Implementasi fuzzy c-means dalam pengelompokan tingkat kemiskinan pada kabupaten/kota di provinsi jawa tengah

Rinaldo Dwi Faturahman\*1), Nutriana Hidayati2)

1. S1-Teknik Informatika, Fakultas Teknologi Informasi dan Kominikasi, Universitas Semarang
2. Fakultas Teknologi Informasi dan Komunikasi, Universitas Semarang

|  |  |
| --- | --- |
| **Article Info** | **ABSTRAK** |
| **Kata Kunci:** *Clustering; Fuzzy C Means; Data Mining;* Kemiskinan; Provinsi Jawa Tengah**Keywords:** *Clustering; Fuzzy C Means; Data Mining; Proverty; Central Java Province*Article history:Received 17 August 2018Revised 15 February 2019Accepted 4 April 2019Available online 4 April 2019DOI : <https://doi.org/10.29100/jipi.v4i1.781> \* Corresponding author.Corresponding AuthorE-mail address: rinaldodwifaturahman@email.ac.id  | Kemiskinan merupakan salah satu permasalahan bagi negara berkembang khususnya di Indoensia. Setiap tahunya kemiskinan di Provinsi Jawa Tengah terdapat kenaikan ataupun penurunan. Kemikinan di Provinsi Jawa Tengah cukup tinggi. Dibuktikan pada data Badan Pusat Statistik Provinsi Jawa Tengah kemiskinan pada tahun 2022 sebesar 10,93% dan pada tahun 2021 sebesar 10,77%. Tujuan dari penelitian ini untuk mengklasifikasikan tingkat kemiskinan di Provinsi Jawa Tengah dengan metode *clustering* dengan menggunakan *fuzz c means.* clustering sendiri adalah salah satu teknik data mining. Dimana data ini merupakan sebuah metode dari data mining untuk mengelompokan data menjadi beberapa kelompok berbeda berdasarkan karakteristik yang sama. Data penelitian yang dugunakan diambil dari Badan Pusat Statistik Provinsi Jawa Tengah dari tahun 2021 – 2023 dengan total 108 data dan 5 atribut indikator kemiskinan yaitu garis kemiskinan(Rp/kapita/bln), jumlah penduduk miskin(ribu jiwa), pengeluaran, rata-rata pendidikan dan jumlah pengangguran. Hasil penelitian ini menghasilkan 4 cluster dengan beberapa data di dalamnya. Pada *cluster 0* dengan tingkat kemiskinan rendah terdapat 13 kabupaten/kota. Kemudian pada *cluster 1* terrdapat 31 kabupaten/ kota dengan tingkat kemiskinan sedang. *Cluster 2* terdapat 20 kabupaten/kota dengan tingkat kemiskinan tinggi. Dan *cluster 3* dengan 44 kabupaten/kota dengan kemsikinan yang sangat tinggi. |
|  |
| **ABSTRACT** |
| Poverty is one of the problems for developing countries, especially in Indonesia. Every year poverty in Central Java Province increases or decreases. Poverty in Central Java Province is quite high. As evidenced by the data from the Central Java Provincial Statistics Agency, poverty in 2022 was 10.93% and in 2021 it was 10.77%. The purpose of this study is to classify the level of poverty in Central Java Province using the clustering method using fuzz c means. clustering itself is one of the data mining techniques. Where this data is a method of data mining to group data into several different groups based on the same characteristics. The research data used was taken from the Central Java Provincial Statistics Agency from 2021 - 2023 with a total of 108 data and 5 poverty indicator attributes, namely the poverty line (IDR / capita / month), the number of poor people (thousand people), expenditure, average education and the number of unemployed. The results of this study produced 4 clusters with some data in them. In cluster 0 with a low poverty rate there are 13 districts/cities. Then in cluster 1 there are 31 districts/municipalities with a medium poverty level. Cluster 2 has 20 districts/municipalities with a high poverty rate. And cluster 3 with 44 districts/cities with very high poverty. |

.

# Pendahuluan

s

alah satu permasalahan sosial yang kompleks dan beragam yang sering dihadapi banyak negara, termasuk Indonesia adalah kemiskinan. Kemiskinan adalah ketidakmampuan seseorang atau sekelompok orang untuk memenuhi kebutuhan dasar seperti pangan, kesehatan, pendidikan, perumahan, sumber daya alam, lingkungan hidup dan keamanan[1]. Pendapatan bukan satu-satunya faktor yang berkontribusi terhadap kemiskinan, faktor ini juga mencakup kendala pada pendidikan, kesehatan, pangan, air bersih, dan aspek lain yang berkaitan dengan kualitas hidup manusia. Sebagai salah satu provinsi terbesar di Indonesia, kemiskinan di Jawa Tengah bersifat kompleks dan berbeda-beda antar daerah/kota.

