



Klasterisasi Kabupaten Berdasarkan Index Pendidikan Penduduk Menggunakan Fuzzy C-Means

^{1*}Asmaul Husnah Nasrullah

¹Universitas Negeri Makassar, Jl. Mallengkeri, Parang Tambung, Kota Makassar 90241, Sulawesi Selatan, Indonesia

Email: ¹asmaul.husnah@unm.ac.id

Diterima: 30-07-2024

Disetujui: 28-08-2024

Dipublikasikan: 30-08-2024

ABSTRAK

Penelitian ini menganalisis penggunaan algoritma *Fuzzy C-Means* untuk mengelompokkan kabupaten di Sulawesi Selatan berdasarkan tingkat pendidikan penduduk. Ditemukan dua kelompok berbeda dengan beberapa kabupaten masuk ke masing-masing kelompok setelah 17 iterasi mencapai solusi optimal. Hasil pengelompokan divisualisasikan dengan grafik titik sebar. Algoritma *Fuzzy C-Means* dieksekusi menggunakan *Python* dengan parameter tertentu. Penelitian bertujuan meningkatkan kualitas pendidikan dengan alokasi sumber daya yang tepat dan identifikasi kabupaten berdasarkan pendidikan tertinggi. Data yang digunakan mencakup indikator pendidikan dan Upah Minimum Kabupaten atau Kota. Hasil penelitian diharapkan memberikan masukan untuk kebijakan pendidikan yang lebih tepat sasaran di Sulawesi Selatan. Algoritma *Fuzzy C-Means* efektif untuk analisis dan pengelompokan data pendidikan dalam pengambilan keputusan kebijakan pendidikan.

Kata Kunci: Fuzzy, C-Means, Klasterisasi, Pendidikan, Kabupaten

ABSTRACT

This study analyses the use of Fuzzy C-Means algorithm to cluster districts in South Sulawesi based on the education level of the population. Two distinct groups were found with several districts falling into each group after 17 iterations to reach the optimal solution. The clustering results were visualised with a point spread graph. The Fuzzy C-Means algorithm was executed using Python with certain parameters. The research aims to improve the quality of education with proper resource allocation and identification of districts based on the highest education. The data used includes education indicators and district minimum wage. The results are expected to provide input for a more targeted education policy in South Sulawesi. Fuzzy C-Means algorithm is effective for analysing and clustering education data in education policy decision making.

Keywords: Fuzzy, C-Means, Clustering, Education, Distric

This is an open access article under the [CC BY-SA](#) license



1. PENDAHULUAN

Sulawesi Selatan adalah salah satu provinsi pada Indonesia yang memiliki banyak tantangan dalam meningkatkan kualitas pendidikan. Tantangan pendidikan di Sulawesi Selatan tidak terlepas dari sejumlah faktor penyebab. Salah satunya penyebabnya adalah disparitas ekonomi dan infrastruktur antara kabupaten-kabupaten di provinsi ini [1]. Kabupaten yang lebih terpencil atau kurang berkembang cenderung menghadapi kesulitan dalam menyediakan akses pendidikan yang sama dengan kabupaten yang lebih maju secara ekonomi [2]. Selain itu, dari tahun 2007 hingga 2021, jumlah populasi Provinsi Sulawesi Selatan bertambah sebanyak 1.255.926 jiwa [3].

Pertumbuhan populasi memiliki dampak yang signifikan terhadap jumlah orang miskin [4]. Akibat dari ketidaksetaraan akses dan kualitas pendidikan antar kabupaten di Sulawesi Selatan sangatlah signifikan. Terbatasnya akses pendidikan, terutama pada tingkat Pendidikan Anak Usia Dini (PAUD) dan pendidikan dasar, dapat mengakibatkan rendahnya tingkat partisipasi pendidikan dan tingkat kelulusan yang rendah pula. Akibatnya, potensi dan masa depan anak-anak di daerah tersebut terhambat dalam mencapai kesempatan yang setara dengan anak-anak di daerah yang lebih berkembang [5]. Disparitas ini juga berdampak pada pertumbuhan ekonomi dan kesejahteraan sosial masyarakat di wilayah tersebut, memperkuat siklus kemiskinan dan ketidaksetaraan yang sulit dipecahkan. Kualitas dari penduduk pada provinsi bisa dilihat dari tingkat Pendidikannya.

Dalam menghadapi masalah ini, metode *Fuzzy C-Means* muncul sebagai metode yang tepat dan efisien sebagai jawaban untuk menghadapi permasalahan ini. Metode ini memungkinkan kita untuk menentukan kabupaten yang memiliki tingkat pendidikan terakhir dan mengetahui akurasi pengelompokan kabupaten menggunakan *Fuzzy C-Means* (FCM) [6]. Salah satu metode pengelompokan data adalah *Fuzzy C-Means* (FCM), di mana derajat keanggotaan setiap titik data menentukan apakah titik data tersebut ada dalam sebuah *cluster* atau tidak. Jim Bezdek pertama kali mempresentasikan metode ini pada tahun 1981 [7].

