

Pengelompokan Faktor-Faktor yang Mempengaruhi Curah Hujan di Provinsi Sumatera Utara Menggunakan Metode *Fuzzy C-Means*

Bravina Aulia Damiri¹, Wulan Nur Ramadhan², dan Kamalina Rosyida Supriyanti³

^{1,2,3}Statistika, Universitas Muhammadiyah Semarang
E-mail: bravinaaaliadamiri@gmail.com¹, wulannr098@gmail.com²,
kamalinarosyida2609@gmail.com³

Diajukan 9 Januari 2024 *Diperbaiki* 21 Juni 2024 *Diterima* 28 Juni 2024

Abstrak

Latar Belakang: Di Indonesia, kondisi curah hujan dipengaruhi oleh berbagai faktor, termasuk kondisi fisiografis, pola angin, dan perubahan iklim. Kondisi curah hujan di Sumatera Utara pada tahun 2020 bervariasi di berbagai wilayah. Pada tahun 2020 curah hujan di Sumatera Utara berada pada angka 4.380 mm di mana angka tersebut termasuk angka tertinggi dari curah hujan normal yang berada dalam kisaran 1000-3000 mm per tahun.

Tujuan: Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi tingkat curah hujan di Provinsi Sumatera Utara.

Metode: Metode analisis data yang digunakan adalah *Fuzzy C-Means*. Penelitian ini menggunakan data sekunder yang bersumber dari BPS Sumatera Utara. Data yang diambil berupa data curah hujan, kelembapan udara, kecepatan angin, penyinaran matahari, dan penguapan di beberapa stasiun di Provinsi Sumatera Utara pada tahun 2020.

Hasil: Hasil analisis yang didapat yaitu menetapkan dua *cluster* pada penelitian ini, stasiun BMKG dapat dikelompokkan menjadi *cluster* 1 (Deli Serdang, BMKG Wilayah I, Marihat, Pinangsori, dan Gunung Sitoli) dan *cluster* 2 (Belawan dan Tanjung Morawa). Dengan begitu, metode *Fuzzy C-Means Clustering* memberikan hasil yang baik dalam mengelompokkan faktor-faktor curah hujan.

Kesimpulan: Penerapan metode *Fuzzy C-Means* memberikan hasil yang baik dalam mengelompokkan faktor-faktor curah hujan.

Kata kunci: Curah Hujan, Sumatera Utara, Pengelompokkan, *Fuzzy C-means*.

Abstract

Background: In Indonesia, rainfall conditions are influenced by various factors, including physiographic conditions, wind patterns and climate change. Rainfall conditions in North Sumatra in 2020 varied in various regions. In 2020, rainfall in North Sumatra was 4,380 mm, which is the highest figure of normal rainfall which is in the range of 1000-3000 mm per year.

Objective: This research aims to analyse the factors that influence rainfall levels in North Sumatra Province.

Methods: The data analysis method used is *Fuzzy C-means*. This research uses secondary data sourced from BPS North Sumatra. The data taken is rainfall, air humidity, wind speed, solar radiation and evaporation at several stations in North Sumatra Province in 2020.

Results: The results of the analysis obtained were to determine two clusters in this research, BMKG stations can be grouped into cluster 1 (Deli Serdang, BMKG Region I, Marihat, Pinangsori, and Gunung Sitoli) and cluster 2 (Belawan and Tanjung Morawa). In this way, the *Fuzzy C Means Clustering* method provides good results in grouping rainfall factors.

Conclusion: The application of the *Fuzzy C-Means* method provides good results in grouping rainfall factors.

Keywords : Rainfall, North Sumatra, Grouping, *Fuzzy C-means*.

PENDAHULUAN

Kondisi curah hujan pada negara Indonesia dipengaruhi oleh beberapa faktor yaitu kondisi fisiografis, pola angin, dan perubahan iklim. Berdasarkan karakteristik geografis Indonesia, terdapat tiga tipe curah hujan utama, yaitu tipe ekuatorial, monsun, dan lokal. Tipe ekuatorial dicirikan oleh dua kali maksimum curah hujan bulanan dalam setahun dan terjadi di wilayah Sumatera dan Kalimantan. Sementara tipe monsun dipengaruhi oleh angin laut dalam skala yang sangat luas, dengan adanya perbedaan yang jelas antara periode musim hujan dan kemarau, dan hanya terjadi satu kali maksimum curah hujan bulanan dalam setahun. Selain itu, terdapat juga tipe curah hujan lokal yang dipengaruhi oleh kondisi lokal seperti bentang perairan atau lautan, pegunungan yang tinggi, serta pemanasan lokal yang intensif Indonesia memiliki iklim tropis karena dilewati oleh garis khatulistiwa, yang menyebabkan sebagian besar wilayahnya memiliki iklim tropis. Suhu rata-rata tahunan di Indonesia telah mengalami peningkatan sekitar 0,3°C sejak tahun 1988 (Tukidi, 2010).