 Menurut data BPS Provinsi Jawa Tengah (Badan Pusat Statistik ) sendiri tingkat kemiskinan pada Provinsi Jawa Tengah Tahun 2022 sebesar 10,93% dan pada tahun 2023 sebesar 10,77%. Meskipun mengalami penurunan dibandingkan tahun sebelumnya, namun masih terdapat jarak yang signifikan dalam tingkat kemiskinan antar daerah. Sebagai contoh pada tahun 2022 presentase populasi yang paling terdapat di Kabupaten Wonosobo dengan presentase 16,17%, sementara di Kota Semarang hanya 4,25 %. Perbedaan tingkat kemiskinan antar daerah ini terpengaruhi oleh macam-macam faktor, seperti kondisi geografis, sumber daya alam, instruktur, akses terhadap pendidikana kan kesehatan, serta karakteristik sosial-ekonomi masyarakat setempat. Oleh karena itu, upaya penanggulangan kemiskinan tiap daerah memerlukan stategi dan kebijakan yang disesuaikan dengan kondisi spesifik di setiap daerah.

 Salah satu pendekatan yang digunakan untuk mengidentifikasi pola dan faktor-faktor yang mempengaruhi kemiskinan di setiap daerah adalah dengan memanfaatkan teknik data mining. Data mining suatu proses mencari serta menganalisis data untuk mengidentifikasi pola yang menarik yang bertujuan menadapatkan informasi serta pengetahuan yang akurat dan potensial, agar mudah dipahami serta bermanfaat dalam membuat keputusan[2]. Proses pengumpulan dan ekstrasi informasi ini dapat dilakukan menggunakan software dengan bantuan perhitungan statistika, matematika ataupun menggunakan Artifical intelligence (AI). Data mining sendiri memiliki banyak sekali metode algortima yang dapat digunakan salah satunya adalah clustering. Clustering adalah teknik untuk membagi data dalam beberapa grup yang memiliki karakteristik objek yang sama didalamnya[3]. Tujuan utama dari metode clustering ialah mengelompokan beberapa data atau obyek ke dalam cluster sehingga pada setiap cluster dapat berisikan data yang mirip atau sama. Terdapat banyak algoritma dalam metode clustering yang dapat digunakan pada penelitian sebelumnya seperti Fuzzy C-Means, K-Means, K-Medoids, Hierrarchical Agglomerative, dan sebagainya. Fuzzy C Means (FCM) merupakan teknik pengelompokan data di mana tiap titik-titik data dalam suatu cluster ditentukan oleh derajat keanggotaanya[4]. Fuzzy C Means sendiri memiliki kelebihan, yaitu tingkat akurasi yang tinggi. Sehingg metode ini memungkinkan identifikasi pola-pola yang mungkin tersembunyi dalam data dan memberikan pemahaman yang lebih baik tentang distribusi kemiskinan di berbagai wilayah di Provinsi Jawa Tengah.

 Beberapa penelitian sebelumnya telah dilakukan mengenai kemiskinan itu sendiri. Pada penelitian bertajuk “Implementasi Data Mining untuk Menentukan Tingkat Kemiskinan Menggunakan Fuzzy C-Means” yang dilakukan oleh Nurrahman Fitriani Kahar, Lillyan Hadjaratie, Sitti Suhada dan Indhitiya R. Padiku, hasilnya dibagi menjadi 3 cluster yaitu 50% sangat miskin, 34% miskin, 16% hampir miskin[5].

 Dalam analisis berjudul “Penerapan Algoritma Clustering Pengelompokan Tingkat Kemiskinan di Provinsi Banten” yang diterbitkan oleh Tb Ai Munandar. Dalam penelitian ini, tiga kelompok berbeda dihasilkan. Tingkat kemiskinan lebih rendah terdapat di Kota Serang, Kota Cilegon, dan Kota Tangerang Sur, sedangkan kelompok miskin sedang terdapat di Kabupaten Bandelan, Lebak, dan Sealand. Kota Serang, Kota Cilegon, dan Kota Tangerang Selatan mempunyai kelompok dengan tingkat kemiskinan yang rendah[6].