Pada penelitian terdahulu yang dilakukan oleh Hanniva dkk, [8]. Berdasarkan faktor Indeks Pembangunan Manusia (IPM), penelitian ini mengelompokkan kabupaten dan kota di Indonesia dengan menggunakan metodologi K-Means dan Fuzzy C-Means. Temuan penelitian menunjukkan bahwa, dengan total kesesuaian sebesar 81,91%, teknik pengelompokan Fuzzy C-Means (FCM) menunjukkan tingkat penerapan yang tinggi dalam mengelompokkan wilayah berdasarkan indikator IPM. Hasil pengelompokan yang konsisten dengan kategorisasi BPS, menunjukkan kemampuan pendekatan FCM dalam mengelompokkan wilayah berdasarkan tingkat IPM. [8].

Pada penelitian terdahulu yang dilakukan oleh Martin dan Yessica Nataliani [9]. Pada penelitian ini menggunakan algoritma Fuzzy C-Means (FCM) untuk menganalisa performa kerja karyawan berdasarkan tiga kriteria: ekspektasi kinerja, disiplin, dan waktu yang dihabiskan untuk mengerjakan tugas-tugas yang berhubungan dengan pekerjaan. Berdasarkan penelitian tersebut, algoritma Fuzzy C-Means (FCM) digunakan untuk mengevaluasi karyawan berdasarkan performa kerja yang akurat. FCM memiliki kemampuan untuk memberikan hasil karyawan yang sebanding dengan hasil pengawasan dengan berfokus pada tiga kriteria kinerja utama: kehadiran, kedisiplinan, dan jadwal kerja. Dengan demikian, dapat dikatakan bahwa penggunaan FCM dalam penelitian ini memberikan hasil yang akurat untuk seleksi karyawan berdasarkan pengalaman kerja mereka. Oleh karena itu, dapat dikatakan bahwa penggunaan FCM dalam penelitian tersebut memberikan hasil yang akurat dalam pengelompokan karyawan berdasarkan kinerja mereka [9].

Pada penelitian terdahulu yang dilakukan oleh Devi Lolita Pardosi dan Irma Damayanti Siagian [10]. Dalam penelitian ini telah dibuktikan bahwa penggunaan *Fuzzy C-Means* (FCM) dapat menghasilkan rekomendasi pekerjaan yang tepat dengan tingkat kesalahan yang rendah. Temuan pengujian menunjukkan bahwa akurasi hasil rekomendasi meningkat seiring dengan jumlah lamaran dan minat untuk setiap pekerjaan. Selain itu, hasil prediksi *Click Through Rate* (CTR) dari kebanyakan pekerjaan juga cukup akurat, menunjukkan bahwa penggunaan FCM dalam klasterisasi data lowongan pekerjaan memberikan hasil yang memuaskan dalam memberikan rekomendasi yang relevan dan efektif.

Dengan mengetahui tingkat pendidikan terakhir di setiap kabupaten, pemerintah dapat mengambil keputusan yang lebih tepat dalam alokasi sumber daya pendidikan. Kabupaten-kabupaten dengan tingkat pendidikan terakhir yang rendah mungkin membutuhkan lebih banyak dukungan dan investasi untuk meningkatkan kualitas pendidikan mereka. Penerapan metode *Fuzzy C-Means* ini diharapkan dapat membuat pengelompokan kabupaten di Sulawesi Selatan dapat dilakukan secara lebih akurat dan lebih relevan dengan kondisi nyata di lapangan,

sehingga memungkinkan penyusunan kebijakan yang lebih terarah dan menuju arah yang lebih baik dalam mencapai visi Indonesia Emas pada tahun 2045.

2. METODE PENELITIAN

Metode yang diterapkan bertujuan untuk mengelompokkan kabupaten-kabupaten di Sulawesi Selatan berdasarkan tingkat pendidikan penduduk dengan pendekatan yang lebih fleksibel, yaitu Fuzzy C-Means (FCM). Pendekatan ini dipilih karena kemampuannya untuk menangani data yang memiliki ketidakpastian dan memberikan derajat keanggotaan pada setiap objek, yang memungkinkan setiap kabupaten untuk memiliki kemungkinan bergabung dalam lebih dari satu kluster. Penerapan metode FCM bertujuan untuk menghasilkan analisis yang lebih mendalam mengenai kesenjangan pendidikan di Sulawesi Selatan, serta memberikan gambaran yang lebih jelas mengenai karakteristik pendidikan yang dimiliki oleh setiap kabupaten. Hasil penelitian ini diharapkan dapat memberikan kontribusi bagi kebijakan pendidikan daerah dalam merumuskan strategi yang lebih efektif untuk meningkatkan kualitas pendidikan di Sulawesi Selatan.

Berikut penjelasan secara sistematis mengenai tahapan-tahapan yang dilakukan mulai dari pengumpulan dan preprocessing data, pemilihan metode Fuzzy C-Means, hingga penerapan algoritma klusterisasi untuk memperoleh kluster-kluster kabupaten berdasarkan indeks pendidikan. Selain itu, penentuan jumlah kluster yang optimal, teknik validasi hasil, serta interpretasi hasil klusterisasi juga akan dibahas secara mendalam. Semua tahapan ini diharapkan dapat memberikan pemahaman yang komprehensif tentang metode yang digunakan untuk mencapai tujuan penelitian.

2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data penduduk berdasarkan Pendidikan terakhir periode 2023/2024 Genap Sulawesi Selatan yang diambil dari website Direktorat Jenderal Anak Usia Dini, Pendidikan Menengah Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi [8] dan data Upah Minimum Kabupaten/Kota (UMK) dari website Gajimu [9]. Berdasarkan jumlah daerah di Provinsi Sulawesi Selatan dengan tiga variabel, maka dipilih 24 titik data. Tingkat pendidikan dimulai dari Taman Kanak-kanak (TK) hingga Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP) hingga Sekolah Menengah Atas (SMA), dan juga Upah Minimum Kabupaten/Kota (UMK) pada tahun 2024 adalah variabel yang dipertimbangkan dalam analisis ini.

