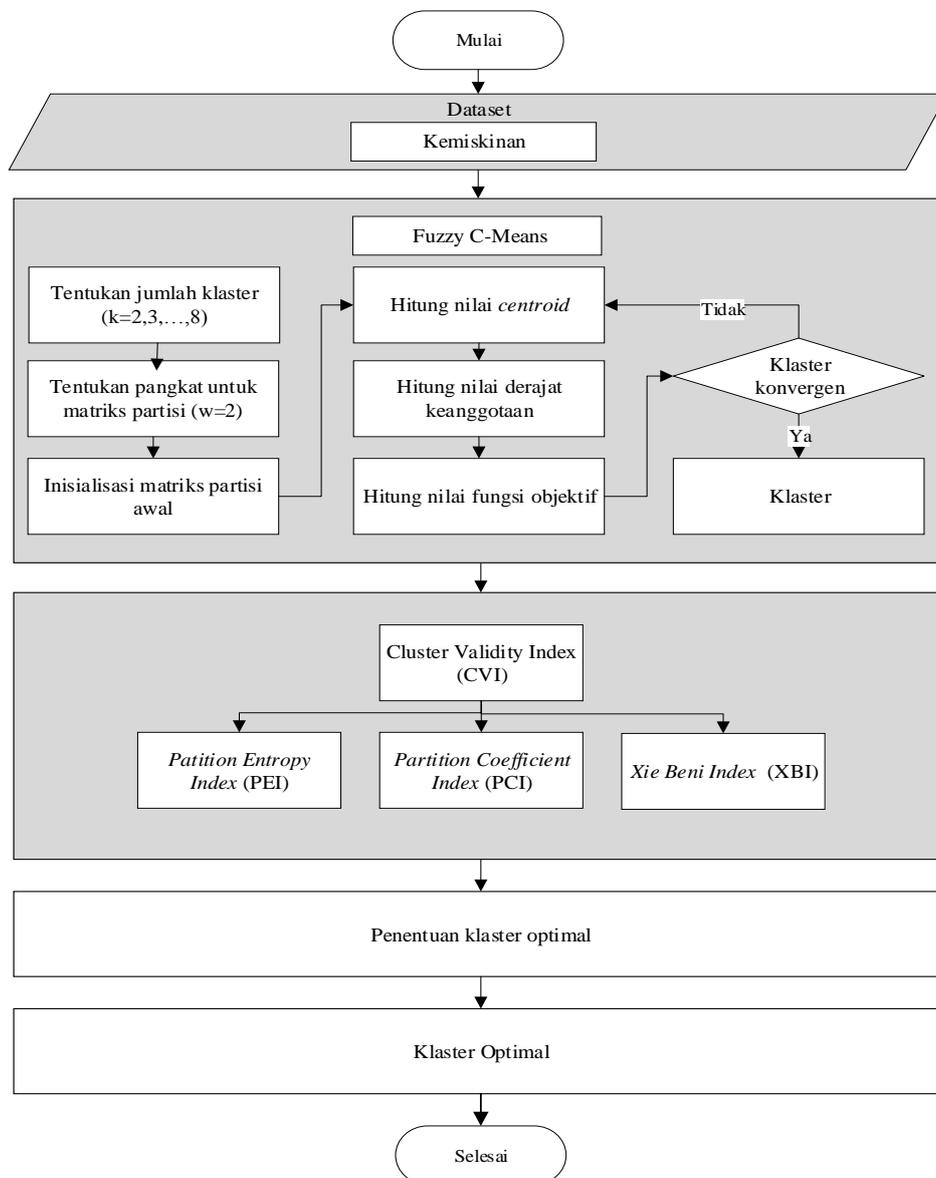
Pada umumnya kluster optimal didapatkan jika nilai yang diperoleh semakin besar (mendekati 1)