 Penelitian yang ditulis oleh Nurhafizah Sepriyanti, Rahma Sani Nahampun dan rekannya ini diberi judul “Penerapan K-means clustering pada pengelompokan tingkat kemiskinan di Provinsi Riau”. Dengan hasil sebanyak 3 *cluster*. Klaster 1 masuk dalam kategori tingkat kemiskinan rendah karena jumlah penduduk miskinnya sedikit, angka penganggurannya rendah, dan rata-rata lama pendidikannya sedang. Penduduk miskin yang termasuk dalam kelompok kemiskinan moderat 2 termasuk dalam kemiskinan sedang, dengan tingkat pengangguran yang lebih tinggi dan rata-rata lama pendidikan yang lebih tinggi. Sedangkan klaster 3 tergolong memiliki tingkat kemiskinan tinggi karena memiliki jumlah penduduk miskin terbanyak, tingkat pengangguran sedang, dan rata-rata lama pendidikan lebih rendah.[7].

 Penelitian yang dilakukan oleh Edy Widodo, Putri Ermayani, dkk dengan judul Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative clustering. Hasil penelitian ini terdapat 3 cluster dengan cluster 1 sebanyak 25 daerah yang memiliki tingkat kemiskinan rendah, cluster 2 sebanyak 7 daerah yang memiliki tingkat kemiskinan sedang, dan cluster 3 sebnayak 2 daerah yang memiliki tingkat kemiskinan tinggi[8].

 Studi yang diselesaikan oleh F. Fajriani dengan judul Persebaran Tingkat Kemiskinan di Sulawesi Selatan Menggunakan K-Means Clustering Analysis. Hasil penelitian tersebut adalah 24 Kabupate/Kota di Sulawesi Selatan, yang memiliki tingkat kemiskinan tinggi berjumlah 7 wilayah, sedangkan terdapat 18 daerah dengan kategori kemiskinan yang rendah[9].

 Tujuan dari studi ini adalah menggunakan metode *fuzzy c means* untuk mengelompokan tingkat kemiskinan di masing-masing kabupaten atau kota yang berada di Provinsi Jawa Tengah dengan berdasarkan indikator yang mempengaruhi. Dengan demikian pengelompokan ini, diharapkan dapat diperoleh informasi yang berguna bagi pemerintah dalam memahami pola dan karakteristik di setiap wilayah, sehingga dapat merumuskan kebijakan dan program penanggulangan kemiskinan yang lebih tepat sasaran dan efisien.

# metode penelitian

Penelitian ini merupakan dari sebuah proses data mining yang menggunakan sebuah metode clustering dan menggunakan algoritma Fuzzy C-Means.

## Metode Fuzzy C-Means Clustering

Fuzzy C-Means (FCM) merupakan algoritma yang termasuk dalam kategori pembelajaran tanpa pengawasan[10]. Metode ini memungkinkan data menjadi anggota cluster yang berbeda-beda dan mempunyai derajat keanggotaan yang berbeda-beda. Keuntungan dari fuzzy c-means adalah menempatkan pusat cluster lebih cepat dibandingkan metode clustering lainnya[11]. Pada gambar 1 menujukan tahapan fuzzy c-means.



Gambar. 1. PROSES FUZZY C-MEANS

Pada gambar 1 tahapan fuzzy c-means clustering ini, dimana pada gambar ini terdapat beberapa tahapan mulai dari data source, dimana data source ini digunakan untuk pengambilan data untuk penelitian ini. Kemudian terdapat data understanding, yang digunakan untuk mengidentifikasi jenis data. Data distribution digunakan untuk melihat persebaran data pada data set. Data preprocessing untuk mengubah data untuk di analisis guna melakukan modeling. Yang terkhir terdapat modeling yang menerapkan fuzzy c-means serta menentukan jumlah cluster dari data yang sudah siap.

## Data Source

Penelitian ini menggunakan data indikator kemiskinan Provinsi Jawa Tengah dari tahun 2021-2023 yang diperoleh dari Badan Pusat Statistik (BPS)[12]. Populasinya adalah kabupaten/kota yang berada di Provinsi Jawa Tengah. Dalam dataset ini terdiri dari 108 data dan 5 atribut data yang dijadikan indikator kemiskinan.

## Data Understanding

Data understanding merupakan tahapan awal, untuk melakukan pengecekan pada data yang digunakan, mengumpulkan data awal, dan mengidentifikasi kualitas data. Dalam tahapan data understanding sebagai berikut:

### Missing Value Identification

Untuk menidentifikasi adanya missing value (data hilang) pada data yang ada.

### Duplicated Data Identification.

Untuk mengidentifikasi adanya Duplicate data (data yang sama) pada data yang ada.

## Data Distribution

Data distribusi (data distribution) adalah cara data tersebar atau terdistribusi di dalam suatu rentangan nilai. Ini menggambarkan frekuensi atau probabilitas kemunculan nilai-nilai tertentu dalam suatu set data. Distribusi data penting untuk memahami karakteristik dasar dari data dan untuk melakukan analisis statistik yang lebih lanjut.