Tren peningkatan frekuensi dan intensitas curah hujan ekstrem juga telah terjadi di seluruh Indonesia dalam beberapa dekade terakhir. Hal ini disebabkan oleh perubahan iklim global, seperti fenomena La Nina yang mempengaruhi curah hujan di wilayah Indonesia. Peningkatan curah hujan ekstrem ini menuntut adaptasi masyarakat dan berbagai pihak terkait untuk mengelola dampaknya. Beberapa kejadian hujan ekstrem di Indonesia juga berhubungan dengan siklon tropis, yang dapat meningkatkan curah hujan hingga 340 mm/hari (Nugroho, 2022).

Sumatera Utara adalah sebuah provinsi di Indonesia yang terletak di Pulau Sumatera dengan koordinat 1° - 4°

Lintang Utara dan 98° - 100° Bujur Timur. Provinsi ini memiliki ibu kota di kota Medan dan terdiri dari 33 kabupaten/kota. Letaknya yang strategis di sekitar garis khatulistiwa, dikelilingi Pegunungan Bukit Barisan dan berbatasan dengan Selat Malaka dan Samudera Hindia, menyebabkan iklim dan curah hujan di provinsi ini dipengaruhi faktor-faktor iklim global seperti *Indian Ocean Dipole* (IOD), *Intertropical Convergence Zone* (ITCZ), *Madden Julian Oscillation* (MJO), dan *El Nino Southern Oscillation* (ENSO). Selain itu, faktor iklim regional seperti monsun, gangguan tropis, dan daerah konvergensi juga berpengaruh terhadap kondisi iklim Sumatera Utara. Di skala lokal, pertumbuhan awan dan curah hujan juga dipengaruhi kondisi geografis daerah serta pergerakan semu matahari (Prasetyo et al., 2018).

Kondisi curah hujan di Sumatera Utara pada tahun 2020 bervariasi di berbagai wilayah. Sebagian wilayah memiliki curah hujan kategori tinggi, sementara yang lain memiliki curah hujan kategori menengah (Chaniago, 2023). Pada tahun 2020 curah hujan di Sumatera Utara berada pada angka 4.380 mm dimana angka tersebut dimaksudkan angka tertinggi dari curah hujan normal yang berada dalam kisaran 1000-3000 mm per tahun (Sumut, 2020). Curah hujan ini juga dipengaruhi oleh iklim yang mempengaruhi fluktuasi curah hujan di Provinsi Sumatera Utara. Pola iklim, termasuk pola angin monsun, dapat menentukan musim hujan dan kemarau di suatu wilayah. Di Indonesia, pola angin musim barat dan timur memengaruhi musim hujan dan kemarau. Selain itu, perubahan iklim global juga berdampak pada pola curah hujan di berbagai wilayah, termasuk Sumatera Utara. Perubahan iklim dapat menyebabkan fluktuasi curah hujan dan meningkatkan frekuensi kejadian cuaca ekstrem, seperti hujan lebat dan banjir (BPS, 2021).

Mempertimbangkan kerugian besar yang ditimbulkan akibat ketidakmampuan mengantisipasi curah hujan tinggi yang berpotensi memicu bencana di suatu daerah,

maka penting dilakukan upaya untuk mempelajari pola dan karakteristik curah hujan. Selain itu, pengelompokan karakteristik curah hujan ke dalam klaster yang sesuai juga diperlukan agar langkah-langkah pencegahan dan mitigasi terhadap risiko bencana akibat curah hujan dapat dilakukan dengan lebih efektif. Salah satu metode statistika yang banyak digunakan dalam pengelompokan yaitu metode *clustering*.

Clustering merupakan metode pengelompokan data berdasarkan ukuran kedekatan atau kesamaan karakteristik. Berbeda dengan pengelompokan biasa yang membentuk *group*, klaster tidak mensyaratkan anggotanya memiliki kondisi yang identik. Pengelompokan dilakukan berdasarkan kemiripan karakteristik data, salah satunya dengan menghitung jarak Euclidean. *Clustering* sangat bermanfaat karena dalam menganalisis suatu permasalahan atau pengambilan keputusan, data jarang bersifat persis sama, namun cenderung memiliki kemiripan tertentu (Sanusi et al., 2020).

Penelitian sebelumnya dilakukan oleh Tanjung et al. (2021) yaitu mengelompokkan curah hujan di Kalimantan Timur menggunakan metode *Fuzzy C-Means*. Hasil penelitian yang diperoleh menunjukkan bahwa metode *Fuzzy C-Means clustering* dengan tiga *cluster* merupakan jumlah yang ideal dibandingkan dengan hasil *clustering* berjumlah dua *cluster*. Penelitian ini menghasilkan informasi tentang kategori tingkat curah hujan di Kalimantan Timur, dimana daerah Long Iram memiliki tingkat curah hujan yang tinggi dan Tenggara, Tenggara Seberang, dan Temindung memiliki kategori curah hujan terendah.