3. *Xie Beni Index (XBI)*

Sebuah fungsi validitas diusulkan oleh Xie dan Beni (1991) (Xie and Beni, 1991) dan telah dimodifikasi oleh Pal dan Bezdek (1995) (Pal and Bezdek, 1995) didefinisikan dengan persamaan berikut

$$XBI = \frac{\sum_{j=1}^K \sum_{i=1}^N u_{ij}^m \times d(x_i, c_j)^2}{N \times \min_{i,j} (d(c_i, c_j)^2)} = \frac{J_m(u, c) / N}{Sep(c)} \tag{1.3}$$

Gambar 1.berikut ini adalah gambar tahapan metode yang diusulkan



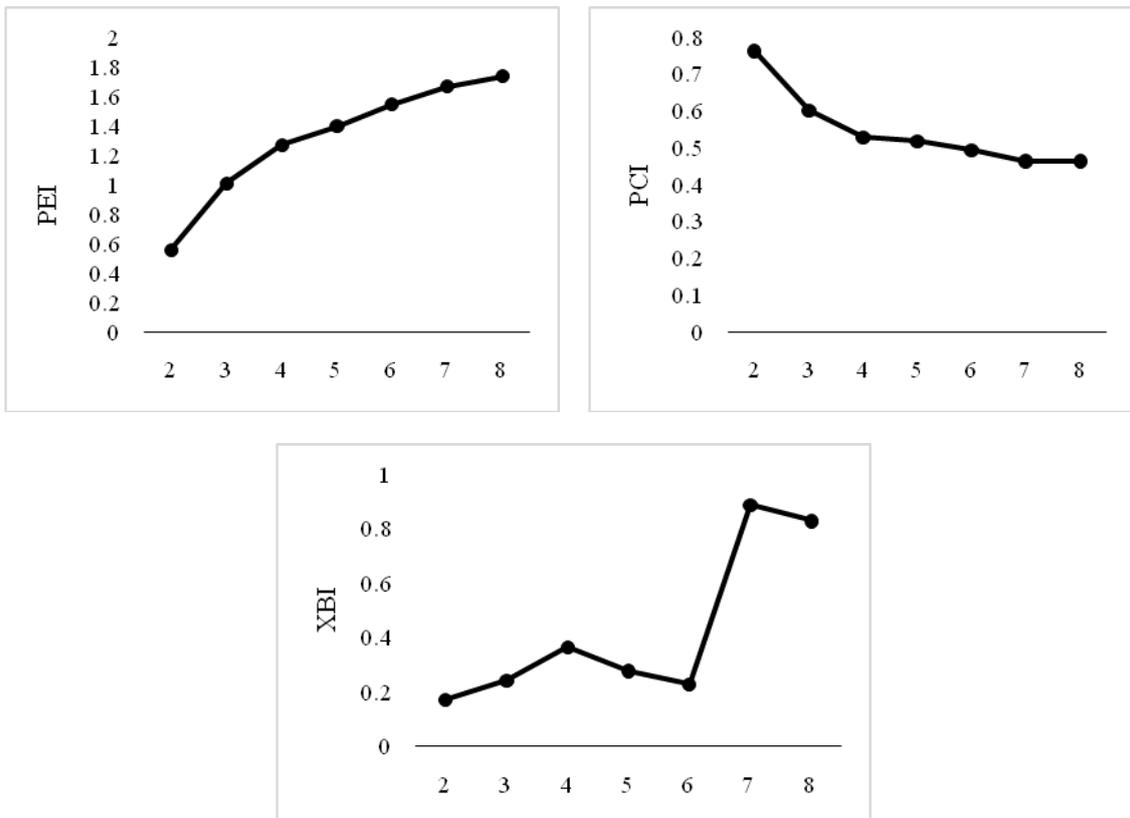
Gambar 1.TahapanMetode yang Diusulkan

3. HASIL DAN PEMBAHASAN

Analisis algoritma Fuzzy C-Means dan validasi kluster optimal menggunakan aplikasi *Matlab R2013a* dengan ukuran jarak Euclidean distance dan bobot (*weighting*) default, $w = 2$. Pada penelitian ini jumlah kluster ditentukan yakni $k = 2, 3, \dots, 8$ dan kluster konvergen diperoleh ketikaperubahan pada fungsi objektif sangat kecil sekali serta dilakukan perulangan sebanyak sepuluh kali karena inisialisasi matriks partisi awal dilakukan secara random. Hasil klustering dan validasi ditunjukkan pada tabel 3.1. Dari ketiga CVI PEI, PCI dan XBI menunjukkan bahwa kluster optimal pada pengelompokan kabupaten/kota di Jawa Tengah berdasarkan indikator status kesejahteraan TNP2K adalah dua kluster.