## Data Preprocessing

Data preprocessing adalah memepersiapkan data mentah atau data yang diperloeh dari hasil crawling yang memiliki keberagaman bentuk data menjadi data yang teratur atau mengeliminasi data yang tidak diperloeh atau tidak sesuai menjadi bentuk lain sehingga data mudah diproses atau dikenali atau mudah diterapkan[14]. Dalam proses data preprocessing ada beberapa tahapan seperti :

### Correlation Data

Suatu ukuran statistik yang menggambarkan sejauh mana dua variable berhubungan satu sama lain.

### Scalling Data

Proses transformasi data sehingga data tersebut berada dalam skala atau rentang tertentu.

## Modeling

Modeling ini dilakukan untuk menentukan dan menerapkan teknik data mining dan algoritma yang akan digunakan dengan menentukan parameter deengan nilai yang optimal[15]. Modeling ini menggunakan fuzzy c-means yang merupaka pengembangan dari metode clustering k-means, dimana setiap data menjadi anggota dari semua cluster dengan derajat keanggotaan tertentu. Dasar dari menghitung teknik fuzzy c-means adalah sebagai berikut[16]:

### Input data yang akan di-cluster X, berupa matriks berukuran n x p (n=jumlah sampel data, p=atribut setiap data). Xkj = data sampel ke-k (𝑘 = 1,2,…,𝑛), atribut ke-j (𝑗 =1,2,3, . . , 𝑚).

### Menentukan:

#### Jumalh Cluster = c;

#### Pangkat pembobot = m;

#### Maksimum iterasi = MaxIter;

#### Error terkecil yang diharapkan = ξ;

#### Fungsi object awal = $P\_{0}$= 0;

#### Iterasi awal = t = 1

### Bangkitakan bilangan random ($μ\_{ik}, I=1,2,…,c;k=1,2,…,n)$ sebagai elemen matriks partisi awal U

$$U\_{0}= \left[\begin{matrix}μ\_{11}(X\_{1})&μ\_{12}(X\_{2})&\cdots \\\vdots &\vdots &\\μ\_{11}(X\_{1})&μ\_{12}(X\_{2})&\cdots \end{matrix} \begin{matrix}μ\_{1c}(X\_{c})\\\vdots \\μ\_{1c}(X\_{c})\end{matrix} \right]$$

Matriks partisi pada *fuzzy clustering* harus memenuhi kondisi sebagai berikut:

$$μ\_{ik}=\left[0,1\right];(1 \leq i \leq c ;1\leq k \leq n$$

$$\sum\_{i=1}^{n}μ\_{ik}=1;1 \leq i \leq c$$

$$0 <\sum\_{i=1}^{c}μ\_{ik}=c;1 \leq k \leq n$$

Hitung jumlah setiap kolom (atribut):

$$Q\_{j}=\sum\_{i=1}^{c}μ\_{ik}$$

Dengan j = 1,2,3,..,m

Kemudian hitung :

$$μ\_{ik}= \frac{μ\_{ik}}{Q\_{j}}$$

### Hitunglah pusat cluster ke-k: Vij, dimana 𝑖=1,2,3,..,𝑐 dan 𝑗=1,2,3,..,𝑚

$$V\_{ij}= \frac{\sum\_{k=1}^{n}(\left(μ\_{ik}\right)^{m}\* X\_{kj}}{\sum\_{k=1}^{n}\left(μ\_{ik}\right)^{m}}$$

$$V=\left[\begin{matrix}V\_{11}&\cdots &V\_{1m}\\\vdots &\ddots &\vdots \\V\_{c1}&\cdots &V\_{cm}\end{matrix}\right]$$

### Hitung fungsi objektif pada iterasi ke-t, Pt dengan menggunakan persamaan sebagai berikut:

$$P\_{t}= \sum\_{k=1}^{n}\sum\_{i=1}^{c}([\sum\_{j=1}^{m}\left(X\_{kj}- V\_{ij}\right](μ\_{ik})^{m}$$

### Hitung perubahan matriks partisi:

$$μ\_{ik}= \frac{[\sum\_{j=1}^{p}(X\_{kj}- V\_{ij})^{2}]^{\frac{-1}{p-1}}}{\sum\_{i=1}^{c}[\sum\_{j=1}^{p}(X\_{kj}- V\_{ij})^{2}]^{\frac{-1}{p-1}} }$$

### Cek kondisi berhenti:

#### Jika $(\left|P\_{t}- P\_{t-1}\right|< ξ$ atau (t < iterasi maksimal) maka berhenti;

#### Jika tidak: maka t=t+1 kemudian ulangi langkah ke-4;

# hasil dan pembahasan penelitian

Jumlah data yang digunakan untuk studi ini adalah sebanyak 108 data dan 5 atribut data yang di ambil *website* BPS (Badan Pusat Statistik) Provinsi Jawa Tengah mulai tahun 2021-2023. 5 atribut data yang dijadikan indikator kemiskinan ini adalah garis kemiskinan (Rp/Kapita/Bln), jumlah penduduk miskin(ribu jiwa), pengeluaran, rata-rata pendidikan, serta jumlah pengangguran. Data dapat dilihat dalam Tabel 1.

