PENERAPAN METODE *FUZZY C-MEANS CLUSTER ING* UNTUK PENGELOMPOKAN KABUPATEN DI JAWA TIMUR BERDASARKAN INDIKATOR KESEJAHTERAAN RAKYAT

Rio Bahtiyar, Sri Lestanti, Saiful Nur Budiman
Program Studi Teknik Informatika, Universitas Islam Balitar
pongahjoly02@gmail.com

ABSTRAK

Setiap negara pasti memiliki tujuan utama dalam pembangunan adalah peningkatankesejahteraan rakyat. Seperti di Indonesia, kesejahteraan rakyat adalah salah satu tujuan negara yang tertulis dalam Pembukaan UUD 1945 alinea IV. Dalam melaksanakan program pembangunan perlu adanya identifikasi berdasarkan karakteristik tingkat kesejahteraan rakyat tiap daerah agar dalam mengambil kebijakan danstrategi pembangunan bisa tepat sasaran dan tepat guna. Istilah kesejahteraan mencakup berbagai aspek kehidupan yang sangat luas dan tidak semuanya dapat diukur. perlu adanya analisis untuk mengidentifikasi karakteristik berdasarkan indikator-indikator kesejahteraan rakyat dengan pengelompokan daerah. Metode yang digunakan untuk mengelompokkan kabupaten di jawa timur menggunakan *Fuzzy C-means cluster* ing. *Cluster* ing adalah metode mengelompokan data kedalam *cluster*, diamana objek dengan kesamaan tinggi berada pada *cluster* yang sama, tetapi objek yang tidak sama berada pada *cluster* yang berbeda. Jumlah *cluster* yang digunakan dalam penelitian ini ada 3 *cluster* yaitu tidak baik, sedang dan baik. Keismpulan yang didapat sebagai berikut cluter 1 tidak terdapat kabupaten yang masuk kedalam *cluster* tersebut, pada *cluster* 2 terdapat 20 kabupaten yaitu: Kab. Pacitan, Kab. Ponorogo, Kab. Trenggalek, Kab. Tulungagung, Kab. Blitar, Kab.Kediri, Kab. Lumajang, Kab. Bondowoso, Kab. Situbondo, Kab. Probolinggo, Kab. Jombang, Kab. Nganjuk, Kab. Madiun, Kab. Magetan, Kab. Ngawi, Kab. Lamongan, Kab. Bangkalan, Kab. Sampang, Kab. Pamekasan, Kab. Sumenep, pada *cluster* 3 terdapat 9 kabupaten yaitu: Kab. Malang, Kab. Jember, Kab. Banyuwangi, Kab. Pasuruan, Kab. Sidoarjo, Kab. Mojokerto, Kab. Bojonegoro, Kab. Tuban, Kab. Untuk mengetahui tingkat homogenitas dalam *cluster* menggunakan silhouette coefficient. Dari hasil penguujian menghasilkan data yang akurat dengan rata-rata nilai silhouette coefficient yang bernilai positif yaitu 2,914.

Kata kunci: *Fuzzy C-Maeans, Silhouette Coefficient, Cluster ing.*

1. PENDAHULUAN

Setiap negara pasti memiliki tujuan utama dalam pembangunan adalah peningkatan kesejahteraan rakyat. Seperti di Indonesia, kesejahteraan rakyat adalah salah satu tujuan negara yang tertulis dalam Pembukaan UUD 1945 alinea IV. Berdasarkan data yang dirilis oleh BPS Jawa Timur persentase penduduk miskin (penduduk yang berada di bawah Garis Kemiskinan) pada bulan September Tahun 2022 persentase jumlah penduduk miskin di Jawa Timur mengalami kenaikan sebesar 0,11%, angka ini melebihi persentase kemiskinan di bulan Maret dan September Tahun 2022.

Tingginya angka kemiskinan di Jawa Timur tersebut perlu dilakukan upaya-upaya pengentasan kemiskinan yang berbasis pada data dengan memahami karakteristik dari masing-masing wilayah Kabupaten, sehingga upaya pengentasan yang dilakukan bisa efektif dan tepat. Aspek yang bisa diukur dalam indikator kesejahteraan rakyat diantaranya adalah kepadatan penduduk, pengangguran terbuka, pengeluaran perkapita disesuaikan, angka harapan hidup, harapan lama sekolah, jumlah presentase penduduk miskin dan PDRB [1].

Dalam melaksanakan program pembangunan perlu adanya identifikasi berdasarkan karakteristik tingkat kesejahteraan rakyat tiap daerah agar dalam mengambil kebijakan dan strategi pembangunan bisa tepat sasaran dan tepat guna. Secara umum kesejahteraan dapat diartikan sebagai suatu keadaaan dimana segenap warga negara selalu berada dalam kondisi serba kecukupan segala kebutuhannya, baik material maupun spiritual. Oleh karenanya, pengentasan kasus kemiskinan di Kecamatan Sanankulon perlu adanya analisis untuk mengidentifikasi karakteristik berdasarkan indikatorindikator kesejahteraan rakyat dengan pengelompokan daerah. Salah satu metode yang dapat digunakan untuk mengelompokkan variabel atau objek adalah analisis klaster (*cluster* analysis) [1].

Klasterisasi (*Cluster* ing) merupakan proses mengelompokkan objek berdasrkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip, untuk memaksimalkan kesamaan antar anggota kelas dan meminimumkan kesamaan antar kelas/*cluster* [2]. Secara umum metode dalam klaster dibedakan menjadi dua yakni Hierarchical *Cluster* ing (metode hirarki) dan Non-Hierarchical *Cluster* ing (metode tak hirarki). Metode hirarki jumlah kelompok yang akan dibentuk belum ditentukan. K-means *Cluster* adalah salah satu metode *cluster* ing non-hierarki yang berusaha mengelompokkan data kedalam klaster sehingga data

yang memiliki karakteristik sama dikelompokkan ke dalam satu klaster yang sama. C-means adalah metode yang lebih baik daripada metode yang lain dikarenakan sederhana, memiliki nilai pengelompokan yang lebih besar dibanding metode yang lain, mudah untuk diterapkan dalam metode tersebut serta memiliki perhitungan yang linear [3].