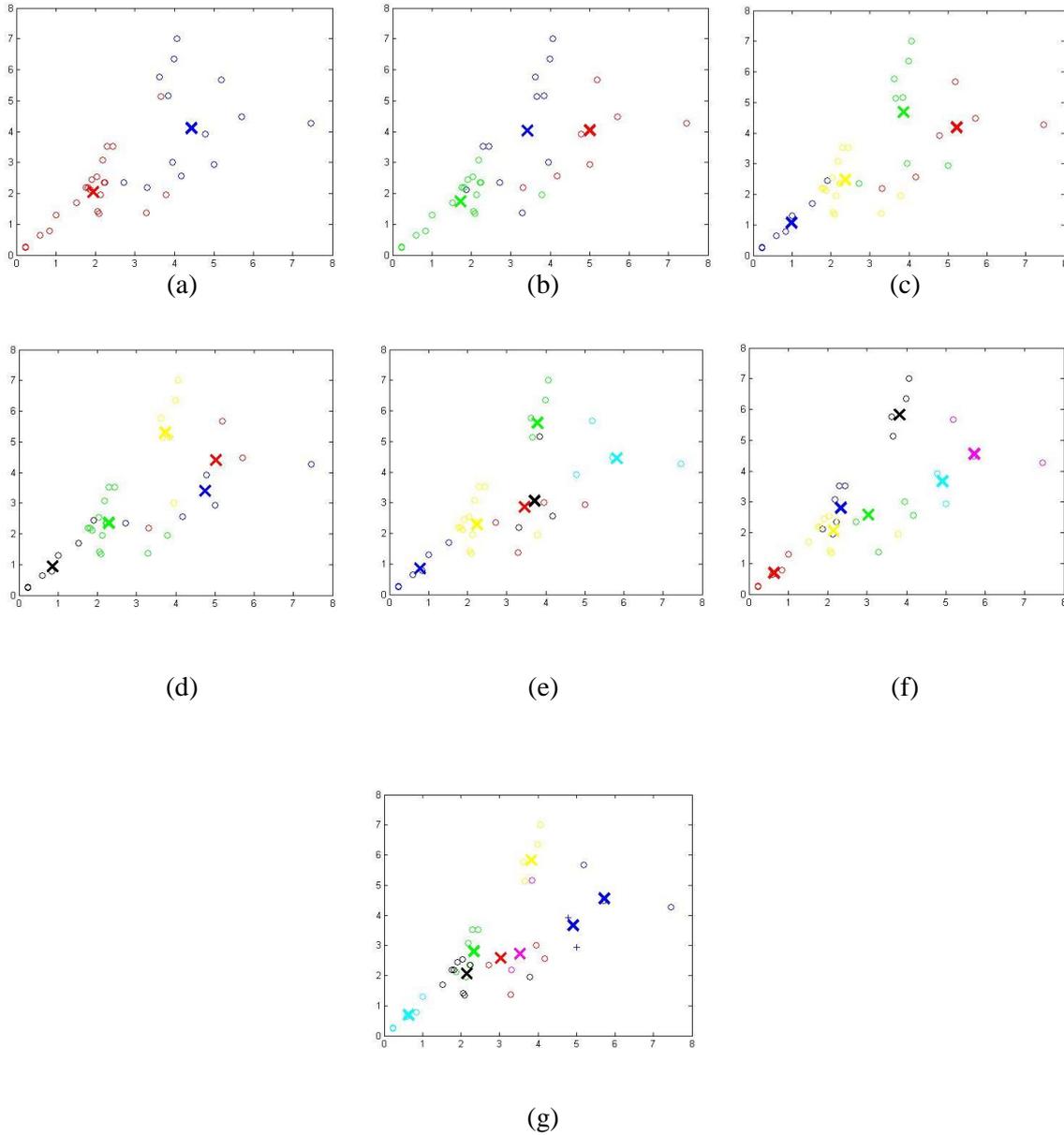
Tabel 1. Hasil Klustering dan Validasi

k	PEI	PCI	XBI
2	0.5551	0.761	0.172
3	1.0055	0.6	0.2415
4	1.2688	0.5273	0.3631
5	1.3955	0.5178	0.277
6	1.5463	0.4935	0.2275
7	1.6667	0.4636	0.8848
8	1.7342	0.4632	0.8266



Gambar 2. Grafik untuk Tabel 1.

Gambar 3. merupakan grafik untuk Tabel 1. yang menunjukkan indeks validitas untuk *dataset* tingkat kesejahteraan penduduk Jawa Tengah tahun 2015. Secara keseluruhan $k=2$ merupakan kluster optimal dan hasilnya sesuai dengan gambar 3.2 yang merupakan gambar struktur *dataset* tingkat kesejahteraan penduduk Jawa Tengah tahun 2015.



Gambar 3. Hasil klastering dengan jumlah kluster yang berbeda:
(a) $k=2$, (b) $k=3$, (c) $k=4$, (d) $k=5$, (e) $k=6$, (f) $k=7$, (g) $k=8$

Tabel 2. Hasil Klaster Kabupaten/Kota

ID	Klaster yang terbentuk	Kabupaten/Kota
1	Klaster 2	Cilacap
2	Klaster 2	Banyumas
3	Klaster 2	Purbalingga
4	Klaster 2	Banjarnegara
5	Klaster 2	Kebumen
6	Klaster 1	Purworejo
7	Klaster 1	Wonosobo
8	Klaster 2	Magelang
9	Klaster 1	Boyolali
10	Klaster 2	Klaten
11	Klaster 1	Sukoharjo
12	Klaster 1	Wonogiri
13	Klaster 1	Karanganyar
14	Klaster 1	Sragen
15	Klaster 2	Grobogan
16	Klaster 1	Blora
17	Klaster 1	Rembang
18	Klaster 2	Pati
19	Klaster 1	Kudus
20	Klaster 2	Jepara
21	Klaster 1	Demak
22	Klaster 1	Semarang
