

Analisis Pola Kondisi Atmosfer Berdasarkan Data Cuaca Historis Menggunakan Metode K-Means Clustering

Analysis of Atmospheric Condition Patterns Based on Historical Weather Data Using the K-Means Clustering Method

Yudisti Prayigo Permana ^{a,1*}, Fingki Marwati^{a,2}

¹²Sistem Informasi, Universitas Pamulang, Jl. Raya Puspitek, Tangerang Selatan 15310, Indonesia

¹dosen03142@unpam.ac.id; ²dosen02817@unpam.ac.id

*Corresponding author

Received:	Revised:	Accepted:	Available online:
21.09.2025	14.10.2025	21.10.2025	30.10.2025

Abstrak: Curah hujan merupakan parameter penting dalam kajian kondisi cuaca karena berpengaruh terhadap berbagai bidang, seperti pertanian, transportasi, dan mitigasi bencana. Penelitian ini bertujuan untuk mengidentifikasi pola kondisi atmosfer yang berkaitan dengan karakteristik curah hujan menggunakan data cuaca historis melalui metode K-Means Clustering. Dataset yang digunakan terdiri atas 96.453 data pengamatan dengan variabel utama meliputi suhu, kelembapan, dan kecepatan angin. Tahap awal penelitian meliputi prapemrosesan data berupa pemilihan fitur dan normalisasi. Jumlah cluster optimal ditentukan menggunakan metode Elbow dan diperoleh tiga cluster sebagai konfigurasi terbaik. Proses pengelompokan dilakukan dengan algoritma K-Means menggunakan jarak Euclidean. Hasil pengelompokan menunjukkan bahwa setiap cluster memiliki karakteristik kondisi atmosfer yang berbeda, yaitu kondisi lembap dengan suhu rendah, kondisi peralihan dengan dominasi kecepatan angin tinggi, serta kondisi relatif kering dengan suhu tinggi dan kelembapan rendah. Evaluasi kualitas clustering menggunakan Silhouette Score menghasilkan nilai rata-rata sebesar 0,368 yang menunjukkan kualitas pengelompokan berada pada kategori cukup baik. Selain itu, nilai terendah pada Davies–Bouldin Index diperoleh pada jumlah tiga cluster, sehingga mendukung pemilihan konfigurasi tersebut sebagai jumlah cluster optimal. Berdasarkan hasil tersebut, metode K-Means Clustering terbukti mampu mengelompokkan data cuaca historis ke dalam pola kondisi atmosfer yang berbeda dan dapat digunakan sebagai pendekatan pendukung dalam analisis karakteristik curah hujan berbasis data historis.

Kata kunci: kondisi atmosfer, clustering, K-Means, Elbow Method, Silhouette Score, Davies–Bouldin Index

Abstract: Rainfall is an important parameter in weather condition analysis because it affects various sectors, including agriculture, transportation, and disaster mitigation. This study aims to identify atmospheric condition patterns related to rainfall characteristics using historical weather data through the K-Means Clustering method. The dataset consists of 96,453 observations with the main variables including temperature, humidity, and wind speed. The initial stage of the study involves data preprocessing, which includes feature selection and data normalization. The optimal number of clusters is determined using the Elbow method, resulting in three clusters as the best configuration. The clustering process is then carried out using the K-Means algorithm with Euclidean distance. The clustering results indicate that each cluster represents different atmospheric condition characteristics, namely humid conditions with low temperature, transitional conditions dominated by high wind speed, and relatively dry conditions with high temperature and low humidity. The evaluation of clustering performance using the Silhouette Score produces an average value of 0.368, indicating a fairly good clustering quality. In addition, the lowest Davies–Bouldin Index value is obtained at three clusters, supporting the selection of this configuration as the optimal number of clusters. Based on these results, the K-Means Clustering method is able to group historical weather data into distinct atmospheric condition patterns and can be used as a supporting approach for analyzing rainfall characteristics based on historical data.

Keywords: atmospheric conditions, clustering, K-Means, Elbow Method, Silhouette Score, Davies–Bouldin Index

1. PENDAHULUAN

Curah hujan merupakan komponen iklim yang berperan penting dalam berbagai sektor, seperti pertanian, pengelolaan sumber daya air, perencanaan wilayah, dan mitigasi bencana hidrometeorologi. Variasi dan ketidakpastian pola curah hujan dapat menyebabkan permasalahan seperti banjir, kekeringan, serta gangguan aktivitas sosial dan ekonomi. Oleh karena itu, analisis

terhadap karakteristik curah hujan menjadi penting sebagai dasar perencanaan dan pengambilan keputusan yang tepat [1].