Metode C-means memiliki kelemahan dalam beberapa hal, seperti langsung tidak dapat melalui penempatan objek tepat pada sebuah partisi. Pada dasarnya objek terletak di antara dua atau lebih dalam partisi yang lainnya, sehingga membutuhkan pembuatan partisi sehingga setiap objek dapat berada pada partisi yang dibuat dengan pembobotan fuzzy, yang bisa disebut dengan metode *Fuzzy C-means Cluster* ing [3]. Kelebihan dari metode *Fuzzy C-means Cluster* ing adalah bisa melakukan *cluster* ing lebih dari satu variabel sekaligus.

Analisis Penelitian Metode *Fuzzy C-means* Pada Pengelompokkan Algen Pulsa. Dari penelitian tersebut diperoleh hasil akhir pengelompokkan algen pulsa dengan metode *Fuzzy C-means* dengan nilai-nilai konvergen pada iterasi ke-55. Algen pulsa terkelompok dalam 3 *cluster* : *cluster* 1 adalah kelompok yang beranggotakan 16 algen dengan jumlah transaksi sedang, *cluster* 2 adalah kelompok yang beranggotakan 29 algen dengan jumlah transaksi kecil, dan *cluster* 3 adalah kelompok yang beranggotakan 3 algen dengan jumlah transaksi besar [4].

Pengelompokan Kualitas Air di Provinsi Jawa Tengah Berdasarkan Indikator Kimia dengan C-means Dan *Fuzzy C-means Cluster* ing. Dari penelitian tersebut dihasilkan perbandingan kedua metode yang terbalik dari nilai iterasi untuk metode c-means sebesar 0,0579 dan metode *Fuzzy C-means* adalah 0,0524. Sehingga pada penelitian itu metode yang lebih baik adalah *Fuzzy C-means* dengan hasil 5 *cluster* [5].

Penelitian ini ditulis bertujuan untuk mengelompokkan desa di Kecamatan Sanan Kulon berdasarkan indikator kesejahteraan rakyat menggunakan metode *Fuzzy C-means Cluster* ing. Dari penelitian sebelumnya yang telah dilakukan, terdapat ketertarikan dalam penelitian ini yaitu ditalmbahkan beberapa indikator yang berpengaruh terhadap kesejahteraan rakyat, yaitu kepadatan penduduk, pengangguran terbuka, pengeluaran per kapita disesuaikan, angka harapan hidup, harapan lama sekolah, PDRB, dan jumlah persentase penduduk miskin. Berdasarkan penjelasan di atas, penulis akan membuat penelitian tentang "Penerapan Metode *Fuzzy C-means Cluster* ing Untuk Pengelompokan Desa di Kecamatan Salamkanci Kulon Berdasarkan Indikator Kesejahteraan Rakyat.

2. TINJAUAN PUSTAKA

2.1. *Cluster ing*

Cluster ing merupakan salah satu metode dalam data *mining*. *Cluster ing* merupakan klasifikasi yang mengidentifikasi sejumlah daerah pada data dalam rangka menemukan pola distribusi keseluruhan dan korelasi yang mungkin tersembunyi di antara atribut [6]. *Cluster ing* dapat memiliki peran penting dalam kehidupan sehari-hari, karena tidak dapat dilepaskan dari sejumlah data yang menghasilkan informasi untuk memenuhi kebutuhan hidup. Analisis *cluster* telah banyak digunakan dalam berbagai hal, sebagai contoh pengelompokan data yang digunakan untuk menganalisis data statistik seperti pengelompokan untuk pembeli mesin, data mining, pengenalan pola, analisis citra, dan bioinformatika [7]. Selain sebagai salah satu metode dari data *mining*, analisis *cluster* dapat digunakan sebagai alat untuk mendapatkan wawasan mengenai distribusi data dengan menghasilkan sejumlah *cluster* yang fokus pada kelompok tertentu untuk analisis lebih lanjut.

2.2. Data

Data merupakan kumpulan fakta yang diperoleh dari suatu pengukuran. Suatu pengambilan keputusan yang baik merupakan hasil dari penarikan kesimpulan yang didasarkan pada data/fakta yang akurat. Untuk mengumpulkan data yang akurat diperlukan suatu alat ukur atau instrumen yang baik. Alat ukur atau instrumen yang baik adalah alat ukur/instrumen yang valid dan reliabel [8]. Syarat dari sebuah data dianggap baik dan bisa digunakan adalah sebagai berikut:

a. Data harus objektif, artinya data menggambarkan keadaan yang sebenarnya.
b. Data harus memiliki akurasi, artinya data merujuk pada perbedaan yang kecil (standar error) yang kecil. Perbedaan kecil ini merupakan sumber kesalahan yang digunakan untuk mengukur tingkat ketelitian.
c. Data harus tepat waktu, artinya keberadaan waktu ini sangat penting, terutama jika data tersebut akan digunakan untuk mengontrol pelaksanaan dan perencanaan, sehingga persoalan yang terjadi dapat segera diketahui untuk segera dikoreksi dan diperbaiki.