Penelitian Curah Hujan juga pernah dilakukan oleh Setiawan (2021) yaitu Analisis Curah Hujan di Indonesia untuk Memetakan Daerah Potensi Banjir dan Tanah Longsor dengan Metode *Cluster*

Fuzzy C-Means dan *Singular Value Decomposition* (SVD). Hasil yang diperoleh Menggunakan *Fuzzy C-means* dapat dibentuk tiga *cluster*, dengan kriteria daerah curah hujan tinggi, sedang, dan rendah. Sedangkan Analisis SVD menghasilkan empat pola curah hujan dominan di Indonesia dengan variansi sebesar 25,59%, dimana Papua dan Papua Barat merupakan wilayah dengan curah hujan tertinggi. Berdasarkan hasil analisis, diketahui bahwa terdapat daerah-daerah yang berpotensi mengalami bencana banjir dan longsor akibat curah hujan yang tinggi. Oleh karena itu, upaya pencegahan (preventif) dan penanggulangan (mitigasi) bencana banjir dan longsor perlu dilakukan dengan lebih baik di daerah-daerah tersebut.

Penelitian selanjutnya dilakukan oleh Hadi et al. (2022) dengan judul Penerapan *Fuzzy C-Means* dan *Fuzzy Sugeno* dalam memprediksi cuaca. Hasil yang diperoleh tingkat akurasi dari model prediksi cuaca pada penelitian ini yaitu menghasilkan akurasi yang baik yaitu sebesar 71,10 % dengan jumlah data prediksi benar sebesar 251 dari total 353 data cuaca.

Metode *Fuzzy C-Means* merupakan salah satu teknik pengelompokan data yang menggunakan logika *fuzzy* untuk menentukan derajat keanggotaan setiap data dalam suatu *cluster*. Pendekatan ini memungkinkan data untuk termasuk ke dalam lebih dari satu *cluster* dengan tingkat keanggotaan yang berbeda-beda. Dalam konteks Stasiun BMKG Sumatera Utara, *Fuzzy C-Means* merupakan salah satu metode *cluster* data yang mengelompokkan data berdasarkan ciri dan karakteristiknya. Kelebihan dari *Fuzzy C-Means* adalah memiliki perhitungan dan algoritma yang sederhana serta tidak membutuhkan komputasi yang rumit, sehingga mudah untuk diterapkan (Satya Nugraha & Amelia Riyandari, 2020). Penerapan metode ini dapat membantu dalam mengelompokkan data iklim, seperti curah hujan, suhu, kelembaban, dan variabel iklim lainnya, sehingga memungkinkan untuk mendapatkan pemahaman yang lebih baik mengenai pola iklim di wilayah tersebut.

Pengelompokan Faktor-Faktor yang Mempengaruhi Curah Hujan di....

Metode ini telah banyak digunakan dalam berbagai studi terkait pengelompokan data iklim, termasuk dalam memprediksi cuaca dan mengkaji pola gempa bumi (Latipa Sari & Suranti, 2016).

METODE

Desain Penelitian

Penelitian ini mengimplementasikan pendekatan kuantitatif. Metode analisis data yang digunakan adalah *Fuzzy C-means* dengan bantuan perangkat lunak R-Studio untuk mengelompokkan faktor-faktor yang berpengaruh terhadap curah hujan di Provinsi Sumatera Utara. Pendekatan kuantitatif dipilih untuk menganalisis data secara numerik. Metode ini cocok digunakan untuk pengelompokan faktor-faktor curah hujan karena mampu menangani ketidakpastian dan ketidakjelasan dalam data.

Populasi dan Sampel

Populasi penelitian ini mencakup beberapa stasiun di Sumatera Utara, sementara sampel penelitian terdiri dari variabel-variabel seperti kelembapan udara, curah hujan, penyinaran matahari, kecepatan angin, dan penguapan. Data yang diambil yaitu data Tahun 2020 yang bersumber dari Badan Pusat Statistik Provinsi Sumatera Utara (Sumut, 2020).

Teknik Sampiling

Metode pengambilan sampel dalam penelitian ini yaitu menggunakan teknik *Simple Random Sampling*. Teknik *Simple Random Sampling* merupakan metode pengambilan sampel dari populasi secara acak tanpa memperhatikan strata di dalam populasi, di mana setiap anggota populasi memiliki kesempatan yang sama untuk terpilih menjadi sampel (Hidayat, 2012). Dalam konteks

penerapan studi kasus penelitian ini, sampel data historis digunakan untuk menganalisis faktor-faktor yang mempengaruhi curah hujan di beberapa stasiun Badan Meteorologi, Klimatologi dan Geofisik (BMKG) Provinsi Sumatera Utara.