Tabel 1 Data Penduduk berdasarkan Pendidikan terakhir

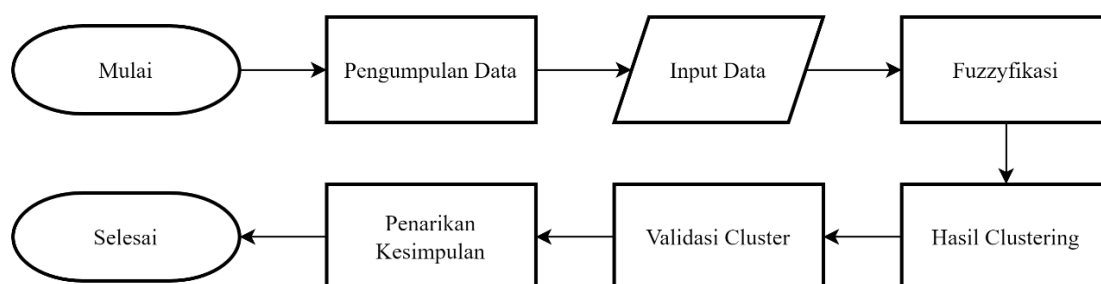
No.	Wilayah	TK - SD	SMP-SMA	UMK
1.	Kota Makassar	1185	449	Rp 3,643,321
2.	Kab. Bone	1229	189	Rp 3,434,298
3.	Kab. Gowa	918	197	Rp 3,434,298
4.	Kab. Bulukumba	830	107	Rp 3,434,298
5.	Kab. Wajo	767	108	Rp 3,434,298
6.	Kab. Pangkajene Kepulauan	616	142	Rp 3,434,298
7.	Kab. Jeneponto	558	121	Rp 3,434,298
8.	Kab. Luwu	519	157	Rp 3,434,298
9.	Kab. Sinjai	598	72	Rp 3,434,298
10.	Kab. Maros	534	131	Rp 3,434,298
11.	Kab. Pinrang	561	90	Rp 3,434,298

No.	Wilayah	TK - SD	SMP-SMA	UMK
12.	Kab. Tana Toraja	493	127	Rp 3,434,298
13.	Kab. Takalar	524	82	Rp 3,434,298
14.	Kab. Luwu Utara	471	110	Rp 3,434,298
15.	Kab. Sidenreng Rappang	483	79	Rp 3,434,298
16.	Kab. Enrekang	476	74	Rp 3,434,298
17.	Kab. Toraja Utara	403	116	Rp 3,434,298
18.	Kab. Barru	460	56	Rp 3,434,298
19.	Kab. Soppeng	426	60	Rp 3,434,298
20.	Kab. Luwu Timur	384	76	Rp 3,434,298
21.	Kab. Bantaeng	339	63	Rp 3,434,298
22.	Kab. Kepulauan Selayar	316	75	Rp 3,434,298
23.	Kota Palopo	200	56	Rp 3,434,298
24.	Kota Parepare	182	47	Rp 3,434,298
Total	Total	13472	2784	Rp 82.632.175

2.2 Proses dan Alur Penelitian

Pada penelitian ini merupakan penelitian dalam bidang statistik terapan yang bertujuan memberikan solusi praktis dalam mengelompokkan kabupaten pada provinsi Sulawesi Selatan berdasarkan indikator indeks pendidikan terakhir penduduk dan pengaruhnya terhadap upah minimum tiap Kabupaten/Kota. Penelitian ini menggunakan data dari web Direktorat Jenderal Pendidikan Anak Usia Dini, Pendidikan Dasar dan Pendidikan Menengah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi [13] dan data Upah Minimum Kabupaten/Kota (UMK) dari website Gajimu [14].

Proses penelitian dimulai dengan mengumpulkan literatur atau referensi terkait pengelompokan data menggunakan metode *Fuzzy C-Means*. Selanjutnya, data pendidikan terakhir penduduk dan data Upah minimum tiap kabupaten/kota dikumpulkan dan diklasterkan menggunakan metode *Fuzzy C-Means* [15]. Langkah berikutnya adalah validasi hasil *clustering* dengan menganalisis ketepatan objek yang telah dikelompokkan. Lalu melakukan validasi secara keseluruhan. Akhirnya, kesimpulan dari penelitian ini diambil berdasarkan hasil *clustering* yang telah dihitung secara seksama dan divalidasi. Tahapan-tahapan penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 Flowchart Penelitian

Keterangan:

1. Mulai: Penelitian dimulai dengan mengumpulkan literatur terkait pengelompokan data menggunakan metode *Fuzzy C-Means*. Data pendidikan terakhir penduduk dan data upah minimum tiap kabupaten/kota dikumpulkan untuk analisis klusterisasi.
2. Pengumpulan Data: Data penduduk berdasarkan pendidikan terakhir periode tertentu di Sulawesi Selatan dan data upah minimum Kabupaten/Kota dikumpulkan dari sumber yang telah disebutkan sebelumnya.
3. Input Data: Data yang telah terkumpul dijadikan sebagai input untuk proses analisis. Data ini mencakup informasi pendidikan terakhir penduduk dan upah minimum kabupaten/kota yang akan digunakan dalam proses klusterisasi.
4. Fuzzyfikasi: Data yang telah diinput diubah menjadi bentuk fuzzy atau nilai keanggotaan. Proses ini memungkinkan data diinterpretasikan dalam konteks klusterisasi menggunakan metode *Fuzzy C-Means*.
5. Hasil *Clustering*: Setelah proses Fuzzyfikasi, data dikelompokkan menjadi beberapa kluster berdasarkan karakteristik yang dimiliki. Hasil klusterisasi ini menunjukkan bagaimana data terorganisir dan dikelompokkan berdasarkan pola yang ada.
6. Validasi *Cluster*: Tahap validasi dilakukan untuk memastikan keakuratan hasil klusterisasi. Proses ini bertujuan memeriksa seberapa baik klusterisasi dilakukan dan apakah hasilnya dapat dipercaya.
7. Penarikan Kesimpulan: Kesimpulan diambil berdasarkan hasil klusterisasi yang telah divalidasi. Analisis lebih lanjut atau keputusan dapat dibuat berdasarkan hasil klusterisasi yang telah dilakukan.
8. Selesai: Tahapan penelitian selesai setelah kesimpulan diambil berdasarkan hasil klusterisasi yang telah divalidasi, dan analisis lebih lanjut dilakukan.

2.3 Fuzzy C-Means

FCM adalah algoritma yang digunakan untuk pengklasteran data yang terawasi, karena cluster ditentukan sesuai dengan input cluster dan hasil output yang diharapkan [11]. Tingkat keikutsertaan dalam sebuah cluster menentukan apakah sebuah titik data termasuk dalam Fuzzy C-Means (FCM), sebuah pendekatan pengelompokan data. Jim Bezdek adalah orang memperkenalkan teknik FCM pada tahun 1981. Ide dasar dari FCM adalah fase awal penentuan pusat cluster, dimana keadaan awal pusat cluster masih belum tepat dan menandai lokasi rata-rata untuk setiap cluster [12]. Pusat cluster dapat ditunjukkan untuk bermigrasi ke arah lokasi yang diinginkan dengan memperbaiki pusat cluster dari hasil yang didapatkan dan nilai keanggotaan setiap data secara iteratif.

Deretan pusat kluster dan derajat keanggotaan setiap objek dalam setiap kluster adalah hasil dari pendekatan Fuzzy C-Means.

Algoritma Fuzzy C-Means sebagai berikut:

- a. Matriks $n \times m$ yang berisi data input yang akan di-cluster (n = jumlah sampel data, m = variabel input untuk setiap set data) digunakan. X_{ij} = variabel input ke- j ($j = 1, 2, \dots, m$), data sampel ke- i ($i = 1, 2, 3, \dots, n$).
- b. Tentukan fungsi objektif awal ($P_0 = 0$), jumlah cluster (c), pangkat (w), iterasi maksimum ($\max Iter$), galat terkecil yang diharapkan (ϵ), dan iterasi awal ($t=1$).
- c. Membuat bilangan acak μ_{ik} untuk $i = 1, 2, \dots, n$ dan $k = 1, 2, \dots, c$ sebagai elemen-elemen matriks partisi pertama. U. Jumlahkan semua kolom:

$$Q_i = \sum_{k=1}^c \mu_{ik} \text{ Dengan } i=1, 2, \dots, n.$$
$$\text{Hitung: } \mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (1)$$

- a. Menghitung pusat cluster ke- k : V_k dengan $k=1, 2, \dots, c$ dan $j=1, 2, \dots, m$, yang dimana X_{ij} adalah variabel fuzzy yang digunakan dan w adalah bobot.

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (2)$$

- b. Penghitungan fungsi objektif pada iterasi ke-t, P_t :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (3)$$

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-\frac{1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{-\frac{1}{w-1}}} \quad (4)$$

- c. Melakukan pengecekan kondisi berhenti:

- Jika: $(|P_t - P_{t-1}| < \text{error terkecil})$ atau $(t > \text{MaxIter})$ maka berhenti;
- Jika tidak: $t=t+1$, ulangi langkah ke-4

3. HASIL DAN PEMBAHASAN

Hasil klasterisasi akan mengelompokkan kabupaten berdasarkan tingkat pendidikan penduduknya, di mana setiap kabupaten memiliki derajat keanggotaan yang dapat menunjukkan kedekatannya dengan klaster-klaster tertentu. Hasil ini diharapkan dapat memberikan gambaran yang lebih jelas mengenai distribusi pendidikan di Sulawesi Selatan serta identifikasi wilayah yang memerlukan perhatian khusus dalam upaya peningkatan kualitas pendidikan.

Pada bagian pembahasan, setiap klaster yang terbentuk akan dianalisis lebih lanjut untuk melihat karakteristik dan pola-pola yang ada di dalamnya. Pembahasan akan mencakup faktor-faktor yang mempengaruhi pembentukan klaster, seperti variabel-variabel pendidikan yang digunakan dalam penelitian ini, serta pengaruh faktor sosial, ekonomi, dan infrastruktur pendidikan yang mungkin turut berperan dalam membentuk pola tersebut. Selain itu, perbandingan antar klaster juga akan dilakukan untuk mengidentifikasi kesenjangan pendidikan antar daerah dan memberikan wawasan terkait kebijakan pendidikan yang tepat untuk masing-masing klaster.