2.3. Analisis *Cluster*

Analisis *cluster*, yaitu analisis untuk mengelompokkan elemen yang mirip, sehingga objek penelitian menjadi kelompok (*cluster*) yang berbeda. Perbedaan analisis diskriminan dengan analisis *cluster* adalah analisis diskriminan merupakan analisi dimana kelompok telah ditentukan, kemudian fungsi diskriminan digunakan untuk menentukan elemen yang harus masuk ke dalam kelompok. Sedangkan analisis *cluster*, adalah analisis berdasarkan kriteria tertentu, memisahkan data yang ada untuk ditunjukkan oleh nilai-nilai variabel, membentuk kelompok (*cluster*). Tujuan dari analisis kluster adalah untuk mengelompokkan objek-objek berdasarkan kesamaan karakteristik di antara objek-objek tersebut. Objek bisa berupa produk, benda, atau orang [9]. Klaster yang baik adalah klaster yang memiliki ciri [10]:

a. Homogenitas, kesamaan yang tinggi antara anggota dalam sebuah kluster (*within cluster*).

b. Heterogenitas, perbedaan yang tinggi antara kluster yang satu dengan yang lainnya (*between cluste)*.

2.4. *Fuzzy C-means*

Fuzzy C-means Cluster ing (FCM) atau juga dikenal sebagai fuzzy ISODATA, merupakan varian dari metode *K-means*. Metode *K-means* adalah salah satu metode *cluster ing* non-hierarki yang membagi data ke dalam satu atau lebih *cluster* menjadi sedemikian rupa, sehingga data dalam setiap *cluster* memiliki tingkat kesamaan yang tinggi dan berbeda dari *cluster* lainnya. Algoritma *Fuzzy C-means* diusulkan pertama kali oleh Dunn pada tahun 1973 dan kemudian diperbaiki oleh Bezdek pada tahun 1981. Algoritma *Fuzzy C-means* merupakan algoritma yang bekerja dengan menggunakan model *fuzzy* sehingga memungkinkan semua data dari semua anggota kelompok terbentuk dengan derajat keanggotaan yang berbeda di antara 0 dan 1 [11].

Metode *Fuzzy C-means* merupakan metode yang cocok untuk analisis data dan konstruksi model, dimana setiap objek dapat menjadi anggota dari beberapa kelompok, tidak harus sepenuhnya dimiliki oleh satu kelompok dengan derajat keanggotaan objek antara 0 hingga 1 [12]. Tujuan dari metode *Fuzzy C-means* adalah untuk meminimalkan sisa fungsi serta mendapatkan pusat *cluster* yang nantinya akan digunakan untuk mengetahui data yang masuk ke dalam sebuah *cluster*. Karakteristik algoritma *Fuzzy C-means* adalah pengelompokkan data yang kabur dan menghasilkan pengelompokan data yang keanggotaannya bervariasi dari data tersebut [13].

2.5. *Silhouette Coefficient*

Silhouette Coefficient merupakan salah satu metode validasi untuk menguji kualitas sekelompok *cluster* [14]. Metode ini menggunakan rumus Euclidean distance dalam proses perhitungannya. Secara singkat, metode *Silhouette Coefficient* ditunjukkan pada persamaan (1) berikut:

$$s\,(i) = \frac{b(i)-a(i)}{\max(\mathrm{a(i).b(i)})} \qquad (1)$$

si = *Silhouette Coefficient*
ai = rata-rata jarak dari objek *i* dengan seluruh objek yang berada pada *cluster* yang sama.
bi = nilai terkecil dari rata-rata jarak objek *i* dengan objek lain pada *cluster* yang berbeda

Nilai *Silhouette Coefficient* berada dalam rentang (-1) hingga 1. Semakin tinggi nilainya, maka semakin baik kualitasnya.

2.6. Indikator Kesejahteraan Rakyat

Menurut Undang-Undang No. 11 Tahun 2009, tentang Kesejahteraan Masyarakat, kesejahteraan masyarakat adalah kondisi terpenuhinya kebutuhan material, spiritual, dan rakyat warga negara agar dapat hidup layak dan mampu mengembangkan diri, sehingga dapat melaksanakan fungsi kesejahteraan rakyatnya. Tingkat kesejahteraan rakyat pada suatu daerah dapat dilihat dari berbagai aspek, seperti kepadatan penduduk, pengangguran terbuka, pendapatan domestik regional bruto, pengeluaran per kapita disesuaikan, angka harapan hidup, harapan lama sekolah, dan jumlah persentase penduduk miskin. Salah satu indikator kesejahteraan masyarakat berdasarkan tingkat kepuasan adalah keterbebasan dari kemiskinan. Artinya, kesejahteraan masyarakat dapat tercapai jika kebutuhan dalam keseluruhannya terpenuhi agar keluar dari kerentanan tersebut.

2.7. Jawa Timur

Jawa Timur adalah salah satu dari 38 provinsi di Indonesia yang memiliki struktur tanah berupa pegunungan dan kondisi tanah yang tandus. Secara geografis, Jawa Timur terletak antara 11100 Bujur Timur - 11404' Bujur Timur dan 70 12' Lintang Selatan - 8048' Lintang Selatan. Provinsi ini terdiri dari 29 Kabupaten dan 9 Kota, dengan luas total 47.963 km2. Dari luas tersebut, 88,70% atau 42.541 km2 adalah wilayah daratan Jawa Timur, sementara sisanya, 11,30% atau 5.422 km2, adalah Kepulauan Madura.

Menurut Badan Pusat Statistik, pada tahun 2023, jumlah penduduk Jawa Timur mencapai 41.416.407 jiwa. Penduduk Jawa Timur memiliki sejumlah pekerjaan yang beragam, seperti petani, pedagang, pekerja profesional, serta pegawai pemerintah negeri dan swasta. Namun, mayoritas penduduknya bekerja sebagai petani karena provinsi ini memiliki tanah yang subur, terutama di sektor pertanian. Peningkatan kesejahteraan dan pendapatan masyarakat sangat bergantung pada hasil pertanian.

2.8. Python

Python adalah bahasa pemrograman yang dirancang dengan fokus pada keterbacaan kode. Bahasa python menggabungkan kapabilitas, kemampuan, sintaksis kode yang jelas, dan dilengkapi dengan fungsi pustaka standar yang besar. Python merupakan bahasa pemrograman umum yang dibuat untuk membuat *source code* yang mudah dibaca. Dengan *library* yang lengkap, *programmer* dapat membuat aplikasi yang canggih meskipun menggunakan *source code* yang sederhana [15].