Subjek Penelitian

Penelitian ini menggunakan data sekunder yang bersumber dari Badan Pusat Statistik Sumatera Utara (BPS SUMUT). Data yang diambil berupa data curah hujan, kelembapan udara, kecepatan angin, penyinaran matahari, dan penguapan di 7 stasiun di Provinsi Sumatera Utara pada tahun 2020. Data ini berskala rasio. Pemilihan variabel tersebut disajikan pada Tabel 1 berikut :

Tabel 1. Iklim di Stasiun Provinsi Sumatera Utara

Stasiun	Variabel				
	Curah Hujan (Y)	Kelembapan Udara (X1)	Penyinaran Matahari (X2)	Kecepatan Angin (X3)	Penguapan (X4)
Deli Serdang	85	2 741	55	1.15	3.8
BMKG Wilayah I	85	3 727	49	1.08	3
Belawan	84	2 173	67	1.94	2.7
Tanjung Morawa/KNIA	87	1 722	65	2.4	4.1
Marihat	84	2 413	55	0	3.1
Pinangsori	85	4 380	56	0.78	5.1
Gunungsitoli	89	3 179	59	1.23	4.2

Teknik Analisis Data

Data mining merupakan suatu proses analisis yang bertujuan untuk menemukan pola menarik dan pengetahuan yang tersembunyi dalam kumpulan data yang besar. Sumber data yang digunakan dapat melibatkan berbagai *platform* seperti *database*, gudang data, web, repositori informasi, atau data yang mengalir ke dalam sistem dinamis (Rouza & Fimawahib, 2020). Analisis data yang digunakan dalam penelitian ini adalah *Fuzzy C-Means*. Analisis *Fuzzy C-Means* (FCM) adalah teknik *clustering* (pengelompokan) yang menggunakan pendekatan *fuzzy* untuk menetapkan tingkat keanggotaan setiap data terhadap kelompok tertentu. Konsep

dari *Fuzzy C-Means* pertama kali adalah menentukan pusat *cluster*, yang akan menandai lokasi rata-rata untuk tiap-tiap *cluster*. Pada tahap awal *cluster Fuzzy C-Means*, pusat *cluster* yang terbentuk masih belum akurat. Setiap data memiliki derajat keanggotaan untuk setiap *cluster*. Dengan melakukan perbaikan posisi pusat *cluster* dan nilai derajat keanggotaan setiap data secara berulang, maka pusat *cluster* akan bergerak menuju lokasi yang lebih tepat (Rahakbauw et al., 2019). Output yang dihasilkan dari metode *Fuzzy C-Means* berupa pusat *cluster* dan derajat keanggotaan setiap data pada tiap *cluster*. Informasi ini dapat dimanfaatkan untuk membangun sebuah sistem inferensi *fuzzy*. Pusat *cluster* menunjukkan titik pusat dari masing-masing kelompok data yang terbentuk. Sementara derajat keanggotaan menunjukkan seberapa besar tingkat kedekatan atau kemiripan setiap data terhadap pusat *cluster* (Novianti et al., 2022). Dalam mengimplementasikan metode *Fuzzy C-Means*, langkah pertama adalah menentukan jumlah *cluster* yang diinginkan. Setelah itu, setiap data diberikan nilai keanggotaan yang berbeda untuk setiap *cluster* (Turrahma et al., 2023). Berikut adalah langkah-langkah umum dalam analisis *Fuzzy C-Means*:

1. Inputkan data yang akan di- cluster, yaitu X , berupa matrix berukuran $n \times m$ (n = jumlah sampel data, m = atribut setiap data).
2. Periksa Missing Value
3. Tentukan jumlah *cluster* (c), derajat bobot (w), iterasi maksimum ($MaxIter$), *error minimum* (ϵ), nilai awal fungsi objektif ($P_0 = 0$), dan iterasi awal ($t = 1$).
4. Bangkitkan bilangan acak (μ_{ik}), dengan $i=1,2,\dots,n$; $k=1,2,\dots,c$; dan c sebagai elemen elemen matriks partisi awal.

5. Hitung pusat *cluster* ke- k (V_{kj}) dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,m$ pada Persamaan 1 di bawah ini :

$$V_{kj} = \frac{\sum_{i=1}^n ((U_{ik})^w + X_{ik})}{\sum_{i=1}^n (U_{ik})^w} \quad (1)$$

Dengan :

V_{kj} = Pusat *cluster* ke- k dan atribut ke- j ,

U_{ik} = Derajat keanggotaan sampel data ke- i dan *cluster* ke- k , Konsisten menggunakan *cluster* atau klaster

X_{ik} = Data pada sampel ke- i dan atribut ke- j .