TABEL I

DATA KEMISKINAN

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kabupaten/Kota | Garis Kemiskinan (Rp/Kapita/Bln) | Jumlah Penduduk Miskin (Ribu Jiwa) | Pengeluaran | Rata-rata pendidikan | Jumlah Pengangguran | Tahun |
| PROVINSI JAWA TENGAH | 409193.00 | 4109.75 | 1048609.00 | 7.75 | 1128223.00 | 2021 |
| Kabupaten Cilacap | 363367.00 | 201.71 | 830026.00 | 7.09 | 82714.00 | 2021 |
| Kabupaten Banyumas | 417086.00 | 232.91 | 1023973.00 | 7.63 | 52390.00 | 2021 |
| Kabupaten Purbalingga | 384183.00 | 153.08 | 900102.00 | 7.25 | 30450.00 | 2021 |
| Kabupaten Banjarnegara | 328679.00 | 150.19 | 975274.00 | 6.75 | 29281.00 | 2021 |
| Kabupaten Kebumen | 390599.00 | 212.92 | 920590.00 | 7.55 | 37408.00 | 2021 |
| Kabupaten Purworejo | 376127.00 | 88.80 | 876976.00 | 8.21 | 14898.00 | 2021 |
| ------------- | --------------------- | ----------------------------- | --------------------- | ---------------- | ------------------- | --------- |
| Kota Semarang | 642456.00 | 80.53 | 1787507.00 | 10.81 | 55656.00 | 2023 |
| Kota Pekalongan | 565998.00 | 21.36 | 1348573.00 | 9.29 | 9065.00 | 2023 |
| Kota Tegal | 623617.00 | 19.22 | 1750336.00 | 9.24 | 8805.00 | 2023 |

Data yang di kumpulakan harus melewati tahapan data understanding terlebih dahulu dimana pada proses data understanding melibatkan pengecekan pada data untuk mengetahui apakah terdapat data yang misssing value (data yang hilang) dan duplicated data (data yang sama).

## Data Distribution

Tahapan ini digunakanuntuk melihat tersebarnya data-data dari indikator kemiskinan. Disini penulis menggunakan hisplot untuk melihat penyebarannya. Hisplot sendiri merupakan alat yang kuat dalam eksplorasi data dan analisis deskriptif, serta memberikan gambaran visual tentang bagaimana data itu tersebar. Berikut adalah gambaran dari *data distribution* menggunakan hisplot pada indikator kemiskinan.



Gambar. 2 DATA DISTRIBUTION PADA INDIKATOR KEMISKINAN

## Data Preprocessing

Pada tahapan ini, proses yang dilakukan adalah untuk persiapan data dilakukan dengan menyesuaikan *dataset* untuk memenuhu syarat yang akan digunakan saat tahapan modeling. pada tahapan ini terdiri dari:

### Correlation Matrix

Corelation matrix ini digunakan untuk merepresentasikan visual dari korelasi antar setiap pasangan dalam sebuah dataset. Pada penelitian ini korelasi matriks dapat digunakan untuk memahami hubungan antar variable-variable dalam indikator tingkat kemiskinan. Berikut adalah visualisai dari korelasi matrik.



Gambar. 3 CORRELATION MATRIX

### Scaling data

*Scaling data* ini digunakan untuk proses transformasi fitur dalam dataset sehingga berada dalam skala atau rentang yang sama. Dimana *scaling ini* menggunakan *StandartScaler,* hasil dari ini adalah untuk menghitung meaan dan standar deviasi dari dataset tersebut. Berikut hasil dari *scaling data:*