3. METODE PENELITIAN

Metode penelitian yang digunakan pada penelitian ini adalah metode penelitian kuantitatif. Metode kuantitatif adalah suatu pendekatan dalam penelitian yang menggunakan data numerik untuk menjawab pertanyaan penelitian, dan menguji hipotesis dengan pengumpulan data yang sistematis serta analisis statistik. Penelitian kuantitatif melibatkan pengumpulan data atau analisis data numerik yang dihasilkan dari survey, eksperimen, atau data sekunder [16]. Dalam penelitian ini, subjek penelitian adalah pengelompokan kabupaten di Jawa

Timur. Tahapan penelitian pada metode ini dapat dilihat pada gambar 1.

Gambar 1. Tahapan Penelitian

Pada gambar 1 dapat dilihat bahwa tahapan penelitian metode kuantitatif terdapat 5 tahap. Pertama tahap perencanaan, pada tahap ini hal pertama yang dilakukan adalah mengidentifikasikan suatu masalah. Kemudian penulis menentukan langkah selanjutnya untuk menangani masalah dari penelitian tersebut. Kabupaten di Jawa Timur dipilih sebagai lokasi oleh penulis yang ditemukan permasalahan yaitu bagaimana penerapan metode *Fuzzy C-means Cluster ing* untuk pengelompokan Kabupaten di Jawa Timur berdasarkan Indikator Kesejahteraan dan Rakyat. Penulis memberikan solusi untuk mengatasi permasalahan ini dengan menggunakan metode yang telah dipelajari di masa kuliah atau sebelumnya.

Kedua tahap studi literatur, penulis mencari bahan referensi untuk penelitiannya dari berbagai sumber seperti buku, jurnal ilmiah, dan referensi lainnya. Referensi ini penting untuk membantu mengarahkan dan mendukung pada penelitian yang dilakukan. Dilakukanya literature ini menentukan kualitas suatu tulisan atau argumen, apabila literature yang digunakan tidak valid, logis, maka laporan atau karya ilmiah yang dibuat tidak akan berkualitas.

Ketiga tahap pengumpulan data, penulis melakukan pengumpulan data yang dibutuhkan dalam penelitian. Pengumpulan data ini dilakukan dengan cara observasi melalui pengamatan website Badan Pusat Statistik Jawa Timur untuk menentukan variable apa saja yang dibutuhkan. Penulis mengambil 7 variabel data kesejahteraan rakyat kabupaten di Jawa Timur.

Keempat tahap implementasi metode, pada tahap ini dilakukannya perhitungan menggunakan pengujian *Silhouette Coefisient* dan mengolah data sesuai metode *Fuzzy C-means Clustering*. Tahap awalnya menggunakan perhitungan *Fuzzy C-means Clustering*

Pada gambar 2 merupakan *flowchart* algoritma *Fuzzy C-means Clustering*, yang menggambarkan proses berjalannya algoritma *Fuzzy C-means*. Langkah pertama yang dilakukan adalah penginputan data yang akan di-*cluster*.

Selanjutnya, langkah kedua adalah menginisiasi beberapa parameter yang dibutuhkan dalam algoritma *Fuzzy C-means*. Kemudian, langkah ketiga adalah membilang angka acak sebagai inisiasi awal keanggotaan setiap data.

Setelah itu, langkah keempat adalah menghitung pusat *cluster* dengan persamaan tertentu.
Kemudian, langkah kelima adalah menghitung perubahan matriks partisi dengan persamaan lain. Proses berikutnya adalah melakukan pemeriksaan kondisi berhenti, jika telah mencapai error terkecil dan maksimal iterasi, maka proses iterasi akan berhenti.

Yang terakhir adalah hasil dan pembahasan, tahap terakhir ini berdasarkan implementasi perhitungan menggunakan metode *Fuzzy C-means Cluster ing* yang dibuat untuk perangkingan alternatif desa yang sudah di-*cluster*.

Hingga mengeluarkan output sehingga dilakukan pengujian menggunakan *Silhouette Coeficient.* Dengan hasil output yang valid dengan perhitungan sebelumnya.

Gambar 2. *Flowchart* Algoritma *Fuzzy C-means*

4. HASIL DAN PEMBAHASAN

4.1. Hasil dan Pembahasan Data

Pada tahap ini menjelaskan tentang data atau nilai yang digunakan untuk proses perhitungan yang menggunakan metode *Fuzzy C-means.* Data ini disajikan dalam bentuk tabel dibawah ini.

