

Pemanfaatan K-Means Clustering Dalam Pengelompokan Cuaca Di Wilayah Kediri

^{1*}Muhammad Firmandani Pramoedya, ²Risky Aswi Ramadhani, ³Ahmad
Bagus Setiawan

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹_1danipramoedya2002@gmail.com, ²riskyaswiramadhani@gmail.com,

³ahmadbagus@unpkediri.ac.id

Penulis Korespondens : Muhammad Firmandani Pramoedya

Abstrak — Cuaca merupakan faktor penting yang memengaruhi berbagai aspek kehidupan, terutama di wilayah Kediri, di mana perubahan cuaca berdampak signifikan pada pertanian dan aktivitas masyarakat. Penelitian ini bertujuan meningkatkan keakuratan informasi cuaca dengan mengelompokkan data cuaca harian menggunakan metode *k-means clustering* berdasarkan suhu, kelembapan, kecepatan angin, dan arah angin. Data dikumpulkan dari unggahan harian sosial media BMKG Kediri tahun 2024–2025 dan melalui tahap *preprocessing* berupa transformasi dan normalisasi data. Pemilihan jumlah *cluster* optimal dilakukan dengan metode *elbow* dan evaluasi kualitas *cluster* menggunakan *davies bouldin index* (DBI). Hasil menunjukkan bahwa jumlah *cluster* optimal adalah tiga dengan nilai DBI sebesar 0,472, menandakan pemisahan *cluster* yang baik dan kompak. Representasi arah angin yang diubah menjadi nilai derajat meningkatkan kualitas klasterisasi. Dengan demikian, sistem klasterisasi ini dapat memberikan informasi cuaca yang lebih spesifik dan akurat di tingkat kecamatan, membantu masyarakat dan pemangku kepentingan dalam mempersiapkan aktivitas sehari-hari serta mitigasi risiko bencana. Penelitian ini memberikan kontribusi pada pengembangan metode analisis data meteorologi yang aplikatif di wilayah Kediri.

Kata Kunci — Aplikasi Berbasis *Website*, Cuaca, *K-Means*, Pola Cuaca

Abstract — Weather is an important factor that affects various aspects of life, especially in the Kediri area, where weather changes have a significant impact on agriculture and community activities. This study aims to improve the accuracy of weather information by grouping daily weather data using the *k-means clustering* method based on temperature, humidity, wind speed, and wind direction. Data were collected from daily uploads of BMKG Kediri social media in 2024–2025 and went through a *preprocessing* stage in the form of data transformation and normalization. The selection of the optimal number of clusters was carried out using the *elbow* method and cluster quality evaluation using the *Davies-Bouldin Index* (DBI). The results show that the optimal number of clusters is three with a DBI value of 0.472, indicating a good and compact cluster indicator. The representation of wind direction converted into a degree value improves the quality of clustering. Thus, this clustering system can provide more specific and accurate weather information at the sub-district level, helping the community and stakeholders in preparing daily activities and mitigating disaster risks. This study contributes to the development of an applicable meteorological data analysis method in the Kediri area.

Keywords — Website Based Application, Weather, *K-Means*, Weather Pattern

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Pendahuluan Cuaca merupakan salah satu faktor penting terhadap segala aspek kehidupan. Yang dimaksud dengan cuaca adalah sebuah kondisi yang terjadi pada waktu dan tempat tertentu, dan berlangsung dalam jangka waktu yang singkat, seperti kondisi suatu wilayah cenderung panas atau dingin [1]. Karena banyak aktivitas tertentu yang dilakukan secara sengaja ataupun tidak sengaja, akibatnya cuaca menjadi tidak menentu [2]. Perubahan cuaca yang besar bisa memengaruhi kegiatan manusia. misalnya di wilayah Kediri, pada bidang pertanian para petani sering kali menghadapi kesulitan ketika sedang merawat tanaman karena pasokan air cukup terbatas ketika tiba-tiba memasuki cuaca panas tidak ada hujan, atau para petani mengalami kerugian ketika tiba-tiba terdapat badai berangin yang dapat merusak tanaman mereka. Dampak dari cuaca yang tidak menentu juga terjadi pada lingkungan, yaitu pada bencana pohon tumbang. Banyak pohon-pohon besar tumbang dikarenakan cuaca berangin yang tidak terprediksi dengan sempurna tumbang di rumah warga dan di jalan raya. Dampak-dampak tersebut tentu jelas menghambat produktivitas warga, khususnya warga di wilayah Kediri [3]. Ramalan cuaca sangat penting dalam kehidupan sehari-hari. Misalnya, petani bisa tahu kapan waktu yang tepat untuk menanam, transportasi jadi lebih aman, dan pemerintah bisa siap lebih awal jika ada bencana seperti angin kencang [4].