23	Klaster 1	Temanggung
24	Klaster 1	Kendal
25	Klaster 1	Batang
26	Klaster 1	Pekalongan
27	Klaster 2	Pematang
28	Klaster 2	Tegal
29	Klaster 2	Brebes
30	Klaster 1	Kota Magelang
31	Klaster 1	Kota Surakarta
32	Klaster 1	Kota Salatiga
33	Klaster 1	Kota Semarang
34	Klaster 1	Kota Pekalongan
35	Klaster 1	Kota Tegal

Berdasarkan hasil analisis di atas dapat diketahui bahwa:

1. Klaster pertama terdiri dari 22 kabupaten/kota yaitu Purworejo, Wonosobo, Boyolali, Sukoharjo, Wonogiri, Karanganyar, Sragen, Blora, Rembang, Kudus, Demak, Semarang, Temanggung, Kendal, Batang, Pekalongan, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan dan Kota Tegal. Klaster pertama ini memiliki karakteristik jumlah penduduk rendah, jumlah rumah tangga dengan KRT perempuan rendah, jumlah anak tidak sekolah rendah, jumlah individu yang cacat

rendah, jumlah individu yang memiliki penyakit kronis rendah, tingkat pengangguran rendah, jumlah rumah tangga dengan sumber air minum tidak terlindungi rendah, jumlah rumah tangga dengan sumber penerangan selain listrik rendah, jumlah rumah tangga yang menggunakan kayu bakar/arang/minyak tanah rendah, dan jumlah rumah tangga yang tidak memiliki fasilitas untuk BAB rendah. Berdasarkan hasil uraian di atas klaster pertama ini merupakan kelompok kabupaten/kota yang berkategori hampir miskin.