6. Hitung fungsi objektif pada iterasi ke- t , untuk melihat jarak data ke- i terhadap pusat *cluster* dengan Persamaan 2 berikut :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (X_{ik} - V_{kj})^2] (U_{ik})) \quad (2)$$

7. Hitung perubahan matriks dengan $i=1,2,\dots,n$; dan $k=1,2,\dots,c$ pada Persamaan 3 berikut :

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ik} - V_{kj})^2]^{\frac{-1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ik} - V_{kj})^2]^{\frac{-1}{w-1}}} \quad (3)$$

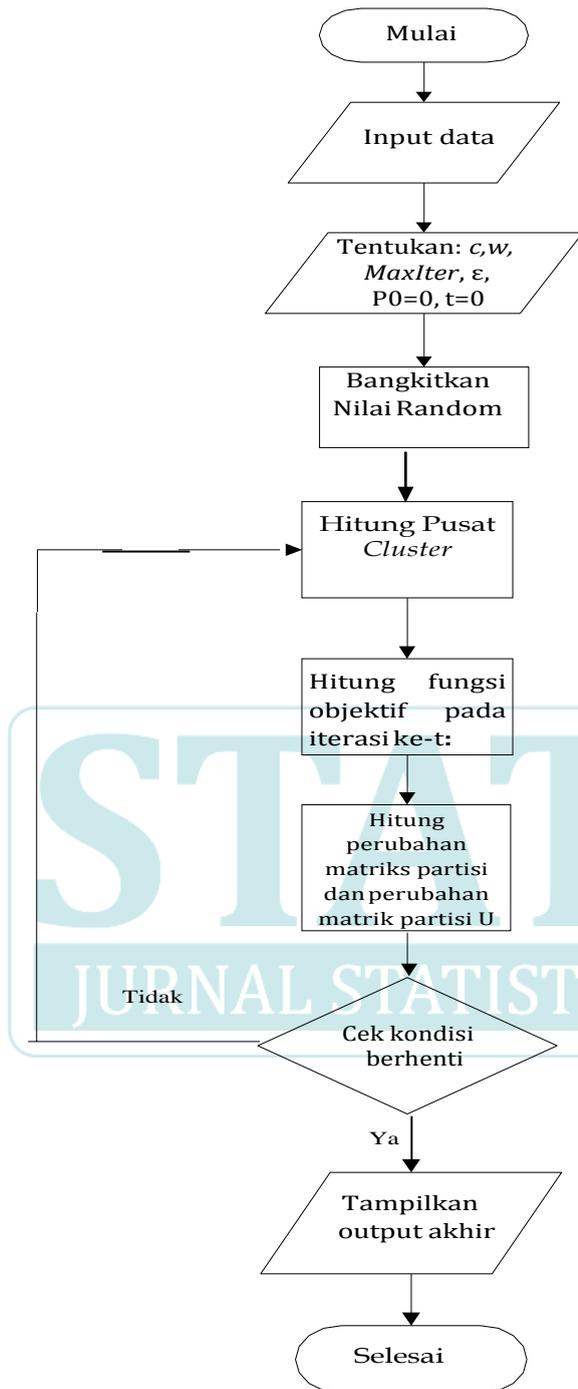
Dengan :

w = Derajat bobot

8. Cek kondisi berhenti

- a. Jika $(|P_t - P_{t-1}|) > \epsilon$ atau $t > maxIter$ maka berhenti.
- b. Jika tidak $t = t + 1$ dan diulangi Langkah ke-4,

Diagram Alir



Gambar 1. Diagram Alir

HASIL DAN PEMBAHASAN

Menginput data

Berikut merupakan data asli faktor-faktor yang mempengaruhi curah hujan di Provinsi Sumatera Utara berdasarkan pada 7 Stasiun BMKG yang kemudian diolah menggunakan

bantuan *software* R Studio.

Memeriksa *Missing Value*

Memeriksa *missing value* bertujuan untuk mengecek apakah terdapat kesalahan dalam pengukuran atau terdapat kekuranglengkapan pada data tersebut. Apabila hal tersebut terjadi maka perlu dilakukan upaya untuk mengatasinya. membandingkan hasil penelitian dengan teori dan hasil penelitian sebelumnya.

```

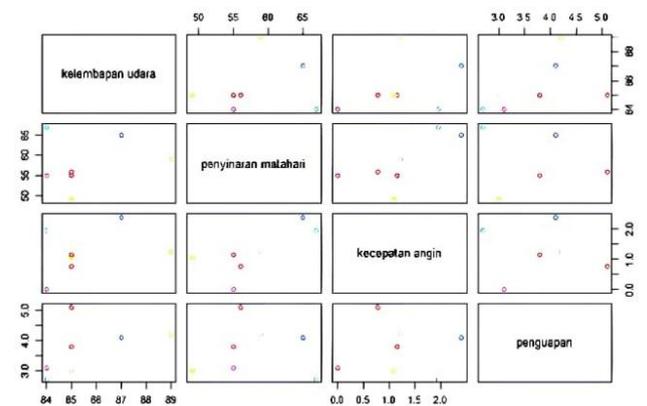
    [1,] kelembapan udara penyinaran matahari kecepatan angin penguapan
    [2,] FALSE FALSE FALSE FALSE FALSE
    [3,] FALSE FALSE FALSE FALSE FALSE
    [4,] FALSE FALSE FALSE FALSE FALSE
    [5,] FALSE FALSE FALSE FALSE FALSE
    [6,] FALSE FALSE FALSE FALSE FALSE
    [7,] FALSE FALSE FALSE FALSE FALSE
  
```

Gambar 2. Mengecek *Missing Value*

Berdasarkan pada output R Studio tersebut dapat ditarik kesimpulan jika data pada penelitian ini tidak mengandung *missing value*. Sehingga tidak perlu dilakukan tindakan lain yang bertujuan untuk mengubah data. Maka dapat dilanjutkan ke tahapan berikutnya.