Namun adanya sebuah ramalan cuaca sering kali tidak benar-benar akurat dalam memprediksi sehingga masalah dan bencana yang ada masih sulit untuk di atasi secara sempurna, walaupun dengan adanya ramalan cuaca tersebut banyak yang pihak terbantu. Maka dari itu penelitian ini dilakukan untuk dapat meningkatkan keakuratan informasi mengenai cuaca dari suatu wilayah. Agar bisa tahu dan memantau informasi mengenai cuaca lebih lanjut, dibutuhkan cara untuk mengumpulkan dan mengolah data. Dari situ, informasi cuaca bisa dibuat dan dibagikan ke masyarakat supaya bisa dimanfaatkan dalam kehidupan sehari-hari [5]. Pada penelitian ini akan dilakukan sebuah teknik bernama *clustering*. *Clustering* adalah proses mengelompokkan data yang belum punya label ke dalam beberapa grup. Tujuannya adalah agar data yang mirip satu sama lain berada dalam grup yang sama, dan data yang berbeda berada di grup yang berbeda [6]. Jadi dengan menggunakan teknik tersebut, nantinya dapat menampilkan informasi mengenai kondisi cuaca di suatu wilayah Kediri cenderung ke arah mana. Pada teknik klasterisasi terdapat sebuah algoritma *k-means* yang dapat mengelompokkan data berdasarkan kedekatan/kemiripan data sehingga data-data yang memiliki karakteristik yang sama akan dimasukkan ke dalam *cluster* yang sama [7]. Metode ini bekerja dengan mencoba berbagai jumlah kelompok (K) dan mengukur seberapa baik data dikelompokkan. Menurut [8] algoritma *k-means* juga merupakan algoritma yang bersifat sederhana dalam pengimplementasian dan menjalankan, pengerjaan relatif cepat, dapat juga beradaptasi dan sudah sering digunakan. Metode *elbow* juga digunakan pada penelitian ini yang melibatkan penilaian kinerja cluster dengan berbagai nilai K dan kemudian dilakukan identifikasi titik *elbow*. Titik *elbow* adalah saat peningkatan hasil mulai melambat, dan itu jadi petunjuk jumlah kelompok yang paling pas. Dalam evaluasi klaster, salah satu kriteria umum yang digunakan selain *elbow method* adalah *davies-bouldin index* (DBI), yang mengukur dissimilaritas rata-rata antara klaster-klasternya dan membantu dalam menilai pemisahan *cluster*. Evaluasi DBI memberikan wawasan tentang

kualitas penugasan klaster dan membantu mengidentifikasi jumlah klaster yang paling tepat untuk *dataset* yang diberikan [9]. Sistem klasterisasi sebenarnya sudah tersedia pada halaman *website* Badan Meteorologi, Klimatologi dan Geofisika, namun, sistem tersebut tidak menampilkan suatu daerah secara spesifik, melainkan hanya menampilkan bagian kota sebagai perwakilan dari suatu daerah.