3.1 Klasterisasi dengan Algoritma Fuzzy C-Means

Klasterisasi setiap kabupaten di Sulawesi Selatan menggunakan software *Microsoft Excel* dengan langkah-langkah sebagai berikut (dilakukan pengolahan data terlebih dahulu dengan menggunakan normalisasi agar data tidak terlalu besar dan agar mendapatkan nilai akurat)

- a. Menentukan matriks partisi awal, U , berdimensi $n \times p$, di mana n adalah jumlah sampel data (24), dan p adalah atribut/variabel dari setiap set data (tiga). X_{ij} = atribut ke- i ($j = 1, 2, 3, \dots, 24$), sampel data ke- j .
- b. Menentukan beberapa nilai parameter awal untuk perhitungan:
 1. Jumlah klaster = 2:

Penggunaan dua klaster adalah untuk membagi data penduduk berdasarkan pendidikan terakhir dan juga UMK yang berbeda secara jelas. Dengan dibaginya data tersebut, akan memudahkan identifikasi perbedaan secara karakteristik antara kelompok Pendidikan yang berbeda pada wilayah tersebut.

2. Pangkat pembobot = 2:

Penggunaan pangkat pembobot sebanyak 2 menunjukkan bahwa setiap data memiliki Tingkat keanggotaan yang jelas terhadap kluster tertentu. Hal ini dapat memungkinkan kita untuk mendapatkan hasil klasterisasi yang lebih akurat dan terdefinisi dengan baik.

3. Maksimum iterasi = 100:

Penetapan maksimum iterasi yang ditetapkan sebagai 100 kali dilakukan untuk memastikan bahwa algoritma *Fuzzy C-Means* memiliki waktu untuk konvergen dan mencapai hasil dengan solusi yang optimal dalam proses klasterisasi. Dengan pembatasan iterasi, diharapkan hasil dari klasterisasi mendapatkan nilai yang stabil dan akurat.

4. Error terkecil yang diharapkan = 0,1:

Dilakukan penetapan pada 0,1 bertujuan untuk menetapkan batas toleransi dalam melakukan proses klasterisasi. Dengan mengatur error terkecil, dapat dipastikan bahwa algoritma akan berhenti saat mencapai tingkat error terkecil yang telah ditetapkan.

5. Fungsi objektif awal = 0:

Fungsi objektif awal ditetapkan sebagai 0 agar proses klasterisasi memiliki titik awal yang netral. Dengan demikian, algoritma dapat mengoptimalkan proses klasterisasi data dari Pendidikan terakhir penduduk Sulawesi Selatan tanpa bias dari kondisi awal tertentu.

6. Iterasi awal = 1

Penetapan iterasi awal pada angka 1 dilakukan untuk memaulai proses klasterisasi dari langkah pertama. Dengan seperti itu, algoritma dapat berjalan secara sistematis dan dapat mencapai konvergensi yang baik dalam proses klasterisasi data.

c. Pilih sebuah bilangan acak (μ_{ik} , $i = 1,2,3,...,24$, $k = 1,2,3$) untuk menjadi nilai awal dari matriks partisi U. Matriks partisi U pada awalnya dikonstruksi (secara acak) dengan x.

d. Menghitung pusat *cluster* (V):

Dua pusat cluster, V_{ij} , ditemukan dengan menggunakan persamaan (2.1), di mana $i = 1,2,3,...,24$ dan $j = 1,2,3$. Setelah itu, matriks V akan dibuat, dengan setiap baris berfungsi sebagai pusat cluster. Berikut ini adalah hasil dari perhitungan pusat *cluster* pada iterasi pertama:

$$V = [0,080,16-0,170,080,16-0,170,93-0,24-1,17-1,290,060,98-0,49-0,291,81-0,04-0,28-0,02-0,04-0,28-0,02-0,49-0,291,81-0,08-0,251,63-0,260,180,18]$$

e. Dilakukan penghitungan fungsi objektif (P_t):

Nilai fungsi objektif pada iterasi pertama dan seterusnya dapat dilakukan menggunakan rumus berikut:

$$P1 = \sum_{i=1}^{24} \sum_{k=1}^3 \left(\left(\sum_{j=1}^2 (X_{ij} - V_{kj})^2 \right) (\mu_{ik})^2 \right)$$

f. Menghitung Perubahan Matriks Partisi (U):

Dengan menggunakan persamaan (2.3), perubahan matriks partisi (U) dapat dihitung.

g. Dilakukan pemeriksaan kondisi berhenti atau konvergen:

Output yang dihasilkan dengan menjalankan fungsi *ppchust()* pada perangkat lunak *Visual Studio Code* menggunakan *Python* dan melakukan komputasi manual di *Excel* menunjukkan bahwa dibutuhkan 17 iterasi untuk mencapai kondisi konvergen.

3.2 Hasil Klasterisasi

Hasil klasterisasi data indikator indeks pendidikan terakhir penduduk di 24 kabupaten di Sulawesi Selatan menggunakan *Fuzzy C-Means* adalah sebagai berikut:

Klaster 1: Makassar, Gowa, Bone.

Klaster 2: Bulukumba, Wajo, Pangkajene Kepulauan, Jeneponto, Luwu, Sinjai, Maros, Pinrang, Tana Toraja, Takalar, Luwu Utara, Sidenreng Rappang, Enrekang, Toraja Utara, Barru, Soppeng, Luwu Timur, Bantaeng, Kepulauan Selayar, Palopo, Parepare.