Menampilkan Matriks/Berpasangan

Kegunaan pada tahapan ini adalah untuk memperlihatkan visualisasi pada data menggunakan plot berbentuk pair plot yang diperlihatkan dengan corak berwarna.



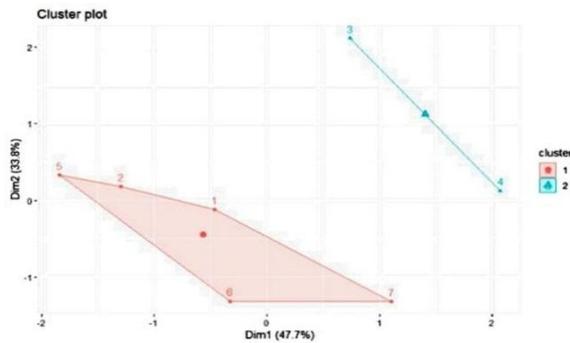
Gambar 3. *Pair Plot*

Korelasi Antar Atribut

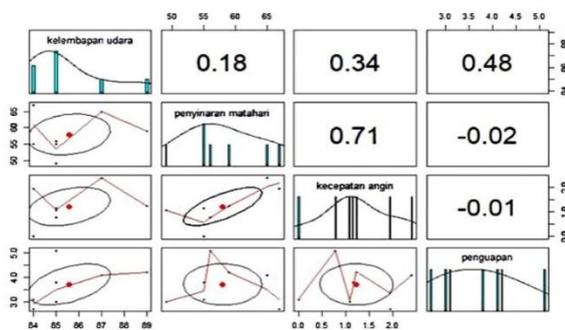
Korelasi antar atribut sebagai penggambaran terkait hubungan antar variabel atau yang memiliki nama lain sebagai atribut pada data yang digunakan dalam penelitian ini. Dan juga melalui hasil

Pengelompokan Faktor-Faktor yang Mempengaruhi Curah Hujan di....

pengecekan korelasi ini berguna dalam membantu kita untuk melihat pola pada suatu data.



Gambar 4. Korelasi Antar Atribut



Gambar 5. Korelasi Antar Atribut

Melalui hasil *output* Rstudio pada Gambar 5, dapat dilihat jika suatu nilai besarnya mendekati 1 maka diartikan bahwa korelasi antar variabel di dalam data semakin kuat. Apabila hasilnya mengandung angka negatif menandakan jika salah satu nilainya bertambah, maka korelasinya akan menurun (berbanding terbalik). Jika angka tersebut positif maka salah satu nilainya turun, korelasinya akan bertambah (berbanding searah).

Menentukan banyaknya cluster (k)

Di dalam melakukan metode *clustering* pada awal langkah yang dilakukan adalah dengan menentukan banyaknya *cluster* yang akan digunakan. Pada penelitian ini kami memutuskan untuk menggunakan $k = 2$.

	kelembapan udara	penyinaran matahari	kecepatan angin	penguapan
Cluster 1	85.09351	54.34710	0.784407	3.799064
Cluster 2	85.83366	65.27946	2.087147	3.485163

Clustering vector:
[1] 1 1 2 2 1 1 1

Gambar 6. Banyaknya Cluster

Melihat pada *output* tersebut diketahui bahwa 7 stasiun BMKG pada Provinsi Sumatera Utara terbagi menjadi 2 *cluster*, yang mana *cluster* 1 terdiri dari Stasiun Deli Serdang, Stasiun BMKG Wilayah I, Stasiun Marihat, Stasiun Pinangsori, dan Stasiun Gunung Sitoli. Sedangkan anggota dari *cluster* 2 yaitu Stasiun Belawan dan Stasiun Tanjung Morawa.

	kelembapan udara	penyinaran matahari	kecepatan angin	penguapan
kelembapan udara	1.0000000	0.17667813	0.34071158	0.4756507
penyinaran matahari	0.1766781	1.00000000	0.713567193	-0.02225324
kecepatan angin	0.3407112	0.71356719	1.000000000	-0.009351054
penguapan	0.4756507	-0.02225324	-0.009351054	1.000000000

Gambar 7. Visualisasi Cluster

Visualisasi *cluster* menggunakan plot tersebut menampilkan dengan jelas hasil pengelompokan menggunakan metode *Fuzzy C-Means Clustering* untuk data pada penelitian ini. Berdasarkan warna pada Gambar 6. Memperlihatkan warna merah berada pada *cluster* 1 dan warna biru menandakan keanggotaan pada *cluster* 2. Selain itu, Gambar 6. juga sebagai gambar yang menunjukkan derajat keanggotaan dari setiap variabel di dalam data. Suatu derajat keanggotaan dikatakan derajat keanggotaan yang paling tinggi jika nilainya besarnya mendekati nilai satu.