Ketersediaan data cuaca historis dalam jumlah besar membuka peluang untuk melakukan analisis iklim secara lebih mendalam. Namun, besarnya volume data menuntut penggunaan metode yang mampu mengolah dan mengekstraksi informasi secara efisien. Teknik data mining, khususnya clustering, dapat digunakan untuk mengelompokkan data berdasarkan kemiripan karakteristik sehingga pola tersembunyi dapat diidentifikasi [2].

Dalam praktiknya, data curah hujan tidak selalu tersedia secara lengkap dan kontinu. Oleh karena itu, parameter atmosfer seperti suhu, kelembapan, dan kecepatan angin dapat dimanfaatkan sebagai indikator tidak langsung untuk menggambarkan kondisi atmosfer yang berkaitan dengan potensi hujan. Metode K-Means digunakan dalam penelitian ini untuk mengelompokkan data cuaca berdasarkan ketiga variabel tersebut guna memperoleh pola kondisi atmosfer yang serupa [3]. Penelitian ini bertujuan untuk mengidentifikasi pola kondisi atmosfer yang berkaitan dengan karakteristik curah hujan menggunakan metode K-Means Clustering serta mengevaluasi kualitas pengelompokan dengan metode Elbow, Silhouette Score, dan Davies-Bouldin Index. Hasil penelitian diharapkan dapat menjadi referensi dalam pengelolaan sumber daya air dan mitigasi risiko bencana terkait curah hujan.

2. TINJAUAN PUSTAKA

Curah hujan merupakan unsur penting dalam sistem iklim yang berpengaruh besar terhadap berbagai sektor, seperti pertanian, pengelolaan sumber daya air, perencanaan wilayah, serta mitigasi bencana hidrometeorologi. Perubahan pola curah hujan dapat menimbulkan dampak negatif, antara lain banjir, kekeringan, dan gangguan aktivitas sosial ekonomi. Oleh karena itu, analisis pola curah hujan diperlukan untuk memahami karakteristik iklim suatu wilayah dan mendukung pengambilan keputusan yang tepat. Data curah hujan berkaitan erat dengan parameter meteorologi lain, seperti suhu, kelembapan, kecepatan angin, tekanan udara, dan tutupan awan, yang bersama-sama memengaruhi intensitas serta distribusi hujan.

Perkembangan teknologi informasi dan ketersediaan data cuaca historis mendorong pemanfaatan teknik data mining dalam analisis iklim. Salah satu teknik utama adalah clustering, yaitu pengelompokan data berdasarkan kemiripan karakteristik tanpa menggunakan label kelas. Dalam bidang meteorologi, clustering digunakan untuk mengelompokkan kondisi cuaca, mengidentifikasi zona iklim, serta menganalisis variasi pola curah hujan [4].

Algoritma K-Means merupakan metode clustering yang banyak digunakan karena sederhana, efisien, dan mudah diterapkan. Metode ini mengelompokkan data berdasarkan jarak terdekat terhadap pusat cluster. Sebelum proses clustering dilakukan, diperlukan tahap prapemrosesan yang meliputi pembersihan data, pemilihan fitur, dan normalisasi agar skala antar variabel seragam.

Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow berdasarkan nilai Within-Cluster Sum of Squares (WCSS). Kualitas hasil clustering kemudian dievaluasi menggunakan Silhouette Score dan Davies-Bouldin Index. Nilai Silhouette yang tinggi dan Davies-Bouldin yang rendah menunjukkan kualitas pengelompokan yang baik. Berdasarkan penelitian sebelumnya, kombinasi metode K-Means dan teknik evaluasi tersebut efektif digunakan untuk menganalisis pola curah hujan berdasarkan data historis [5].

3. METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode data mining untuk menganalisis pola curah hujan berdasarkan data historis. Metode yang digunakan dalam penelitian ini adalah K-Means Clustering, yaitu salah satu algoritma pengelompokan data tanpa supervisi (*unsupervised learning*) yang bertujuan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristik tertentu [6] [7].