Maka dari itu pada penelitian ini akan melakukan klasterisasi lebih spesifik terhadap kecamatan yang ada di Kediri. Pengumpulan data cuaca didapat dari unggahan harian sosial media Badan Meteorologi, Klimatologi dan Geofisika Kediri pada tahun 2024 dan 2025 yang berisikan informasi mengenai cuaca pada setiap hari. Data yang dikumpulkan mencakup beberapa parameter antara lain suhu udara, kelembapan udara, kecepatan angin, dan arah angin. Proses pengumpulan data cuaca dilakukan dengan melakukan pencatatan setiap harinya ketika ada unggahan cuaca harian baru dari sosial media BMKG Kediri, sehingga data yang terkumpul dipastikan selalu terbaru. Data cuaca yang akurat dan analisis yang mendalam sangat diperlukan untuk dapat memahami pola cuaca dan dampaknya terhadap kehidupan [10]. Hasil dari sistem klasterisasi nantinya adalah masyarakat jadi lebih mengetahui mengenai sebuah informasi dari suatu wilayah di Kediri yang cenderung mengarah pada karakteristik cuaca tertentu seperti dingin, panas, lembap, atau kering. Dengan begitu masyarakat dan pihak-pihak yang berpentingan dapat mempersiapkan aktivitas ke depannya dengan lebih matang. Dilakukannya penelitian ini, diharapkan dapat mengenali potensi dari penggunaan *k-means clustering* dalam upaya untuk mengelompokkan data cuaca yang kemudian dapat dilakukan analisis lebih lanjut yang dapat memberikan sebuah informasi yang lebih spesifik terhadap suatu wilayah di Kediri dan dapat digunakan oleh masyarakat umum, khususnya untuk masyarakat Kediri. Hasil pada penelitian ini diharapkan dapat memberikan kontribusi dan informasi yang berarti terhadap ilmu meteorologi, klimatologi, dan geofisika di Kediri. Serta memberikan pandangan yang lebih luas terhadap pemanfaatan teknik klasterisasi dalam konteks ilmu meteorologi, klimatologi, dan geofisika.

II. METODE

A. *Preprocessing*

Setelah data berhasil dikumpulkan, pada tahap *preprocessing*, kegiatan yang dilakukan adalah mengetik kembali data yang diperoleh ke dalam Microsoft Excel untuk mencatat seluruh data yang telah didapatkan [11]. Proses ini penting karena data yang diterima masih berbentuk gambar, acak, dan belum terstruktur, sehingga perlu diolah dan dikelompokkan terlebih dahulu agar sesuai dengan kebutuhan analisis selanjutnya [12]. Salah satu *preprocessing* adalah normalisasi. Normalisasi adalah proses pengubahan skala yang berbeda dari beberapa nilai menjadi satu skala yang sama. Tujuan dari proses ini adalah untuk menyeimbangkan nilai pada setiap atribut dalam *dataset*, supaya tidak ada nilai yang mendominasi atau mengurangi pengaruh atribut lainnya [9].

B. *K-Means*

Algoritma klasterisasi yang digunakan adalah menggunakan *k-means clustering*. Algoritma *k-means* merupakan model yang menggunakan *centroid* untuk membuat *cluster* [13]. Langkah-

langkah dalam menerapkan algoritma *k-means* diawali dengan menentukan nilai cluster awal pada data. Kemudian dilanjutkan dengan penentuan nilai pusat (*centroid*). *Centroid* bisa diibaratkan titik pusat dari *cluster*, juga dikatakan dapat mewakili suatu *cluster*. Penentuan nilai *centroid* pada awalnya ditentukan secara acak yang nantinya setelah iterasi dilakukan, maka *centroid* akan berganti sesuai dengan kesamaan datanya.

$$Euclidian = \sqrt{\sum_i^n (x_i - y_i)^2} \quad (1)$$

Keterangan :