Validasi

Validasi untuk *cluster* terdiri dari beberapa indeks validasi, yaitu *Fuzzy Silhouette Index* (FSI), *Partition Entropy Index* (PEI), *Partition Coefficient Index* (PCI), dan *Modified Partition Coefficient Index* (MPCI). Validasi bertujuan untuk mengukur pada kriteria indeks subjektif pengukuran cluster. Kriteria indeks disajikan dalam Tabel 2.

Tabel 2. Kriteria Indeks Subjektif Pengukuran Cluster

Nilai	Interpretasi
0,71 – 1,00	Struktur <i>cluster</i> sangat baik
0,51 – 0,70	Struktur <i>cluster</i> baik
0,26 – 0,50	Struktur <i>cluster</i> lemah
< 0,25	Struktur <i>cluster</i> buruk

Pengelompokan Faktor-Faktor yang Mempengaruhi Curah Hujan di....

Tabel 3. Validasi Cluster

Indeks	Nilai
FSI	0,8257
PEI	0,2268
PCI	0,8685
MPCI	0,7371

Berdasarkan Tabel 3. menunjukkan hasil validasi *cluster* menggunakan metode *Fuzzy C-Means Clustering* dengan FSI, PCI, dan MPCI menghasilkan nilai yang mengandung kualitas *cluster* yang sangat baik. Sedangkan pada PEI terlihat jika kualitas *cluster* dikatakan buruk. Maka berdasarkan indeks-indeks validasi *cluster* dapat ditarik kesimpulan suatu nilai validasi yang besarnya mendekati 1 memperlihatkan jika indeks FSI, PCI, dan MPCI telah dilakukan dengan sangat baik.

PENUTUP

Kesimpulan

Berdasarkan hasil penelitian menggunakan metode *Fuzzy C-Means Clustering* pada 7 Stasiun BMKG di Provinsi Sumatera Utara didapatkan kesimpulan hasil pengelompokan dengan 2 *cluster*, yang mana *cluster* 1 terdiri dari Stasiun Deli Serdang, Stasiun BMKG Wilayah I, Stasiun Marihat, Stasiun Pinangsori, dan Stasiun Gunung Sitoli. Sedangkan anggota dari *cluster* 2 yaitu Stasiun Belawan dan Stasiun Tanjung Morawa. Kemudian, mendapat hasil validasi *cluster* yang cocok digunakan pada metode ini dengan hasil yang menandakan kualitas *cluster* data sangat baik dengan nilai yang berada pada 0,71 – 1,00 atau dengan kata lain nilainya mendekati 1 yaitu FSI sebesar 0,826, PCI sebesar 0,868, dan MPCI sebesar 0,868. Dengan demikian, metode *Fuzzy C Means Clustering* dapat digunakan untuk mengelompokkan faktor-faktor yang mempengaruhi curah hujan dengan lokasi pada 7 Stasiun BMKG di Provinsi Sumatera Utara.

Saran

Membertimbangkan kemungkinan yang ada untuk menyertakan analisis tentang bagaimana faktor-faktor yang mempengaruhi curah hujan dapat dipengaruhi oleh seperti topografi, penggunaan lahan, atau faktor lain yang relevan di wilayah tersebut. Melakukan analisis geospasial untuk memahami lebih lanjut pola spasial dari *cluster* yang terbentuk. Ini dapat memberikan wawasan tambahan tentang sebaran faktor-faktor curah hujan di wilayah tersebut. Mengintegrasikan analisis spasial menggunakan peta dapat memberikan visualisasi yang kuat terkait pola spasial dari *cluster* yang terbentuk. Ini dapat memudahkan pemahaman bagi pembaca yang tidak terbiasa dengan analisis data.s.i.

DAFTAR PUSTAKA

- BPS. (2021). Produksi Kelapa Sawit Di Indonesia. Badan Pusat Statistik, 21(1), 7–18.
- Chaniago, N. (2023). The Effect of Rainfall on Rice Production and Productivity in Percut Sei Tuan District, Deli Serdang Regency, North Sumatra. *AGRILAND Jurnal Ilmu Pertanian*, 11(3), 130–136. <https://jurnal.uisu.ac.id/index.php/agriland>
- Hidayat, A. (2012). Simple Random Sampling: Pengertian, Jenis, Cara dan Contohnya. *Simple Random Sampling: Pengertian, Jenis, Cara Dan Contohnya*. <https://www.statistikian.com/2018/02/pengertian-simple-random-sampling.html>
- Ispramono Hadi, S., Ermatita, E., & Chamidah, N. (2022). Penerapan Fuzzy C-Means dan Fuzzy Sugeno dalam Memprediksi Cuaca. *Informatik: Jurnal Ilmu Komputer*, 18(1), 11. <https://doi.org/10.52958/iftk.v17i4.4103>
- Latipa Sari, H., & Suranti, D. (2016). Perbandingan Algoritma Fuzzy C-Means (FCM) Dan Algoritma Mixture Dalam Penclustering Data Curah Hujan Kota

Pengelompokan Faktor-Faktor yang Mempengaruhi Curah Hujan di....