Data yang digunakan merupakan data cuaca historis yang diperoleh dari file *weatherHistory.csv* dengan jumlah 96.453 data pengamatan per jam. Dataset ini memiliki beberapa atribut meteorologi, antara lain waktu pengamatan (*Formatted Date*), kondisi cuaca (*Summary*),

jenis presipitasi (*Precip Type*), suhu udara (*Temperature*), suhu semu (*Apparent Temperature*), kelembapan (*Humidity*), kecepatan angin (*Wind Speed*), arah angin (*Wind Bearing*), jarak pandang (*Visibility*), tutupan awan (*Cloud Cover*), tekanan udara (*Pressure*), dan ringkasan harian (*Daily Summary*)[8]. Deskripsi masing-masing atribut disajikan pada Tabel 1.

Tabel 1. Data Set Cuaca

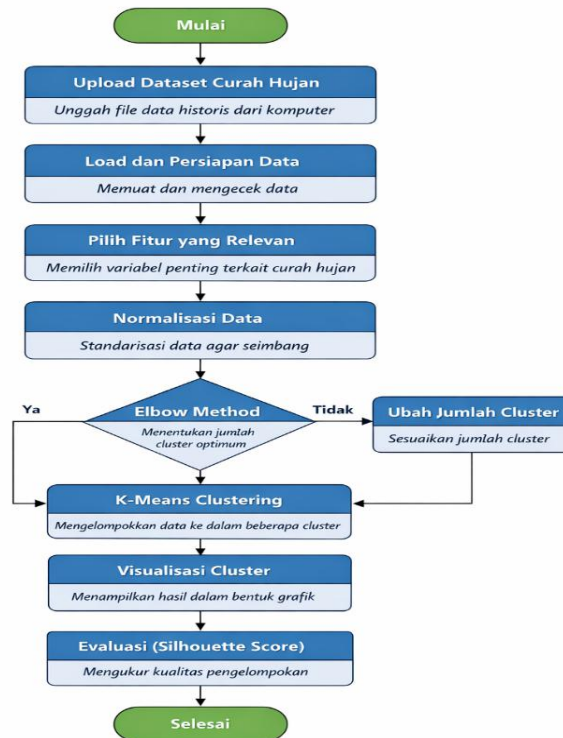
No	1	2	3	4	5
Formatted Date	2006-04-01 00:00:00 +0200	2006-04-01 01:00:00 +0200	2006-04-01 02:00:00 +0200	2006-04-01 03:00:00 +0200	2006-04-01 04:00:00 +0200
Summary	Partly Cloudy	Partly Cloudy	Mostly Cloudy	Partly Cloudy	Mostly Cloudy
Precip Type	Rain	Rain	Rain	Rain	Rain
Temperature (°C)	9.47	9.36	9.38	8.29	8.76
Apparent Temperature (°C)	7.39	7.23	9.38	5.94	6.98
Humidity	0.89	0.86	0.89	0.83	0.83
Wind Speed (km/h)	14.12	14.26	3.93	14.10	11.04
Wind Bearing (°)	251.0	259.0	204.0	269.0	259.0
Visibility (km)	15.83	15.83	14.96	15.83	15.83
Cloud Cover	0.0	0.0	0.0	0.0	0.0
Pressure (millibars)	1015.13	1015.63	1015.94	1016.41	1016.51
Daily Summary	Partly cloudy throughout the day	Partly cloudy throughout the day	Partly cloudy throughout the day	Partly cloudy throughout the day	Partly cloudy throughout the day

Pada penelitian ini, atribut yang digunakan sebagai fitur utama dalam proses clustering adalah suhu (*Temperature*), kelembapan (*Humidity*), dan kecepatan angin (*Wind Speed*), karena ketiga atribut tersebut memiliki keterkaitan langsung dengan karakteristik kondisi hujan. Tahap awal penelitian dilakukan dengan melakukan eksplorasi data dan preprocessing dengan menghapus data yang memiliki nilai kosong. Selanjutnya, data dinormalisasi menggunakan metode *StandardScaler* untuk menyamakan skala antar variabel.

Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow dengan menghitung nilai *Within-Cluster Sum of Squares* (WCSS)[9]. Proses pengelompokan kemudian dilakukan menggunakan algoritma K-Means Clustering dengan jarak Euclidean sebagai ukuran kedekatan antar data. Hasil clustering divisualisasikan dalam bentuk grafik sebar untuk melihat distribusi dan pemisahan antar cluster [10].

Evaluasi kualitas hasil clustering dilakukan menggunakan Silhouette Score dan Davies-Bouldin Index untuk menilai tingkat pemisahan antar cluster dan kualitas pengelompokan data. Tahapan penelitian yang dilakukan dalam penelitian ini disajikan secara ringkas dalam bentuk diagram alur pada Gambar 1.

Gambar 1 memperlihatkan alur tahapan penelitian analisis kondisi atmosfer menggunakan metode K-Means Clustering. Proses dimulai dengan pengunggahan dan pemuatan dataset, dilanjutkan dengan pemilihan fitur serta normalisasi data. Selanjutnya jumlah cluster optimal ditentukan menggunakan metode Elbow, kemudian proses clustering dilakukan dengan algoritma K-Means. Hasil pengelompokan divisualisasikan dan dievaluasi menggunakan Silhouette Score sebelum penelitian diakhiri.



Gambar 1. Flowchart Tahapan Penelitian Analisis Kondisi Atmosfer

4. HASIL DAN PEMBAHASAN

4.1 Deskripsi dan Eksplorasi Data

Tahap awal penelitian dilakukan dengan eksplorasi dataset cuaca historis untuk mengetahui karakteristik data yang digunakan. Statistik deskriptif data ditunjukkan pada **tabel 2**, yang memperlihatkan nilai minimum, maksimum, rata-rata, dan standar deviasi dari setiap variabel numerik.

Tabel 2. Statistik Deskriptif Dataset Cuaca

Statistik	Temperature (°C)	Apparent Temperature (°C)	Humidity	Wind Speed (km/h)	Wind Bearing (°)	Visibility (km)	Cloud Cover	Pressure (millibars)
Count	96.453	96.453	96.453	96.453	96.453	96.453	96.453	96.453
Mean	11.93	10.86	0.73	10.81	187.51	10.35	0.00	1003.24
Std	9.55	10.70	0.20	6.91	107.38	4.19	0.00	116.97
Min	-21.82	-27.72	0.00	0.00	0.00	0.00	0.00	0.00
25%	4.69	2.31	0.60	5.83	116.00	8.34	0.00	1011.90
50% (Median)	12.00	12.00	0.78	9.97	180.00	10.05	0.00	1016.45
75%	18.84	18.84	0.89	14.14	290.00	14.81	0.00	1021.09
Max	39.91	39.34	1.00	63.85	359.00	16.10	0.00	1046.38

Berdasarkan hasil tersebut, diketahui bahwa data terdiri dari 96.453 pengamatan dengan variasi suhu antara $-21,82^{\circ}\text{C}$ hingga $39,91^{\circ}\text{C}$, kelembapan antara 0 hingga 1, serta kecepatan angin hingga 63,85 km/jam.

Selanjutnya dilakukan pemilihan fitur yang relevan terhadap analisis pola curah hujan, yaitu suhu (Temperature), kelembapan (Humidity), dan kecepatan angin (Wind Speed). Contoh

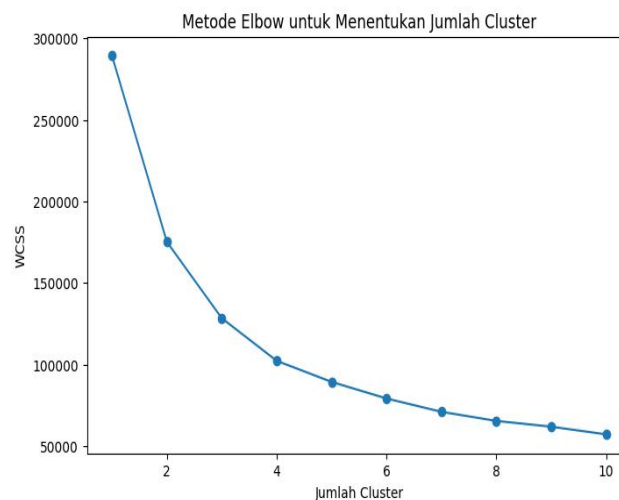
data hasil seleksi fitur ditunjukkan pada **Tabel 3**, yang memperlihatkan nilai ketiga variabel yang akan digunakan dalam proses clustering.