Tabel 1. Data Indikator Kesejahteraan Rakyat di Jawa Timur Tahun 2022

KABUPATEN	X1	X2	X3	X4	X5	X6	X7
Pacitan	592,92	2.04	9184	12,66	13,8	72,48	11722,44
Ponorogo	964,25	5,51	10199	13,76	9,32	73,2	15093,71
Trenggalek	739,67	5,37	10042	12,5	10,96	74,26	13545,41
Tulungagung	1105,34	6,65	11162	13,33	6,71	74,54	28818,91
Blitar	1240,32	5,45	11001	12,64	8,71	73,98	27037,33
Kediri	1656,02	6,83	11565	13,61	10,65	72,97	30800,71
Malang	2685,9	6,57	10326	13,38	9,55	72,95	72136,46
Lumajang	1137,23	4,97	9466	12,02	9,06	70,61	23626,58
Jember	2567,72	4,06	9840	13,44	9,39	69,68	57161,13
Banyuwangi	1731,73	5,26	12320	13,11	7,51	71,06	57923,55
Bondowoso	781,42	4,32	10851	13,31	13,47	67,29	14410,2
Situbondo	691,26	3,38	10263	13,18	11,78	69,62	14318
Probolinggo	1159,97	3,25	11254	12,58	17,12	67,78	24734,19
Pasuruan	1619,04	5,91	10726	12,76	8,96	70,55	113352,08
Sidoarjo	2103,4	8,8	14808	14,95	5,36	74,36	151613,88
Mojokerto	1133,58	4,83	13051	12,96	9,71	72,93	63699,84
Jombang	1335,97	5,47	11579	13,58	9,04	72,86	30086,17
Nganjuk	1117,03	4,74	12349	13,07	10,7	71,5	19543,18
Madiun	757,67	5,84	11848	13,18	10,79	71,9	14169,62
Magetan	678,34	4,33	12031	14,05	9,84	72,97	13939,15
Ngawi	877,43	2,48	11563	12,84	14,15	72,81	14264,44
Bojonegoro	1315,13	4,69	10323	12,84	12,21	72,16	61782,87
Tuban	1209,54	4,54	10703	12,24	15,02	71,97	47890,26
Lamongan	1371,51	6,05	11648	14,01	12,53	72,86	29447,44
Gresik	1332,66	7,84	13384	13,96	11,06	72,99	108796,88
Bangkalan	1086,62	8,05	8971	11,91	19,44	70,54	16959,91
Sampan	984,16	3,11	8944	12,39	21,61	68,38	14308,28
Pamekasan	857,82	1,4	8967	13,67	13,93	68,03	12031,56
Sumenep	1136,63	1,36	9388	13,51	18,76	71,99	14912,62

Pada tabel 1, merupakan tabel data sekunder yang diperoleh dari Website Badan Pusat Statistik Jawa Timur. Tabel tersebut berupa kepadatan penduduk (*X*1), pengangguran terbuka (*X*2, pengeluaran terbuka (*X*3), harapan lama sekolah (*X*4), persentase penduduk miskin (*X*5), angka harapm hidup (*X*6), dan PDRB (*X*7) pada tahun 2022. Setelah di peroleh 29 data kabupaten di Jawa Timur, maka proses selanjutnya adalah membangkitkan bilangan random. Hasil membangkitkan bilangan random tersebut, dapat dilihat pada tabel 2 di bawah ini.

Tabel 2. Iterasi

i	μ_1	μ_2	μ_3
1	0,3	0,5	0,2
2	0,2	0,3	0,5
3	0,5	0,2	0,3
4	0,3	0,5	0,2
5	0,2	0,3	0,5
6	0,5	0,2	0,3
7	0,3	0,5	0,2
8	0,2	0,3	0,5
9	0,5	0,2	0,3
10	0,3	0,5	0,2
11	0,2	0,3	0,5
12	0,5	0,2	0,3
13	0,3	0,5	0,2
14	0,2	0,3	0,5
15	0,5	0,2	0,3
16	0,3	0,5	0,2
17	0,2	0,3	0,5
18	0,5	0,2	0,3
19	0,3	0,5	0,2
20	0,2	0,3	0,5
21	0,5	0,2	0,3
22	0,3	0,5	0,2
23	0,2	0,3	0,5
24	0,5	0,2	0,3
25	0,3	0,5	0,2
26	0,2	0,3	0,5
27	0,5	0,2	0,3
28	0,3	0,5	0,2
29	0,2	0,3	0,5

Tabel 2 merupakan iterasi 1 dari proses membangkitkan bilangan random, dimana pada *cluster* 1 sampai dengan *cluster* 3 nilainya tidak boleh lebih dari 1. Data pada μ1, μ2, dan μ3 merupakan data dari 3 *cluster* yang akan dikelompokkan. Proses dari membangkitkan bilangan random diperoleh dari variabel yang dipangkatkan dengan jumlah *cluster y*aitu ada 3 *cluster* yang akan menghasilkan bilangan random 0,3 pada data kabupaten. Berikut proses perhitungan data kabupaten untuk dikelompokkan menggunakan metode *Fuzzy C-means cluster ing:*

1. Menentukan nilai matriks n × m. (n = jumlah sampel data, m = atribut setiap data). X_{ij} = data sampel ke I (i = 1,2,3,…,n), atribut ke-j (j = 1,2,…,m)

Tabel 3. Menetukan Nilai Matrik

DATA	X1	X2	X3	X4	X5	X6	X7
1	592,92	2,04	9184	12,66	13,8	72,48	11722,44
2	964,25	5,51	10199	13,76	9,32	73,2	15093,71
3	739,67	5,37	10042	12,5	10,96	74,26	13545,41
4	1105,34	6,65	11162	13,33	6,71	74,54	28818,91
5	1240,32	5,45	11001	12,64	8,71	73,98	27037,33
6	1656,02	6,83	11565	13,61	10,65	72,97	30800,71
7	2685,9	6,57	10326	13,38	9,55	72,95	72136,46
8	1137,23	4,97	9466	12,02	9,06	70,61	23626,58
9	2567,72	4,06	9840	13,44	9,39	69,68	57161,13
10	1731,73	5,26	12320	13,11	7,51	71,06	57923,55
11	781,42	4,32	10851	13,31	13,47	67,29	14410,2
12	691,26	3,38	10263	13,18	11,78	69,62	14318
13	1159,97	3,25	11254	12,58	17,12	67,78	24734,19
14	1619,04	5,91	10726	12,76	8,96	70,55	113352,08
15	2103,4	8,8	14808	14,95	5,36	74,36	151613,88
16	1133,58	4,83	13051	12,96	9,71	72,93	63699,84
17	1335,97	5,47	11579	13,58	9,04	72,86	30086,17
18	1117,03	4,74	12349	13,07	10,7	71,5	19543,18
19	757,67	5,84	11848	13,18	10,79	71,9	14169,62
20	678,34	4,33	12031	14,05	9,84	72,97	13939,15
21	877,43	2,48	11563	12,84	14,15	72,81	14264,44
22	1315,13	4,69	10323	12,84	12,21	72,16	61782,87
23	1209,54	4,54	10703	12,24	15,02	71,97	47890,26
24	1371,51	6,05	11648	14,01	12,53	72,86	29447,44
25	1332,66	7,84	13384	13,96	11,06	72,99	108796,88
26	1086,62	8,05	8971	11,91	19,44	70,54	16959,91
27	984,16	3,11	8944	12,39	21,61	68,38	14308,28
28	857,82	1,4	8967	13,67	13,93	68,03	12031,56
29	1136,63	1,36	9388	13,51	18,76	71,99	14912,62

Tabel 3 merupakan data kabupaten yang dihasilkan dari pengkodean data yang sudah diurutkan.