- Bengkulu. In Seminar Nasional Aplikasi Teknologi Informasi (SNATi).
- Novianti, F., Aisyah Yasmin, Y. R., & Novitasari, D. C. R. (2022). Penerapan Algoritma Fuzzy C-Means (FCM) dalam Pengelompokan Provinsi di Indonesia berdasarkan Indikator Penyakit Menular Manusia. *JUMANJI (Jurnal Masyarakat Informatika Unjani)*, 6(1), 23. <https://doi.org/10.26874/jumanji.v6i1.103>
- Nugroho, A. (2022). Banyak Faktor Pengaruhi Tingginya Curah Hujan. Banyak Faktor Pengaruhi Tingginya Curah Hujan. <https://ugm.ac.id/id/berita/22459-pengamat-ugm-banyak-faktor-pengaruh-tingginya-curah-hujan/>
- Prasetyo, B., Irwandi, H., & Pusparini, N. (2018). Karakteristik Curah Hujan Berdasarkan Ragam Topografi Di Sumatera Utara. *Jurnal Sains & Teknologi Modifikasi Cuaca*, 19(1), 11. <https://doi.org/10.29122/jstmc.v19i1.2787>
- Rahakbauw, D. L., Ilwaru, V. Y. I., & Hahury, M. H. (2019). Implementasi Fuzzy C-Means Clustering Dalam Implementation Of Fuzzy C-Means Clustering In. *Jurnal Ilmu Matematika Dan Terapan*, 11, 1-12. <https://media.neliti.com/media/publications/277582-implementasi-fuzzy-c-means-clustering-da-3afa5ba1.pdf>
- Rouza, E., & Fimawahib, L. (2020). Implementasi Fuzzy C-Means Clustering dalam Pengelompokan UKM Di Kabupaten Rokan Hulu. *Techno.Com*, 19(4), 481-495. <https://doi.org/10.33633/tc.v19i4.4101>
- Sanusi, W., Zaky, A., & Afni, B. N. (2020). Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-faktor Penyebab Gizi Buruk. *Journal of Mathematics, Computations, and Statistics*, 2(1), 47. <https://doi.org/10.35580/jmathcos.v2i1.12458>
- Satya Nugraha, G., & Amelia Riyandari, B. (2020). Implementasi Fuzzy C-Means Untuk Pengelompokan Daerah Berdasarkan Indikator Kesehatan. *Jurnal Teknologi Informasi*, 4(1), 56-62.
- Setiawan, D. (2021). Analisis Curah Hujan di Indonesia untuk Memetakan Daerah Potensi Banjir dan Tanah Longsor dengan Metode Cluster Fuzzy C-Means dan Singular Value Decomposition (SVD). *Engineering, Mathematics and Computer Science (EMACS) Journal*, 3(3), 115-120. <https://doi.org/10.21512/emacsjournal.v3i3.7428>
- Sumut, B. (2020). Rata-rata Kelembaban Udara, Curah Hujan, Penyinaran Matahari, Kecepatan Angin, dan Penguapan menurut Stasiun, 2020. Rata-Rata Kelembaban Udara, Curah Hujan, Penyinaran Matahari, Kecepatan Angin, Dan Penguapan Menurut Stasiun, 2020. <https://sumut.bps.go.id/statictable/2021/04/19/2069/rata-rata-kelembaban-udara-curah-hujan-penyinaran-matahari-kecepatan-angin-dan-penguapan-menurut-stasiun-2020.html>
- Tanjung, F. A., Windarto, A. P., & Fauzan, M. (2021). Penerapan Metode K-Means Pada Pengelompokan Pengangguran Di Indonesia. *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, 6(1), 61. <https://doi.org/10.30645/jurasik.v6i1.271>
- Tukidi. (2010). Karakter Curah Hujan Di

Indonesia. Jurnal Geografi, 7(2),
136-145.

<http://journal.unnes.ac.id/nju/index.php/JG/article/view/84>

Turrahma, R. N., Nanda Caesario, A.,
Alfajri, M. D., Gusmanto, R., &
Oktoberza, W. K. (2023).
Implementasi Fuzzy C-Means Untuk
Clustering Data Harga Saham
Harian Pada PT. Astra International
TBK. Jurnal Rekursif, 11, 64-69.
<https://ejournal.unib.ac.id/rekursif/article/view/27167/12023>

STATKOM
JURNAL STATISTIKA DAN KOMPUTASI