Tabel 3. Data Hasil Seleksi Fitur

No	Humidity	Temperature (°C)	Wind Speed (km/h)
1	0.89	9.47	14.12
2	0.86	9.36	14.26
3	0.89	9.38	3.93
4	0.83	8.29	14.10
5	0.83	8.76	11.04

4.2 Penentuan Jumlah Cluster dengan Metode Elbow

Setelah data dinormalisasi, penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow. Hasil perhitungan nilai *Within-Cluster Sum of Squares* (WCSS) ditunjukkan pada Gambar 4.



Gambar 4. Grafik Metode Elbow untuk Menentukan Jumlah Cluster

Grafik tersebut menunjukkan adanya titik siku pada nilai $K = 3$, sehingga jumlah cluster yang digunakan dalam penelitian ini adalah tiga cluster. Pemilihan jumlah cluster ini diharapkan mampu menghasilkan pengelompokan data yang optimal.

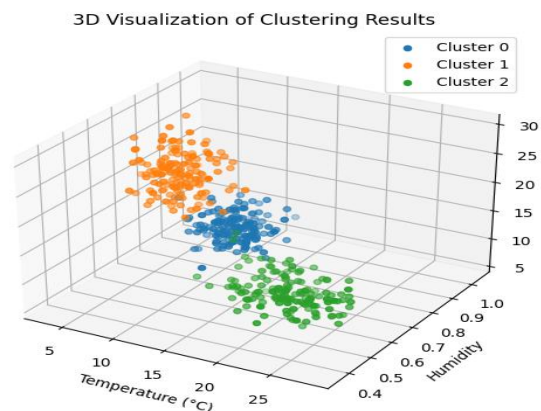
4.3 Hasil K-Means Clustering dan Visualisasi

Proses pengelompokan data dilakukan menggunakan algoritma K-Means dengan jumlah cluster sebanyak tiga. Hasil pemberian label cluster pada data ditunjukkan pada tabel 4.

Tabel 4. Data Hasil Clustering dengan Label Cluster

No	Humidity	Temperature (°C)	Wind Speed (km/h)	Cluster
1	0.89	9.47	14.12	0
2	0.86	9.36	14.26	0
3	0.89	9.38	3.93	0
4	0.83	8.29	14.10	0
5	0.83	8.76	11.04	0

Selanjutnya, hasil clustering divisualisasikan dalam bentuk grafik sebar antara kelembapan dan suhu seperti yang ditunjukkan pada Gambar 6.

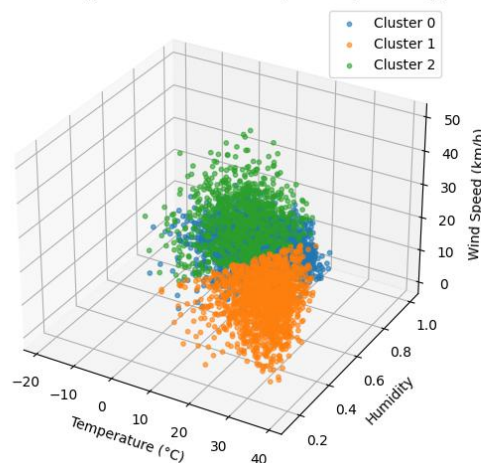


Gambar 6. Visualisasi Hasil Clustering Kondisi Atmosfer (Humidity vs Temperature)

Gambar 6 memperlihatkan visualisasi hasil clustering kondisi atmosfer menggunakan algoritma K-Means berdasarkan variabel kelembapan (Humidity) dan suhu (Temperature). Grafik menunjukkan bahwa data terbagi ke dalam tiga cluster yang ditandai dengan warna merah, biru, dan hijau. Cluster 0 (merah) didominasi oleh data dengan tingkat kelembapan tinggi dan suhu relatif rendah hingga sedang, yang menunjukkan kondisi atmosfer yang sangat lembap dengan kecenderungan terjadinya hujan. Cluster 1 (biru) berada pada rentang kelembapan sedang dengan suhu rendah hingga sedang, yang menggambarkan kondisi peralihan antara cuaca kering dan lembap. Sementara itu, Cluster 2 (hijau) didominasi oleh data dengan kelembapan rendah hingga sedang dan suhu relatif tinggi, yang menunjukkan kondisi cuaca yang lebih kering. Meskipun terdapat beberapa tumpang tindih antar cluster, hasil visualisasi menunjukkan bahwa algoritma K-Means mampu memisahkan data ke dalam kelompok dengan karakteristik cuaca yang berbeda secara cukup jelas.