TABEL II HASIL DATA SETELAH MELAKUKAN SCALING DATA

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Kabupaten/Kota | Garis Kemiskinan (Rp/Kapita/Bln) | Jumlah Penduduk Miskin (Ribu Jiwa) | Pengeluaran | Rata-rata pendidikan | Jumlah Pengangguran | Tahun |
| PROVINSI JAWA TENGAH | -5.36858927e-01 | 6.20066633e+00 | -4.14186451e-01 | -2.96183898e-01 | 6.04456324e+00 | 2021 |
| Kabupaten Cilacap | -1.22073862e+00 | -2.47907132e-02 | -1.17424426e+00 | -8.26369439e-01 | 1.23091340e-01 | 2021 |
| Kabupaten Banyumas | -4.19068536e-01 | 2.49104822e-02 | -4.99850862e-01 | -3.92581269e-01 | -4.86553541e-02 | 2021 |
| Kabupaten Purbalingga | -9.10093144e-01 | -1.02257673e-01 | -9.30575671e-01 | -6.97839611e-01 | -1.72917406e-01 | 2021 |
| Kabupaten Banjarnegara | -1.73840150e+00 | -1.06861405e-01 | -6.69187248e-01 | -1.09949532e+00 | -1.79538296e-01 | 2021 |
| Kabupaten Kebumen | -8.14344614e-01 | -6.93332856e-03 | -8.59334703e-01 | -4.56846183e-01 | -1.33509231e-01 | 2021 |
| Kabupaten Purworejo | -1.03031605e+00 | -2.04654879e-01 | -1.01098950e+00 | 7.33393580e-02 | -2.60999604e-01 | 2021 |
| ------------- | --------------------- | -------------------- | --------------------- | ---------------- | ------------------- | ------ |
| Kota Semarang | 2.94421797e+00 | -2.17828882e-01 | 2.15511306e+00 | 2.16194907e+00 | -3.01576394e-02 | 2023 |
| Kota Pekalongan | 1.80320470e+00 | -3.12085924e-01 | 6.28849743e-01 | 9.40915698e-01 | -2.94036092e-01 | 2023 |
| Kota Tegal | 2.66307605e+00 | -3.15494917e-01 | 2.02586189e+00 | 9.00750127e-01 | -2.95508660e-01 | 2023 |

## Modeling

Pada tahap ini, pola data mining yang dihasilkan melalui pemodelan *clustering* adalah algoritma *fuzzy c means*. Langkah pertama adalah menentukan berapa banyak *cluster* yang ada. Untuk menghitung jumlah *cluster* penulis menggunakan *elbow* seperti yang di tunjukan pada gambar 4 di bawah ini.



Gambar. 4 MENANTUKAN JUMLAH CLUSTER DENGAN ELBOW

Pada gambar 4 dapat dilihat hasil dari grafik tersebut. *Cluster* yang diperoleh sebanyak 4 *cluster* dengan skor 80,135. Skor dan *cluster* didapatkan memalui elbow. jadi penelitian ini menggunakan 4 *cluster* sebagai nilai k untuk pemodelan *fuzzy c means.* selah menentukan *cluster,* tahap selanjutnya adalah pemodelan dengan menggunakan algoritma *fuzzy c means* pada dataset ditunjukan pada gambar 5.

**

Gambar. 5 SOURCODE PEMODELAN FUZZY C MEANS

Kemudian setelah tahapan pemodelan *cluster* menggunakan *fuzzzy c means*. terdapat 4 titik *cluster* yang terbentuk dari hasil pemodelan tersebut dengan tiap titik *cluster* dengan berbgai warna yang merupakan penyebaran datanya. Berikut gambar 6 penyebaran data *cluster*.



Gambar. 6 PENYEBARAN DATA CLUSTER

Setelah melakukan pemodelan *cluster* selesai, hasil *cluster* dari setiap data telah di ketahui. Didapatkan jumlah tiap pengelompokan pada Gambar 7.



Gambar. 7 JUMLAH PADA TIAP CLUSTER

Setelah pengujian keseluruhan selesai, hasil informasi *cluster* yang menunjukan area mana saja yanga termasuk dalam c*cluster* penyebaran serta tiap area dengan masing masing *cluster. Cluster 0* menunjukan tingkat kemiskinan yang rendah, *cluster 1* menunjukan tingkat kemiskinan yang sedang, *cluster 2* menunjukan tingkat kemiskinan yang tinggi, dan *cluster 3* menunjukan tingkat kemiskinan yang saat tinggi. Untuk penyebaran *cluster* berdasarkan kabupaten/kota dapat tdilihat pada tabel 3