Berikut adalah hasil *cluster* dan juga *error threshold* setelah dilakukan perhitungan manual melalui *Excel*:

Pada tabel berikut, memperlihatkan hasil dari perhitungan dari 3 variabel yang dibuat menjadi 2 *cluster*. Ketiga variable tersebut meliputi antara Tamatan TK-SD, Tamatan SMP-SMA, dan UMK dari 24 kabupaten di Sulawesi Selatan. Dan hasil dari tabel diatas memperlihatkan hasil *cluster* yang diambil dari nilai random yang telah ditentukan sebelumnya. Perhitungan ini menggunakan metode manual dengan menggunakan *software Microsoft Excel*.

Tabel 2 Hasil Clustering 3 Variabel

	u_{i1}	u_{i2}	Cluster
1	0.17	0.83	2
2	0.08	0.92	2
3	0.15	0.85	2
4	0.63	0.37	1
5	0.74	0.26	1
6	0.89	0.11	1
7	0.97	0.03	1
8	0.92	0.08	1
9	0.96	0.04	1
10	0.96	0.04	1
11	0.98	0.02	1
12	0.98	0.02	1
13	0.99	0.01	1
14	1	0	1
15	1	0	1
16	1	0	1
17	0.99	0.01	1

	u_{i1}	u_{i2}	Cluster
18	0.99	0.01	1
19	0.99	0.01	1
20	0.99	0.01	1
21	0.97	0.03	1
22	0.97	0.03	1
23	0.93	0.07	1
24	0.92	0.08	1

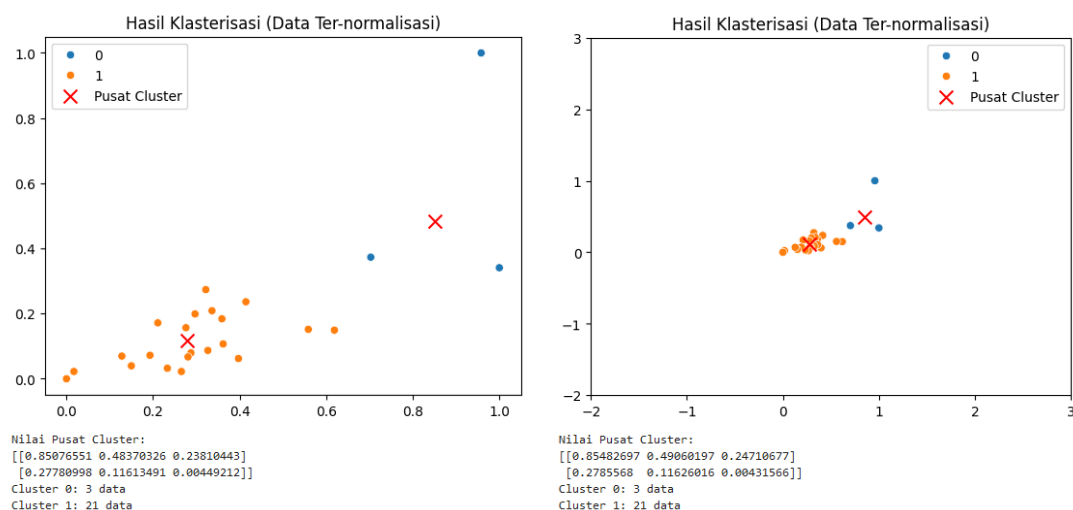
Pada tabel berikut, menunjukkan bahwa selisih dari hasil yang didapatkan dari setiap iterasi yang dilakukan. Data dari variabel awalnya memiliki nilai *default*, tetapi demi mendapatkan nilai yang tidak terlalu besar dan mendapatkan nilai yang spesifik dan detail, maka dilakukan normalisasi pada setiap variabel agar range datanya hanya mencakup $0 \sim 1$. Maka dari itu bisa dilihat pada tabel diatas, hasil dari iterasi pertama yaitu 1.04 dan dikurangi dengan nilai iterasi kedua yaitu 0.88 dan seterusnya hingga mencapai iterasi ke 17 dengan mendapatkan nilai selisih 2.13×10^{-6} . Nilai dari error tiap iterasi juga di dapatkan melalui perhitungan manual menggunakan *software Microsoft Excel*.

Tabel 3 Error Tiap Iterasi

	FO	ERROR
1	1.048625484	-
2	0.883926457	0.164699027
3	0.824543826	0.059382632
4	0.729767547	0.094776279
5	0.657720621	0.072046927
6	0.622475365	0.035245255
7	0.604008365	0.018467
8	0.595298861	0.008709504
9	0.591340759	0.003958101
10	0.589512386	0.001828374
11	0.588831083	0.000681302

	FO	ERROR
12	0.588607701	0.000223382
13	0.588527684	8.00173E-05
14	0.588460963	6.6721E-05
15	0.588437357	2.36057E-05
16	0.588420938	1.64192E-05
17	0.588418803	2.13485E-06

Berikut adalah grafik *scatterpoint*-nya:



Gambar 2 Grafik *Scatterpoint* (Rentang Sumbu 0-1) **Gambar 3** Grafik *Scatterpoint* (Rentang Sumbu 0-3)