2. Menentukan inisialisasi awal pada *cluster ing* :
   a. Jumlah *cluster* : 3
   b. Pangkat (w) : 5
   c. Max_Iter : 3
   d. Error terkecil (e) : 0,01
   e. Fungsi objektif P_0 : 0
   f. Iterasi (t) : 1

3. Membangkitkan bilangan random μ_{ik} $I = 1,2,\ldots, n;k = 1,2,\ldots,c$: sebagai elemen-elemen matriks partisi awal U. Matriks random sebanyak data, dimana c = 3 (ada 3 *cluster*)

Tabel 4. Nilai Matrik Random

i	μ_1	μ_2	μ_3
1	0,3	0,5	0,2
2	0,2	0,3	0,5
3	0,5	0,2	0,3
4	0,3	0,5	0,2
5	0,2	0,3	0,5
6	0,5	0,2	0,3
7	0,3	0,5	0,2
8	0,2	0,3	0,5
9	0,5	0,2	0,3
10	0,3	0,5	0,2
11	0,2	0,3	0,5
12	0,5	0,2	0,3
13	0,3	0,5	0,2
14	0,2	0,3	0,5
15	0,5	0,2	0,3
16	0,3	0,5	0,2
17	0,2	0,3	0,5
18	0,5	0,2	0,3
19	0,3	0,5	0,2
20	0,2	0,3	0,5
21	0,5	0,2	0,3
22	0,3	0,5	0,2
23	0,2	0,3	0,5
24	0,5	0,2	0,3
25	0,3	0,5	0,2
26	0,2	0,3	0,5
27	0,5	0,2	0,3
28	0,3	0,5	0,2
29	0,2	0,3	0,5

Tabel 4 merupakan nilai matrik random pada setiap *cluster* yang totalnya tidak melebihi nilai 1 (satu).

4. Menghitung pusat *cluster* ke-k: V_{kj} dengan $k = 1,2, \ldots, c$: *dan* $j = 1,2, \ldots, m$

Tabel 5. Nilai Centroid tiap *cluster*

v_{kj}	1	2	3	4	5	6	7
1	1336,863	4,968859	11213,7	13,31555	11,85373	71,81289	38835,62561
2	1257,322	4,849229	11132,87	13,15488	11,31066	71,67707	44529,31952
3	1135,011	4,988828	10545,54	13,00274	12,13649	71,61176	32289,87332

Pada tabel 5, V_{kj} merupakan data *cluster*, dimana niali centroid tiap *cluster* dihasilkan dari total nilai matrik yang sudah dipangkatkan dan dikali dengan atribut kemudian dibagi dengan nilai matrik random. Kemudian dari total nilai $\mu 1^{w} * x_1$ dibagi dengan nilai total $\mu 1^{w}$ maka akan dihasilkan nilai centroid tiap *cluster*.

5. Hitung fungsi obyektif pada iterasi ke t, P_t:

Tabel 6 Nilai fungsi objective

Data	μ_1^w	P_{i1}	μ_2^w	P_{i2}	$\mu 3^w$	P_{i3}	P *Cluster*
1	0,00243	1797709,113	0,03125	33766590,46	0,00032	136053,437	35700353
2	0,00032	180751,0485	0,00243	2107813,893	0,03125	9245540,511	11534105,5
3	0,03125	20041391,83	0,00032	307667,3997	0,00243	854788,3793	21203847,6
4	0,00243	243949,8981	0,03125	7713779,593	0,00032	3977,128202	7961706,62
5	0,00032	44561,39436	0,00243	743549,3567	0,03125	868992,4738	1657103,22
6	0,03125	2024535,83	0,00032	60422,53417	0,00243	8573,905105	2093532,27
7	0,00243	2701074,992	0,03125	23901440,34	0,00032	508865,2599	27111380,6
8	0,00032	75011,00637	0,00243	1068513,316	0,03125	2381815,087	3525339,41
9	0,03125	10600818,36	0,00032	52144,41805	0,00243	1509345,58	12162308,4
10	0,00243	888720,7217	0,03125	5657492,458	0,00032	211388,8542	6757602,03
11	0,00032	191053,2832	0,00243	2205145,588	0,03125	9996908,409	12393107,3
12	0,03125	18826081,59	0,00032	292416,3005	0,00243	785533,91	19904031,8
13	0,00243	483286,7812	0,03125	12245979,81	0,00032	18429,09632	12747695,7
14	0,00032	1776966,228	0,00243	11510591,07	0,03125	205354631,6	218642189
15	0,03125	397888790,2	0,00032	3674024,18	0,00243	34645299,63	436208114
16	0,00243	1510600,148	0,03125	11600105,74	0,00032	317716,2644	13428422,2
17	0,00032	24539,65643	0,00243	507407,7899	0,03125	186397,9941	718345,44
18	0,03125	11672990,24	0,00032	200257,8611	0,00243	402726,2561	12275974,4
19	0,00243	1480233,61	0,03125	28827262,99	0,00032	105658,3555	30413155
20	0,00032	198699,5627	0,00243	2276667,8	0,03125	10598880,72	13074248,1
21	0,03125	18877383,23	0,00032	293213,528	0,00243	792223,3098	19962820,1
22	0,00243	1281508,73	0,03125	9323257,532	0,00032	278374,0231	10883140,3
23	0,00032	26324,30174	0,00243	27903,71305	0,03125	7606325,647	7660553,66
24	0,03125	2760245,235	0,00032	72877,27672	0,00243	22722,37593	2855844,89
25	0,00243	11905268,23	0,03125	129231018,5	0,00032	1875653,754	143011940
26	0,00032	154764,5881	0,00243	1858403,867	0,03125	7421542,844	9434711,3
27	0,03125	18964585,64	0,00032	293816,6823	0,00243	791998,9692	20050401,3
28	0,00243	1758676,334	0,03125	33154844,17	0,00032	132149,7357	35045670,2
29	0,00032	184218,7326	0,00243	2138905,698	0,03125	9478403,162	11801527,6