2. Klaster kedua terdiri dari 13 kabupaten/kota yaitu Cilacap, Banyumas, Purbalingga, Banjarnegara, Kebumen, Magelang, Klaten, Grobogan, Pati, Jepara, Pematang, Tegal dan Brebes. Klaster kedua ini memiliki karakteristik jumlah penduduk tinggi, jumlah rumah tangga dengan KRT perempuan tinggi, jumlah anak tidak sekolah tinggi, jumlah individu yang cacat tinggi, jumlah individu yang memiliki penyakit kronis tinggi, tingkat pengangguran tinggi, jumlah rumah tangga dengan sumber air minum tidak terlindungi tinggi, jumlah rumah tangga dengan sumber penerangan selain listrik tinggi, jumlah rumah tangga yang menggunakan kayu bakar/arang/minyak tanah tinggi, dan jumlah rumah tangga yang tidak memiliki fasilitas untuk BAB tinggi. Berdasarkan hasil uraian di atas klaster kedua ini merupakan kelompok kabupaten/kota yang berkategori miskin.

4. KESIMPULAN

Penerapan *Cluster Validity Index* (CVI) yakni *Partition Entropy Index* (PEI), *Partition Coefficient Index* (PCI) dan *Xie Beni Index* (XBI) pada algoritma Fuzzy C-Means dalam pemetaan kabupaten/kota di Jawa Tengah berdasarkan indikator status kesejahteraan penduduk Jawa Tengah pada tahun 2015 menghasilkan dua kelompok/klaster optimal. Kelompok pertama terdiri dari 22 kabupaten/kota dengan kategori hampir miskin dan kelompok kedua terdiri dari 13 kabupaten/kota dengan kategori miskin.

REFERENSI

- Anuraga, G. (2015) 'Hierarchical Clustering Multiscale Bootstrap untuk Pengelompokkan Kemiskinan di Jawa Timur', *Statistika*, 1, pp. 27–33.
- Bezdek, J. C. (1974) 'Numerical Taxonomy with Fuzzy Sets', *Journal of Mathematical Biology*, 1, pp. 57–71.
- Bezdek, J. C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. 1st edn. Edited by M. Nadler. New York: Springer. doi: DOI 10.1007/978-1-4757-0450-1.
- Bezdek, J. C., Ehrlich, R. and Full, W. (1984) 'FCM \square : The Fuzzy C-Means Clustering Algorithm', *Computer & Geosciences*, 10(2), pp. 191–203.
- [BPS] (2016) *Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2016*. Badan Pusat Statistik.
- [BPS] (2017) *Profil Kemiskinan di Indonesia 2016, Berita Resmi Statistik*. Badan Pusat Statistik.
- Erilli, N. A. *et al.* (2011) 'Determining the Most Proper Number of Cluster in Fuzzy Clustering by Using Artificial Neural Networks', *Elsevier*, 38, pp. 2248–2252. doi: 10.1016/j.eswa.2010.08.012.
- Han, J., Kamber, M. and Pei, J. (2012) *Data Mining Concepts and Techniques*. Third. Morgan Kaufmann.