TABEL III PENYEBARAN DATA KABUPATEN/KOTA BERDASARKAN CLUSTERNYA

|  |  |  |
| --- | --- | --- |
| Cluster | Kabupaten/Kota | Kriteria |
| Cluster 0(13 data) | Kota Semarang(2021), Kota Semarang(2022), Kota Semarang(2023), Kota Surakarta(2021), Kota Surakarta(2022), Kota Surakarta(2023), Kota Tegal(2023), Kota Salatiga(2021),Kota Salatiga(2022), Kota Salatiga(2023),Kota Magelang(2021), Kota Magelang(2022), Kota Magelang(2023) | Rendah |
| *Cluster 1**(31 data)* | Kabupaten Banjarnegara(2021), Kabupaten Banjarnegara(2022), Kabupaten Banjarnegara(2023), Kabupaten Batang(2021), Kabupaten Batang(2022), Kabupaten Batang(2023), Kabupaten Blora(2021), Kabupaten Blora(2022), Kabupaten Boyolali(2021), Kabupaten Boyolali(2022), Kabupaten Cilacap(2021), Kabupaten Cilacap(2022), Kabupaten Grobogan(2021), Kabupaten Kebumen(2021), Kabupaten Kendal(2021), Kabupaten Magelang(2021), Kabupaten Magelang(2022), Kabupaten Pemalang(2021), Kabupaten Purbalingga(2021), Kabupaten Purbalingga(2022), Kabupaten Purworejo(2021), Kabupaten Purworejo(2022), Kabupaten Sragen(2021), Kabupaten Tegal(2021), Kabupaten Temanggung(2021), Kabupaten Temanggung(2022), Kabupaten Temanggung(2023), Kabupaten Wonogiri(2021), Kabupaten Wonogiri(2022), Kabupaten Wonogiri(2023), Kabupaten Wonosobo(2021), Kabupaten Wonosobo(2022) | Sedang |
| *Cluster 2**(20 data)* | Kabupaten Banyumas(2023), Kabupaten Demak(2022), Kabupaten Demak(2023), Kabupaten Karanganyar(2023), Kabupaten Klaten(2022), Kabupaten Klaten(2023), Kabupaten Kudus(2021), Kabupaten Kudus(2022), Kabupaten Kudus(2023), Kabupaten Pati(2022), Kabupaten Pati(2023), Kabupaten Semarang(2023), Kabupaten Sukoharjo(2021), Kabupaten Sukoharjo(2022), Kabupaten Sukoharjo(2023) , Kota Pekalongan(2021), Kota Pekalongan(2022), Kota Pekalongan(2023), Kota Tegal(2021), Kota Tegal(2022) | Tinggi |
| *Cluster 3**(44 Data)* | KabupatenBanyumas(2021), Kabupaten Banyumas(2022), Kabupaten Blora(2023), Kabupaten Boyolali(2023), Kabupaten Brebes(2021), Kabupaten Brebes(2022), Kabupaten Brebes(2023), Kabupaten Cilacap(2023), Kabupaten Demak(2021), Kabupaten Grobogan(2022), Kabupaten Grobogan(2023), Kabupaten Jepara(2021), Kabupaten Jepara(2022), Kabupaten Jepara(2023), Kabupaten Karanganyar(2021), Kabupaten Karanganyar(2023), Kabupaten Kebumen(2022), Kabupaten Kebumen(2023), Kabupaten Kendal(2022), Kabupaten Kendal(2023), Kabupaten Klaten(2021), Kabupaten Magelang(2023), Kabupaten Pati(2021), Kabupaten Pekalongan(2021), Kabupaten Pekalongan(2022), Kabupaten Pekalongan(2023), Kabupaten Pemalang(2022), Kabupaten Pemalang(2023), Kabupaten Purbalingga(2023), Kabupaten Purworejo(2023), Kabupaten Rembang(2021), Kabupaten Rembang(2022), Kabupaten Rembang(2023), Kabupaten Semarang(2021), Kabupaten Semarang(2022), Kabupaten Sragen(2022), Kabupaten Sragen(2023), Kabupaten Tegal(2022), Kabupaten Tegal(2023), Kabupaten Wonosobo(2023), Provinsi Jawa Tengah(2021), Provinsi Jawa Tengah(2022), Provinsi Jawa Tengah(2023) | Sangat Tinggi |

# Kesimpulan

Menurut hasil perhitungan algoritma *fuzzy c-means* dengan dataset sebanyak 108 data serta 4 atribut yang menjadi indikator kemiskinan pada Provinsi Jawa Tengah yang digunakan. Pengelompokan tingkat kemiskinan pada Provinsi Jawa Tengah memperoleh hasil 4 *cluster. Dimana cluster 0* teradpat 13 kabputaen/kota yang artinya tingkat kemiskinannya rendah. Kemudian *cluster 1* terdapat 31 kabupaten/kota yang tingkat kemiskinannya sedang,. *Cluster 2* terdapat 20 kabupaten/kota yang tingkat kemiskinanya tinggi dan *cluster 3* terdapat 44 kabupaten/kota dengan tingkat kemiskinan yang sangat tinggi.