i = banyak data

n = jumlah dimensi

y_i, x_i = koordinat titik A dan B dalam dimensi ke- i

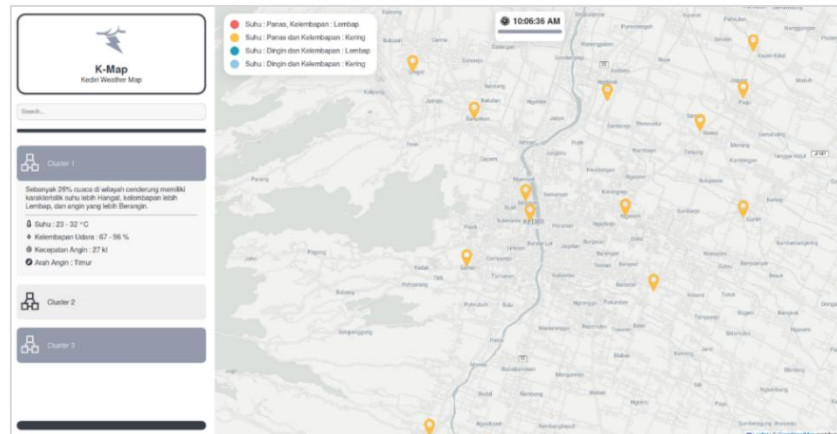
C. Elbow Method

Salah satu metode yang dapat digunakan untuk mengetahui jumlah *cluster* yang optimal yaitu menggunakan *elbow method*. *Elbow method* melakukan sebuah proses evaluasi dengan membandingkan hasil antara penjumlahan cluster yang akan membentuk suatu titik terakhir dan *cluster* nilai yang digunakan sebagai model data untuk *cluster* terbaik [14].

D. Davies Bouldin Index

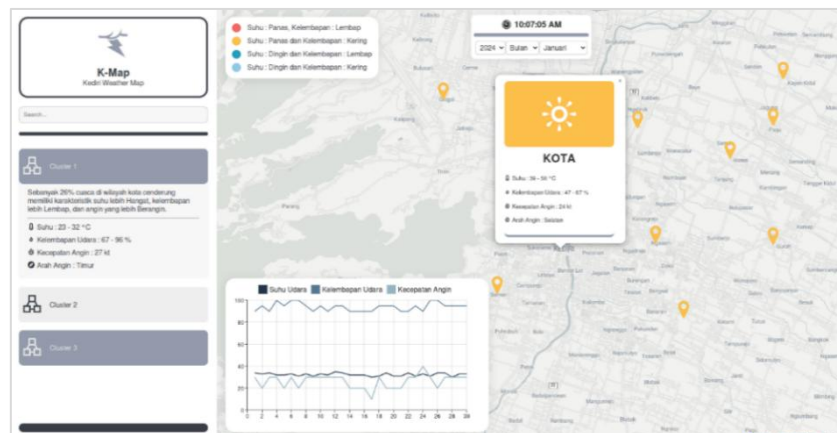
Metode *davies-bouldin index* digunakan untuk melihat apakah suatu *cluster* sudah optimal atau belum. DBI melakukan proses pengukuran suatu hasil pengelompokan berdasar pada nilai kohesi dan separasi [14]. Hasil pengelompokan kohesi mengacu pada seberapa dekat data-datanya dengan titik pusat (*centroid*) dari kelompok yang terbentuk. Sementara itu, pengelompokan separasi didasarkan pada jarak antar *centroid* dari masing-masing kluster. *davies-bouldin index* mengukur kualitas *clustering* dengan cara menilai jarak yang rendah di dalam *cluster* (*intra-cluster*) dan jarak yang tinggi antar *cluster*; semakin kecil nilai indeks ini, maka semakin baik hasil *clustering* tersebut [15].

E. Implementasi



Gambar 1. Tampilan Halaman Utama

Pada gambar (1) merupakan tampilan awal dari sistem klusterisasi cuaca di wilayah Kediri, di mana terdapat 29 titik lokasi yang memberikan informasi mengenai karakter dominan cuaca melalui representasi warna. Informasi mengenai warna dipilih dengan harapan pengguna dapat dengan mudah untuk menelaah informasi. Untuk warna merah cenderung panas dan warna biru untuk dingin. Pada bagian *sidebar* terdapat informasi mengenai hasil *cluster* yang merepresentasikan seberapa besar dominan karakter cuaca pada wilayah Kediri. Pada *sidebar* juga terdapat *searchbar* yang memungkinkan pengguna dapat melakukan pencarian juga pada bagian *searchbar* supaya lebih mudah dalam melakukan pencarian.



Gambar 2. Tampilan Halaman Utama Setelah Aksi

Pada gambar (2) merupakan tampilan setelah pengguna melakukan aksi *click* pada salah satu area. Tampilan tersebut menampilkan informasi tambahan mengenai rata-rata karakteristik cuaca dan grafik mengenai suhu, kelembapan, dan kecepatan angin pada area yang dipilih.