Nilai fungsi objektif pada tabel 6 diperoleh dari nilai $(x_i - v_i)$ merupakan nilai *cluster* yang dikalikan dengan nilai matrik pangkat.

6. Hitung perubahan matrik partisi (updet nilai μ)

Tabel 7. Matriks Hasil Baru

i	μ_1	μ_2	μ_3
1	0,329459	0,481199	0,189342
2	0,326857	0,501941	0,171202
3	0,328119	0,491909	0,179972
4	0,279127	0,686317	0,034556
5	0,294376	0,64684	0,058784
6	0,251951	0,734327	0,013722
7	0,320646	0,220633	0,458721
8	0,312402	0,58602	0,101577
9	0,301989	0,145064	0,552947
10	0,302916	0,149947	0,547137
11	0,327252	0,497403	0,175345
12	0,327499	0,496766	0,175735
13	0,306755	0,604417	0,088827
14	0,329336	0,280932	0,389731
15	0,330961	0,29844	0,370598
16	0,313059	0,186937	0,500003
17	0,26311	0,716425	0,020465
18	0,320612	0,537139	0,142249
19	0,327182	0,495472	0,177345
20	0,327325	0,493885	0,17879
21	0,327167	0,496262	0,176571
22	0,311017	0,175949	0,513035
23	0,243997	0,034059	0,721943
24	0,271427	0,699838	0,028734
25	0,328897	0,277616	0,393487
26	0,325485	0,514687	0,159828
27	0,327864	0,496052	0,176084
28	0,329321	0,482766	0,187913
29	0,327242	0,500346	0,172413

Nilai matriks baru pada tabel 7 diperoleh dari perhitungan nilai dari tiap *cluster* di bagi dengan nilai dari semua *cluster*.

7. *Cluster* yang diperoleh

Tabel 8. Hasil *Cluster*

i	*CLUSTER*
1	2
2	2
3	2
4	2
5	2
6	2
7	3
8	2
9	3
10	3
11	2
12	2
13	2
14	3
15	3
16	3
17	2
18	2
19	2
20	2
21	2
22	3
23	3
24	2
25	3
26	2
27	2
28	2
29	2

Pada tabel 8 merupakan hasil dari nilai *cluster* yang di golongkan menjadi 3, yaitu *cluster* 1 dengan keterangan tidak baik, *cluster* 2 keterangan sedang dan *cluster* 3 dengan keterangan baik.

4.2. Pengujian Menggunakan *Silhouette coefficient*

1. Menghitung Rata-rata Jarak

Hitungan rata-rata jarak dari sauatu objek i dengan seluruh objek yang ada pada satu cluter (a_i) dan hitungan rata-rata jarak dari objek i dengan yang berada pada *cluster* lain dan diambil nilai terkecil (b_i) adalah sebagai berikut:

Tabel 9. Hasil perhitungan nilai (a_i) dan nilai (b_i)

a_i	b_i
2,708	4,613
2,46	8,248
2,549	8,0505
2,336	8,379
2,317	2,318
2,385	2,096
4,45	9
2,325	2,103
3,457	2,725
3,495	4,365
2,493	7,331
2,497	1,147
2,312	4,734
7,748	5,281
1,131	4,284
3,849	3,339
2,363	6,434
2,377	4,252
2,525	3,048
2,545	6,037
2,512	3,233
3,716	3,339
3,029	9,409
2,348	1,791
7,358	5,281
2,424	2,237
2,522	1,351
2,684	4,613
2,48	7,652

2. Hitungan Nilai *Silhouette Coefficient*

Tabel 10. Nilai *Silhouette Coefficient*

s_i
2,662
2,378
2,468
2,252
2,085
2,176
3,55
2,114
3,184
3,058
2,419
2,382
1,839
7,22
7,029
3,515
2,298
1,952
2,494
2,485
2,479
3,382
2,088
5,57
6,83
2,201
2,387
2,638
2,403

Pada tabel 10 merupakan hasil perhitungan *Silhouette Coefficient* yang didapatkan dengan cara menghitung nilai maksimum hasil b_i (jarak antar semua titik) dikurangi dengan hasil a_i (nilai rata-rata dalam satu *cluster*). Dari hasil tabel diatas didapatkan hasil rata-rata dari perhitungan *silhouette coefficient*, yaitu:

Tabel 11. Hasil rata-rata

S_i
2,914

Hasil rata-rata *silhouette coefficient* 2,914 bernilai positif, sehingga bisa disimpulkan bahawa dalam proses tersebut sudah baik.