Untuk memperjelas hasil pengelompokan berdasarkan tiga variabel yang digunakan, ditambahkan visualisasi tiga dimensi pada Gambar 7 yang menampilkan hubungan antara suhu (Temperature), kelembapan (Humidity), dan kecepatan angin (Wind Speed). Visualisasi ini menunjukkan pemisahan cluster yang lebih jelas pada ruang tiga dimensi, sehingga karakteristik masing-masing kelompok kondisi atmosfer dapat diamati secara lebih komprehensif. Setiap warna merepresentasikan satu cluster hasil algoritma K-Means, yang menggambarkan perbedaan pola kondisi atmosfer berdasarkan kombinasi ketiga parameter meteorologi tersebut.

3D Visualization of Clustering Results Based on Temperature, Humidity, and Wind Speed



Gambar 7. 3D Visualization of Clustering Results Based on Temperature, Humidity, and Wind Speed

Pada Gambar 7 terlihat bahwa pemisahan antar cluster menjadi lebih jelas ketika ketiga variabel, yaitu suhu, kelembapan, dan kecepatan angin, dianalisis secara bersamaan. Masing-masing cluster membentuk kelompok data yang relatif terpisah dalam ruang tiga dimensi, meskipun pada beberapa bagian masih terdapat tumpang tindih antar kelompok. Hasil ini menunjukkan bahwa kombinasi ketiga parameter meteorologi tersebut mampu merepresentasikan perbedaan karakteristik kondisi atmosfer secara lebih komprehensif. Dengan demikian, visualisasi tiga dimensi ini memperkuat hasil clustering yang diperoleh dan mendukung interpretasi pola kondisi atmosfer yang berkaitan dengan karakteristik curah hujan.

4.4 Analisis Karakteristik Tiap Cluster

Karakteristik masing-masing cluster ditunjukkan pada Tabel 5, yang memperlihatkan nilai rata-rata suhu, kelembapan, dan kecepatan angin pada setiap cluster.

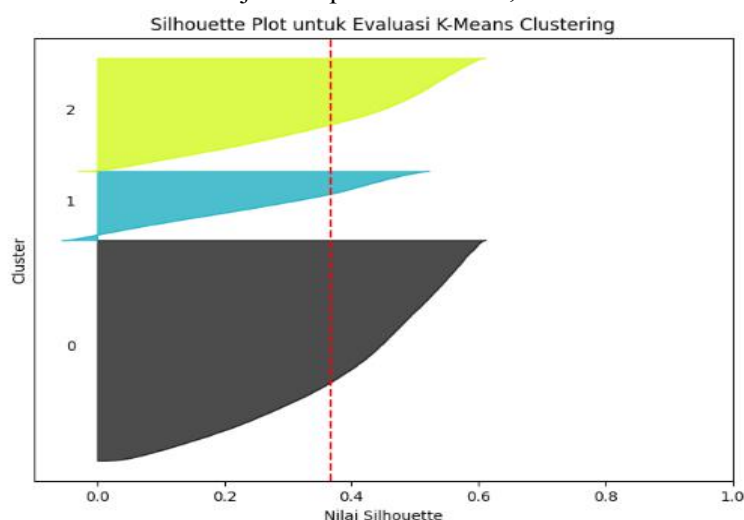
Tabel 5. Rata-Rata Nilai Variabel pada Setiap Cluster

Cluster	Humidity	Temperature (°C)	Wind Speed (km/h)
0	0.861973	7.38	7.65
1	0.715945	9.07	21.67
2	0.498390	22.55	10.38

Cluster 0 memiliki kelembapan tinggi (0.8619), suhu rendah (7.38 °C), dan kecepatan angin rendah (7.65 km/h) yang menunjukkan kondisi atmosfer lembap dengan potensi terjadinya hujan ringan hingga sedang. Cluster 1 memiliki kelembapan sedang (0.7159), suhu relatif rendah (9.07 °C), dan kecepatan angin tinggi (21.67 km/h) yang menggambarkan kondisi peralihan dengan dominasi angin kencang serta kemungkinan terjadinya hujan. Sementara itu, Cluster 2 memiliki kelembapan rendah (0.4983), suhu tinggi (22.55 °C), dan kecepatan angin sedang (10.38 km/h) yang mencerminkan kondisi atmosfer relatif kering.