Grafik tersebut merupakan grafik *scatterpoint* dari hasil klusterisasi melalui program Python. Perbedaan dari kedua grafik tersebut adalah di pengaturan rentang sumbu x dan y-nya. Pada grafik di sebelah kiri rentangnya adalah 0 sampai 1, ini berarti data yang diplot akan dibatasi dalam rentang 0 hingga 1 pada kedua sumbu, pada grafik di sebelah kanan rentangnya adalah 0 sampai 3, ini berarti data yang diplot akan dibatasi dalam rentang 0 hingga 3 pada kedua sumbu. Rentang sumbu dibesarkan karena dengan rentang yang besar, detail pada data yang memiliki perubahan kecil yang mungkin tidak begitu jelas dapat terlihat. Pada program yang digunakan diinisialisasikan objek FCM seperti berikut:

$$\text{FCM} = \text{FCM}(\text{n_clusters}=2, \text{max_iter}=100, \text{error}=0.1, \text{m}=2)$$

Baris kode tersebut menginisialisasi objek FCM dengan pengaturan yang ditentukan, yang kemudian dapat digunakan untuk melakukan *clustering* pada data yang diberikan menggunakan algoritma *Fuzzy C-Means*. Algoritma ini akan bekerja dengan parameter-parameter ini untuk mengelompokkan 3 variabel data ke dalam 2 *cluster* dengan iterasi maksimal 100 kali dan toleransi kesalahan 0.1 untuk mencapai konvergensi, serta menggunakan derajat *fuzziness* 2 untuk menentukan keanggotaan data dalam *cluster*. Kesimpulan yang diperoleh dari kedua proses klusterisasi yang telah dilakukan adalah data terklusterisasi menjadi 2 kluster dengan

klaster 1 menampung sebanyak 21 data dan klaster 2 menampung sebanyak 3 data walaupun terdapat perbedaan pada posisi pusat klaster diantara perhitungan menggunakan *Excel* dan penggunaan program *Python*.

4. KESIMPULAN DAN SARAN

Berdasarkan analisis yang telah dilakukan, penelitian ini berhasil mengelompokkan kabupaten-kabupaten di Sulawesi Selatan ke dalam beberapa klaster yang memiliki karakteristik pendidikan yang serupa. Setiap klaster mencerminkan tingkat kemajuan pendidikan yang berbeda-beda antar kabupaten, yang diukur melalui variabel-variabel seperti angka melek huruf, rata-rata lama sekolah, dan persentase penduduk yang mengenyam pendidikan menengah dan tinggi. Hasil klasterisasi menunjukkan adanya dua klaster yang memungkinkan untuk mengidentifikasi Kabupaten-kabupaten dengan tingkat pendidikan terakhir yang serupa. Dalam penelitian ini, dengan 17 iterasi, algoritma *Fuzzy C-Means* mencapai solusi optimal dalam proses klasterisasi. Proses klasterisasi dilakukan menggunakan perangkat lunak *Excel* dan *Python* yang menghasilkan dua klaster. Klaster 1 menampung sebanyak 21 data dan klaster 2 menampung sebanyak 3 data, meskipun terdapat perbedaan pada posisi pusat klaster di antara masing-masing perangkat lunak yang digunakan. Untuk memvisualisasikan hasil klasterisasi, grafik *scatterpoint* juga dihasilkan oleh program *Python*. Dari hasil klasterisasi tersebut, ditemukan bahwa ada beberapa kabupaten yang tergolong dalam klaster dengan tingkat pendidikan yang lebih tinggi, sementara sebagian lainnya masih tergolong dalam klaster dengan tingkat pendidikan yang lebih rendah. Temuan ini menunjukkan adanya kesenjangan yang signifikan dalam distribusi pendidikan di Sulawesi Selatan, yang bisa dipengaruhi oleh berbagai faktor seperti kondisi ekonomi, infrastruktur pendidikan, serta kebijakan daerah yang diterapkan.

Sebagai implikasi praktis, hasil penelitian ini memberikan wawasan yang penting bagi pemerintah daerah dan pemangku kebijakan untuk merancang program pendidikan yang lebih terfokus dan tepat sasaran. Kabupaten-kabupaten yang tergolong dalam klaster dengan pendidikan rendah dapat menjadi prioritas dalam program intervensi untuk meningkatkan akses pendidikan, kualitas pengajaran, serta fasilitas pendidikan yang memadai. Sementara itu, kabupaten dengan tingkat pendidikan lebih tinggi dapat menjadi model atau pusat pengembangan pendidikan yang lebih maju di daerah tersebut.

Adapun beberapa rekomendasi yang dapat diambil dari hasil penelitian ini, antara lain:

1. Peningkatan Infrastruktur Pendidikan: Kabupaten dengan tingkat pendidikan rendah perlu mendapatkan perhatian khusus dalam hal penyediaan fasilitas pendidikan yang memadai, seperti sekolah, tenaga pengajar, dan sarana pendukung lainnya.
2. Program Pelatihan dan Pengembangan: Perlu adanya program pelatihan bagi tenaga pendidik untuk meningkatkan kualitas pengajaran, terutama di kabupaten-kabupaten dengan hasil pendidikan yang lebih rendah.
3. Kebijakan Pemerataan Pendidikan: Pemerintah daerah perlu merancang kebijakan yang lebih inklusif dan merata dalam hal distribusi anggaran pendidikan agar daerah-daerah yang tertinggal dapat memperoleh perhatian lebih.