5. KESIMPULAN DAN SARAN

Penelitian implementasi pengelompokan kabupaten di Jawa Timur berdasarkan indicator kesejahteraan dengan metode *Fuzzy C-means Cluster ing* menghasilkan 3 *cluster*. *Cluster* 1 tidak terdapat kabupaten yang masuk ke dalam *cluster* tersebut, *cluster* 2 terdapat 20 kabupaten, dan *cluster* 3 terdapat 9 kabupaten. Hasil ini diperoleh menggunakan data random dan menghasilkan data yang berubah-ubah. Data pengujian menggunakan metode *silhouette coefficient* mendapat tingkat homogenitas yang akurat, dengan hasil rata-rata yang bernilai positif yaitu 2,914. Dalam pengujian menggunakan *silhouette coefficient* jika nilainya 1 atau positif, maka objek I berada pada *cluster* yang tepat. Jika nilainya 0 maka objek I berada pada dua *cluster* atau tidak jelas. Namun jika nilainya -1 atau *negative*, maka objek lebih tepat dimasukkan dalam *cluster* lain. Berdasarkan hasil dari pembahasan, terdapat beberapa saran yang di harapkan dari peneliti yaitu perlu dilakukannya pencarian variabel data, agar hasil proses *cluster ing* yang dihasilkan dapat lebih maksimal. Kemudian untuk mendapatkan hasil *cluster* yang baik, perlu dilakukan proses *preprocessing* data untuk menghindari data ambigu atau tidak valid.

DAFTAR PUSTAKA

[1] Soemartini and E. Supartini, "Analisis K-Means Cluster Untuk Pengelompokan Kabupaten / Kota Di Jawabarat Berdasarkan Indikator Masyarakat," *Konf. Nas. Penelit. Mat. dan Pembelajarannya II (KNPMP II)*, no. Knpmp Ii, 2017.

[2] A. Fauzan, A. Y. Badharudin, and F. Wibowo, "Sistem Klasterisasi Menggunakan Metode K-Means dalam Menentukan Posisi Access Point Berdasarkan Posisi Pengguna Hotspot di Universitas Muhammadiyah Purwokerto (Clustering System Using K-Means Method in Determining Access Point Position at Muhammadiyah Un," *Juita*, vol. III, no. 1, pp. 25–29, 2014.

[3] A. S. Rizal and R. F. Hakim, "Metode K-Means Cluster Dan Fuzzy C-Means Cluster (Studi Kasus: Indeks Pembangunan Manusia Di Kawasan Indonesia Timur Tahun 2012)," *Pros. Semin. Nas. Mat. dan Pendidik. Mat. UMS 2015*, pp. 643–657, 2015, [Online]. Available: https://publikasiilmiah.ums.ac.id/xmlui/handle/11617/5803

[4] R. Nuraeni *et al.*, *No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title*, vol. 2, no. 1. 2017. [Online]. Available: http://i-lib.ugm.ac.id/jurnal/download.php?dataId=2227%0A???%0Ahttps://ejournal.unisba.ac.id/index.php/kajian_akuntansi/article/view/3307%0Ahttp://publicacoes.cardiol.br/portal/ijcs/portugues/2018/v3103/pdf/3103009.pdf%0Ahttp://www.scielo.org.co/scielo.ph

[5] R. Mustafidah and R. M. Atok, "Pengelompokan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Kemiskinan Dengan C-Means Dan Fuzzy C-Means Clustering," *Skripsi. Surabaya Inst. Teknol. Sepuluh …*, 2017, [Online]. Available: https://core.ac.uk/download/pdf/291465651.pdf

[6] N. Agustina and P. Prihandoko, "Perbandingan Algoritma K-Means dengan Fuzzy C-Means Untuk Clustering Tingkat Kedisiplinan Kinerja Karyawan," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 3, 2018, doi: 10.29207/resti.v2i3.492.

[7] B. A. Akash *et al.*, "THE INTERNATIONAL ADVISORY BOARD EDITORIAL BOARD SUPPORT TEAM 2 B Language Editor," *Jordan J. Mech. Ind. Eng.*, 2011.

[8] D. A. Setyawan, "Data dan Metode Pengumpulan Data Penelitian," *Metodol. Penelit.*, pp. 9–17, 2013.

[9] J. Supranto, *Analisis multivariat : arti dan interpretasi*. Jakarta : Rineka Cipta, 2004. [Online]. Available: https://opac.perpusnas.go.id/DetailOpac.aspx?id=649801

[10] M. K. Norfai, SKM., *Analisis Data dan Penelitian*. 2021. [Online]. Available: https://www.google.co.id/books/edition/ANALISIS_DATA_PENELITIAN_Analisis_Univar/IY5-EAAAQBAJ?hl=id&gbpv=1&dq=SPSS+22+from+Essential+to+Expert+Skills.&pg=PA185&printsec=frontcover

[11] D. J. Bora and D. A. K. Gupta, "A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm," *Int. J. Comput. Trends Technol.*, vol. 10, no. 2, pp. 108–113, 2014, doi: 10.14445/22312803/ijctt-v10p119.

[12] R. Subhamita and N. Sudipta, "Classification of Power Quality Disturbances using Features of Signals," *Int. J. Sci. Res. Publ.*, vol. 2, no. 11, pp. 7–15, 2012.

[13] H. Vol and N. Barakoska, "International Journal of Cognitive Re- search in Science , Engineering and Education," pp. 1–11, 2016.

[14] M. Tanzil Furqon and L. Muflikhah, "Clustering the Potential Risk of Tsunami Using Density-Based Spatial Clustering of Application With Noise (Dbscan)," *J. Enviromental Eng. Sustain. Technol.*, vol. 3, no. 1, pp. 1–8, 2016, doi:

10.21776/ub.jeest.2016.003.01.1.
[15] M. Ilham, "Pengertian Python, Fungsi, Kelebihan dan Kekurangan," *5 Mei 2019*. 2019.
[16] D. P. Saputra and A. Widiyarta, "Efektivitas Program SIPRAJA Sebagai Inovasi Pelayanan Publik di Kecamatan Sidoarjo Kabupaten Sidoarjo," *JPAP J. Penelit. Adm. Publik*, vol. 7, no. 2, pp. 194–211, 2021, doi: 10.30996/jpap.v7i2.4497.