4.5 Evaluasi Kualitas Clustering

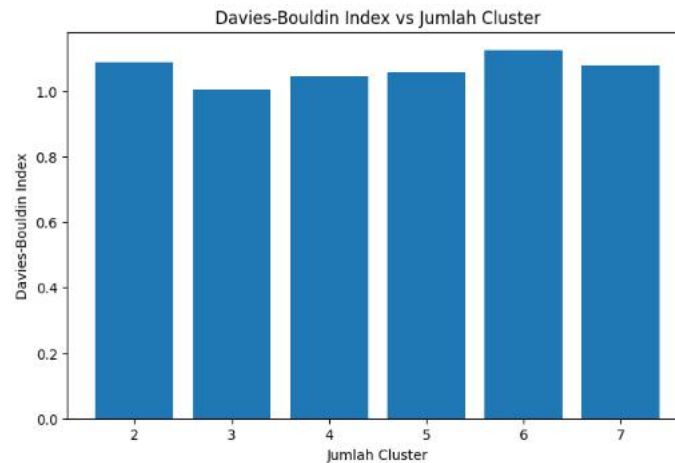
Evaluasi kualitas hasil clustering dilakukan menggunakan Silhouette Score dan Davies-Bouldin Index. Hasil Silhouette Plot ditunjukkan pada Gambar 8,



Gambar 8. Silhouette Plot

Gambar 8. menunjukkan hasil evaluasi kualitas clustering menggunakan Silhouette Plot pada proses K-Means Clustering. Grafik tersebut menampilkan distribusi nilai silhouette untuk setiap cluster, dengan garis putus-putus merah sebagai penanda nilai rata-rata Silhouette Score sebesar 0,368. Sebagian besar data pada masing-masing cluster memiliki nilai silhouette positif,

yang menunjukkan bahwa data telah dikelompokkan ke dalam cluster yang sesuai. Meskipun masih terdapat beberapa data dengan nilai silhouette mendekati nol, hasil ini mengindikasikan bahwa pemisahan antar cluster telah terbentuk dengan cukup baik. Dengan demikian, dapat disimpulkan bahwa hasil clustering yang diperoleh memiliki kualitas pengelompokan yang cukup memadai untuk digunakan dalam analisis kondisi atmosfer yang berkaitan dengan curah hujan.



Gambar 9. Evaluasi Davies – Bouldin Index

Gambar 9 menunjukkan hasil evaluasi kualitas clustering menggunakan Davies–Bouldin Index pada beberapa jumlah cluster. Grafik memperlihatkan bahwa nilai Davies–Bouldin Index terendah diperoleh pada jumlah cluster $K = 3$, yang mengindikasikan bahwa konfigurasi tiga cluster memberikan kualitas pengelompokan terbaik dibandingkan jumlah cluster lainnya. Nilai indeks yang lebih rendah menunjukkan bahwa jarak antar cluster semakin besar dan jarak dalam cluster semakin kecil, sehingga pemisahan antar cluster menjadi lebih baik. Dengan demikian, hasil evaluasi ini mendukung pemilihan tiga cluster. Dengan demikian, hasil evaluasi ini mendukung pemilihan tiga cluster sebagai jumlah cluster optimal dalam penelitian ini untuk analisis kondisi atmosfer yang berkaitan dengan curah hujan.