Secara keseluruhan, penelitian ini menyarankan agar kebijakan pendidikan di Sulawesi Selatan dapat lebih memperhatikan perbedaan karakteristik pendidikan antar kabupaten dan menyesuaikan program pembangunan pendidikan dengan kondisi dan kebutuhan spesifik masing-masing daerah. Hal ini diharapkan dapat membawa dampak positif bagi pemerataan dan peningkatan kualitas pendidikan di seluruh wilayah Sulawesi Selatan.

Penelitian ini juga diharapkan dapat menjadi referensi bagi penelitian selanjutnya yang membahas klasterisasi pendidikan di daerah lain atau bahkan dalam skala nasional, dengan menggunakan metode yang lebih kompleks atau melibatkan lebih banyak variabel untuk analisis yang lebih komprehensif.

5. UCAPAN TERIMA KASIH



Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penelitian ini dapat diselesaikan dengan baik. Pada kesempatan ini, saya ingin mengucapkan terima kasih yang sebesar-besarnya kepada Badan Pusat Statistik Provinsi Sulawesi Selatan atas data dan informasi yang telah diberikan untuk kelancaran penelitian ini. Serta ucapan terimakasih kepada semua pihak yang telah memberikan bantuan, baik secara langsung maupun tidak langsung, yang tidak bisa saya sebutkan satu per satu. Semoga kebaikan dan dukungannya selalu mendapatkan balasan yang setimpal. Semoga hasil penelitian ini dapat memberikan manfaat dan kontribusi yang positif, terutama dalam bidang deteksi kemiripan judul penelitian. Saya menyadari bahwa penelitian ini masih jauh dari sempurna, untuk itu, saya sangat mengharapkan kritik dan saran yang membangun demi perbaikan di masa yang akan datang.

REFERENSI

- [1] Y. Paskissing, M. Nasir, and S. Nujum, 2020, Analisis Pertumbuhan Dan Ketimpangan Ekonomi Kabupaten/Kota Di Provinsi Sulawesi Selatan, *Journal of Management Science (JMS)*, vol. 1, no. 1, hal 35-87
- [2] W. Sanusi et al., 2019, Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-faktor Penyebab Gizi Buruk, <http://www.ojs.unm.ac.id/jmathcos>
- [3] Abdullah, 2023, Badan Pusat Statistik Sulawesi Selatan, <https://sulsel.bps.go.id/>, Accessed: Apr. 11, 2024.
- [4] J. Juardi, Muh. A. A. Ahmad, and B. Iwang, 2023, Analisis Jumlah Penduduk, Tingkat Pendidikan, PDRB, dan Inflasi terhadap Kemiskinan di Provinsi Sulawesi Selatan, *Bulletin of Economic Studies (BEST)*, vol. 3, no. 1, hal. 14–24
- [5] N. Dwitianti, N. Selvia, and F. R. Andrari, 2019, Penerapan Fuzzy C-Means Cluster dalam Pengelompokkan Provinsi Indonesia Menurut Indikator Kesejahteraan Rakyat, *Faktor Exacta*, vol. 12, no. 3, hal. 201
- [6] N. Ulinuha, U. Sunan, and A. Surabaya, “Provincial Clustering in Indonesia Based on Plantation Production Using Fuzzy C-Means.”
- [7] Hanniva, A. Kurnia, S. Rahardianto, and A. A. Mattjik, 2022, Penggerombolan Kabupaten/Kota di Indonesia Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan Metode K-Means dan Fuzzy C-Means, *Xplore: Journal of Statistics*, vol. 11, no. 1, hal. 36–47
- [8] Y. Nataliani, 2020, Klasterisasi kinerja karyawan menggunakan algoritma fuzzy c-means, *AITI: Jurnal Teknologi Informasi*, vol. 17, no. Agustus, hal. 118–129
- [9] D. Lolita Pardosi and I. Damayanti Siagian, Klasterisasi Data Lowongan Pekerjaan Berdasarkan Fuzzy C-Means, *JIKOMSI Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 3, no. 2, hal. 27–31.
- [10] R. Goejantoro and F. Deny Tisna Amijaya, 2019, Comparison of C-Means and Fuzzy C-Means Methods in the Districts/Cities on the Island of Kalimantan Based on the 2019 HDI Indicators, *Jurnal EKSPONENSIAL*, vol. 12, no. 2. Hal 26
- [11] A. Surya Maulana, A. Nazir, L. Handayani, and I. Afrianty, 2023, KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Algoritma Fuzzy C-Means untuk Melihat Pola Penerima Beasiswa Bank Indonesia, *Media Online*, vol. 3, no. 6, hal. 670–679
- [12] 2024, “Data Pokok Pendidikan Direktorat Jenderal Pendidikan Anak Usia Dini, Pendidikan Dasar dan Pendidikan Menengah Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi,” <https://dapo.kemdikbud.go.id/sp/1/190000>, Accessed: May 11 2024
- [13] 2024, “UMP/UMK Sulawesi Selatan,” <https://gajimu.com/garmen/gaji-pekerja-garmen/gaji-minimum/ump-umk-sulsel>, Accessed: May 11, 2024. [Online]. Available:



- [14] N. I. Kalla, S. Annas, and M. Fahmuddin, 2022, Metode Subtractive Fuzzy C-Means (SFCM) dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kemiskinan, *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, vol. 4, no. 2, hal. 95–108