Daftar Pustaka

[1] G. Haumahu and M. Y. Matdoan, “Algoritma K-Medoids Clustering Untuk Mengelompokkan Tingkat Kemiskinan Pada Kabupaten Dan Kota Di Kepulauan Maluku Dan Papua,” *Var. J. Stat. Its Appl.*, vol. 4, no. 2, pp. 81–88, 2023, doi: 10.30598/variancevol4iss2page81-87.

[2] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, “Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review,” *Fakt. Exacta*, vol. 13, no. 1, p. 35, 2020, doi: 10.30998/faktorexacta.v13i1.5548.

[3] A. Bahauddin, A. Fatmawati, and F. Permata Sari, “Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means,” *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.36595/misi.v4i1.216.

[4] W. WIJAYANTI, I. RAHMI HG, and F. YANUAR, “Penggunaan Metode Fuzzy C-Means Untuk Pengelompokan Provinsi Di Indonesia Berdasarkan Indikator Kesehatan Lingkungan,” *J. Mat. UNAND*, vol. 10, no. 1, p. 129, 2021, doi: 10.25077/jmu.10.1.129-136.2021.

[5] I. R. P. Nurrahmah Fitirani Kahar, Lillyan Hadjaratie, Sitti Suhada, “T Ingkat K Emiskinan Di I Ndonesia T Ahun 2007,” *J. Informatics*, vol. 14, no. 2, pp. 27–38, 2020.

[6] T. A. Munandar, “Penerapan Algoritma Clustering Untuk Pengelompokan Tingkat Kemiskinan Provinsi Banten,” *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 109–114, 2022, doi: 10.30656/jsii.v9i2.5099.

[7] S. Nurhafiza, H. Rahma, Sani Nahampun Muhammad, and I. A. A. Rahmadeyan, “Penerapan K-Means Clustering Untuk Mengelompokkan Tingkat Kemiskinan di Provinsi Riau,” *SENTIMAS Semin. Nas. Penelit. dan Pengabdi. Masy.*, vol. 1, no. 1, pp. 59–65, 2022.

[8] E. Widodo, P. Ermayani, L. N. Laila, and A. T. Madani, “Pengelompokkan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative Clustering (Indonesian Province Grouping Based on Poverty Level Using Hierarchical Agglomerative Clustering Analysis),” *Semin. Nas. Off. Stat.*, vol. 2, pp. 557–566, 2021.

[9] F. Fajriani, “Persebaran Tingkat Kemiskinan di Sulawesi Selatan Menggunakan K-Means Clustering Analysis,” *Pros. Semant.*, vol. 2, no. 1, pp. 276–281, 2019, [Online]. Available: https://journal.uncp.ac.id/index.php/semantik/article/view/1526/1337%0Ahttps://journal.uncp.ac.id/index.php/semantik/article/view/1526.

[10] M. K. Sukardi and Indah Manfaati Nur, “Penerapan Algoritma Fuzzy C-Means Untuk Pengelompokkan Kemiskinan Di Kabupaten/kota Provinsi Aceh,” *J. Ilm. Mat. Dan Terap.*, vol. 20, no. 2, pp. 115–126, 2023, doi: 10.22487/2540766x.2023.v20.i2.16494.

[11] P. Algoritma, K. Fuzzy, P. Jawa, and C. Zonyfar, “411-Article Text-836-1-10-20220713,” vol. III, pp. 1–8, 2022.

[12] “BPS Provinsi Jawa Tengah.” https://jateng.bps.go.id/ (accessed May 28, 2024).

[13] F. N. Dhewayani, D. Amelia, D. N. Alifah, B. N. Sari, and M. Jajuli, “Implementasi K-Means Clustering untuk Pengelompokkan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM,” *J. Teknol. dan Inf.*, vol. 12, no. 1, pp. 64–77, 2022, doi: 10.34010/jati.v12i1.6674.

[14] N. Nofiyani and W. Wulandari, “Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining,” *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1621, 2022, doi: 10.30865/mib.v6i3.4332.

[15] D. K. Sitinjak, B. N. Sari, and I. Maulana, “Clustering Daerah Penyumbang Sampah Berdasarkan Provinsi di Indonesia Menggunakan Algoritma K-Means,” *J. Ilm. Wahana Pendidik.*, vol. 8, no. 16, pp. 137–146, 2022, [Online]. Available: https://doi.org/10.5281/zenodo.7059032.

[16] W. Sanusi, A. Zaky, and B. N. Afni, “Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-faktor Penyebab Gizi Buruk,” *J. Math. Comput. Stat.*, vol. 2, no. 1, p. 47, 2020, doi: 10.35580/jmathcos.v2i1.12458.