5. KESIMPULAN

Berdasarkan hasil penelitian terhadap data cuaca historis menggunakan metode K-Means Clustering, dapat disimpulkan bahwa proses pengelompokan data berhasil membentuk tiga cluster sebagai jumlah cluster optimal. Penentuan jumlah cluster ini didukung oleh hasil metode Elbow yang menunjukkan titik siku pada nilai $K = 3$ serta hasil evaluasi menggunakan Davies–Bouldin Index yang memberikan nilai terendah pada jumlah cluster tersebut. Hasil clustering menunjukkan bahwa setiap cluster memiliki karakteristik kondisi atmosfer yang berbeda berdasarkan variabel suhu, kelembapan, dan kecepatan angin. Cluster 0 merepresentasikan kondisi dengan kelembapan tinggi dan suhu rendah yang mengindikasikan potensi hujan yang lebih tinggi, Cluster 1 menunjukkan kondisi peralihan dengan kecepatan angin tinggi, sedangkan Cluster 2 menggambarkan kondisi relatif kering dengan suhu tinggi dan kelembapan rendah. Perbedaan karakteristik ini menunjukkan bahwa metode K-Means mampu mengelompokkan data cuaca ke dalam pola-pola kondisi atmosfer yang bermakna. Evaluasi kualitas clustering menggunakan Silhouette Score menghasilkan nilai rata-rata sebesar 0,368 yang berada pada kategori cukup baik, menunjukkan bahwa pemisahan antar cluster telah terbentuk meskipun masih terdapat tumpang tindih antar kelompok. Secara keseluruhan, metode K-Means Clustering terbukti efektif dalam mengelompokkan data cuaca historis ke dalam pola kondisi atmosfer yang berbeda dan dapat digunakan sebagai pendekatan pendukung dalam memahami

variasi kondisi cuaca Secara keseluruhan, metode K-Means Clustering terbukti efektif dalam mengelompokkan data cuaca historis ke dalam pola kondisi atmosfer yang berbeda dan dapat digunakan sebagai pendekatan pendukung dalam memahami variasi kondisi cuaca serta karakteristik potensi curah hujan dan perencanaan pengelolaan sumber daya air. dalam memahami variasi kondisi cuaca serta karakteristik potensi curah hujan dan perencanaan pengelolaan sumber daya air.

DAFTAR PUSTAKA

- [1] Jurnal Informatika Mulawarman, "Clustering Data Cuaca Ekstrem," *J. Inform. Mulawarman*, vol. 1089, pp. 1–10, 2022.
- [2] J. Suryanto, "Analisa Perbandingan Pengelompokkan Curah Hujan 15 Harian," *J. Agrifor*, vol. 16, no. 2, pp. 229–242, 2017.
- [3] Y. Anzari *et al.*, "PENERAPAN ALGORITMA K-MEANS UNTUK KLASTERISASI POLA IKLIM STUDI KASUS : PROVINSI JAMBI PERIODE 2020-2024," vol. 6, no. 3, pp. 1037–1052, 2025.
- [4] D. Peng, T. Zhou, S. Hu, and J. Zheng, "Ocean-driven shifts in circulation regime frequency modulate South China rainfall," *npj Clim. Atmos. Sci.*, vol. 8, no. 1, pp. 1–13, 2025.
- [5] I. F. Ashari, E. Dwi Nugroho, R. Baraku, I. Novri Yanda, and R. Liwardana, "Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta," *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 89–97, 2023.
- [6] M. Yashaswini, G. Sowmya, and P. Jai, "Weather Clustering Using K-Means," *Int. J. Creat. Res. Thoughts*, vol. 9, no. 6, pp. 2320–2882, 2021.
- [7] Uston Nawawi Christanto, Brina Miftahurrohman, T. Bariyah, H. Kuswanto, and N. Faria, "Cluster-Based Machine Learning Approaches for Predicting Daily Maximum Temperatures in Indonesia Under Climate Change," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 11, no. 1, pp. 236–249, 2025.
- [8] M. I. Drilanang, Z. Indra, A. P. Walidin, and T. S. Warman, "Development of an Adaptive Hybrid Weather Prediction Model Based on Pattern Classification and Deep Learning for Disaster Mitigation in Indonesia," vol. 4, no. 3, pp. 921–931, 2025.
- [9] M. Yana and H. Farsya, "Analisis Curah Hujan Dan Temperatur Rata-Rata Pada Data Meteorologi Bmkg Kota Jambi Menggunakan Metode Clustering K-Means," *J. Komput. dan Inform.*, vol. 15, no. 2, pp. 283–286, 2020.
- [10] N. Wahidah, O. Juwita, and F. N. Arifin, "Pengelompokkan Daerah Rawan Bencana di Kabupaten Jember Menggunakan Metode K-Means Clustering," *INFORMAL Informatics J.*, vol. 8, no. 1, p. 22, 2023.