- Hudaya, D. (2009) *Faktor-Faktor yang Mempengaruhi Tingkat Kemiskinan di Indonesia*. Institut Pertanian Bogor.
- Ishartono and Raharjo, S. T. (2015) ‘Sustainable Development Goals (Sdgs) dan Pengentasan Kemiskinan’, *Social Work Journal*, 6, pp. 159–167.
- Jajuli, R. (2015) *Analisis Faktor-faktor yang Mempengaruhi Kemiskinan di Jawa Tengah Tahun 1991-2013*. Universitas Muhammadiyah Surakarta.
- Kolesnikov, A., Trichina, E. and Kauranne, T. (2014) ‘Estimating the Number of Clusters in a Numerical Data Set Via Quantization Error Modeling’, *Pattern Recognition*. Elsevier, pp. 1–12. doi: 10.1016/j.patcog.2014.09.017.
- Mingoti, S. A. and Lima, J. O. (2006) ‘Comparing SOM Neural Network with Fuzzy C-Means, K-Means and Traditional Hierarchical Clustering Algorithms’, *European Journal of Operational Research*, 174(3), pp. 1742–1759. doi: 10.1016/j.ejor.2005.03.039.
- Nidyasofa, N. (2017) ‘Penerapan Algoritma Fuzzy C-Means untuk Pengelompokan Kabupaten/Kota di Jawa Tengah Berdasarkan Status Kesejahteraan Tahun 2015’, *URECOL*. Magelang, 6, pp. 23–30.
- Pal, N. R. and Bezdek, J. C. (1995) ‘On Cluster Validity for the Fuzzy C-Means Model’, *IEEE Transactions on Information Theory*, 3(3), pp. 370–377.
- Pimentel, B. A. and de Souza, R. M. C. R. (2016) ‘Multivariate Fuzzy C-Means Algorithms with Weighting’, *Neurocomputing*. Elsevier, 174, pp. 946–965. doi: 10.1016/j.neucom.2015.10.011.
- Purnamasari, S. B., Yasin, H. and Wuryandari, T. (2014) ‘Pemilihan Cluster Optimum pada Fuzzy C-Means (Studi Kasus: Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan Indikator Indeks Pembangunan Manusia)’, *Jurnal Gaussian*, 3, pp. 491–498.
- Rezaee, B. (2010) ‘A Cluster Validity Index for Fuzzy Clustering’, *Fuzzy Sets and Systems*. Elsevier, 161(23), pp. 3014–3025. doi: 10.1016/j.fss.2010.07.005.
- Scitovski, R., Vidovi, I. and Bajer, D. (2016) ‘A New Fast Fuzzy Partitioning Algorithm’, *Expert Systems with Applications*. doi: 10.1016/j.eswa.2015.12.034.
- Simhachalam, B. and Ganesan, G. (2015) ‘Performance Comparison of Fuzzy and Non-Fuzzy Classification Methods’, *Egyptian Informatics Journal*. Ministry of Higher Education and Scientific Research. doi: 10.1016/j.eij.2015.10.004.
- Sivarathri, S. and Govardhan, A. (2014) ‘Experiments on Hypothesis “Fuzzy K-Means is Better than K-Means For Clustering”’, *International Journal of Data Mining & Knowledge Management Process (IJDMP)*, 4(5), pp. 21–34. doi: 10.5121/ijdkp.2014.4502.
- [TNP2K] ‘Tim Nasional Percepatan Penanggulangan Kemiskinan. Pengembangan, Pengelolaan, dan Pemanfaatan Basis Data Terpadu (BDT).2015’.
- Wang, W. and Zhang, Y. (2007) ‘On Fuzzy Cluster Validity Indices’, *Fuzzy Sets and Systems*, 158, pp. 2095–2117. doi: 10.1016/j.fss.2007.03.004.

- Wu, C. *et al.* (2014) ‘A New Fuzzy Clustering Validity Index with a Median Factor for Centroid-based Clustering’, *IEE Transactions on Fuzzy System*, 6706(c), pp. 1–17. doi: 10.1109/TFUZZ.2014.2322495.
- Wu, K. L. and Yang, M. (2005) ‘A Cluster Validity Index for Fuzzy Clustering’, *Pattern Recognition Letters*, 26(9), pp. 1275–1291. doi: 10.1016/j.patrec.2004.11.022.
- Xie, X. L. and Beni, G. (1991) ‘A Validity Measure for Fuzzy Clustering’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, pp. 841–847.
- Yang, M. and Nataliani, Y. (2017) ‘Robust-Learning Fuzzy C-Means Clustering Algorithm with Unknown Number of Clusters’, *Pattern Recognition*. Elsevier Ltd, 71, pp. 45–59. doi: 10.1016/j.patcog.2017.05.017.
- Yulianto, S. and Hidayatullah, K. H. (2014) ‘Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat’, *Statistika*, 2(1).
- Zhao, Q. and Fränti, P. (2014) ‘WB-index \square : A Sum-of-Squares Based Index for Cluster Validity’, *Data & Knowledge Engineering*. Elsevier B.V., 92, pp. 77–89. doi: 10.1016/j.datak.2014.07.008.