F. Simulasi Algoritma

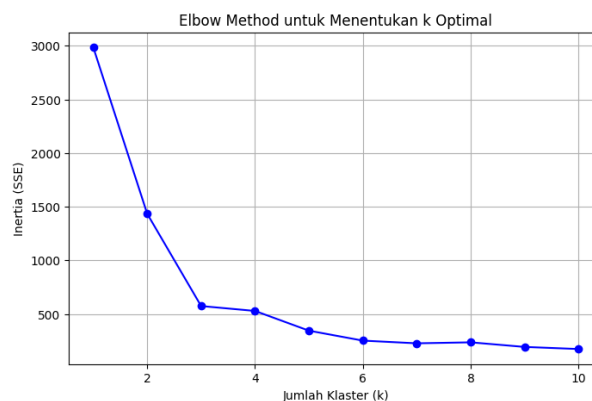
Tabel 1. Sampel Data

Tanggal	Suhu (awal)	Suhu (akhir)	Kelembapan (awal)	Kelembapan (akhir)	Kecepatan angin	Arah Angin
---------	-------------	--------------	-------------------	--------------------	-----------------	------------

2024-01-02	0.08	0.08	0.5	0.8	0.2	0.2
2024-01-03	0.09	0.07	0.4	0.9	0.1	0.4
2024-01-04	0.06	0.08	0.5	0.8	0.2	0.2
2024-01-05	0.08	0.05	0.5	1.0	0.2	0.4
2024-01-06	0.08	0.05	0.5	0.9	0.1	0.4

Pengolahan data yang digunakan pada penelitian menggunakan data cuaca harian BMKG Kediri pada tahun 2024 dan 2025. Pada tabel (1) nilai dari setiap atribut dilakukan normalisasi supaya skala dari setiap atribut berada pada satu skala yang sama. Namun pada atribut arah angin dilakukan proses yang sedikit berbeda. Jadi nilai arah angin akan dilakukan transformasi untuk nilai berubah menjadi numerik baru dilanjutkan dengan normalisasi, sama seperti dengan atribut lainnya.

Jumlah *cluster* awal biasanya diberikan secara acak atau berdasarkan perkiraan awal, namun, nilai *k* yang dipilih secara sembarangan dapat menyebabkan hasil klasterisasi yang kurang akurat atau kurang bermakna. Untuk mengoptimalkan pemilihan jumlah *cluster*, metode *elbow* menjadi salah satu opsi. Metode ini melibatkan pelaksanaan klasterisasi dengan berbagai nilai *k*, lalu menghitung *Within-Cluster Sum of Squares* (WCSS) untuk setiap *k*. WCSS mengukur total jarak kuadrat antara titik data dengan *centroid cluster* mereka masing-masing; semakin kecil nilai WCSS menunjukkan *cluster* yang lebih rapat. Berdasarkan visualisasi dengan menggunakan *elbow method* maka jumlah *cluster* yang dipilih adalah 3 *cluster*. Visualisasi dari jarak dari setiap *k* yang berbeda terdapat pada gambar (3).



Gambar 3. *Elbow Method*

Setelah menentukan banyaknya *cluster*, tahapan selanjutnya yaitu menentukan titik *centroid* dari tiap *cluster*. Pengambilan titik pusat pada mulanya dipilih seacara acak nantinya titik pusat akan berpindah secara otomatis sampai mendapati titik yang optimal. Setelah titik *centroid* telah ditentukan, masuk pada proses inti yaitu proses klasterisasi dengan *k-means clustering*.

Tabel 2. Evaluasi *Davies Bouldin Index*

Jumlah K	<i>Davies Bouldin Index</i>
2	0.828
3	0.472
4	0.766
5	0.832

6	0.819
---	-------

Setelah proses klasterisasi dilakukan, berdasarkan Tabel (2), hasil evaluasi menggunakan *davies bouldin index* (DBI) menunjukkan bahwa jumlah *cluster* yang paling optimal adalah K=3 dengan nilai DBI sebesar 0.472. Nilai ini lebih rendah dibandingkan dengan nilai DBI pada jumlah *cluster* lainnya, seperti K=2 (0.828), K=4 (0.766), K=5 (0.832), dan K=6 (0.819), yang menunjukkan bahwa pada K=3 terdapat pemisahan antar *cluster* yang paling baik dan kompak. Pada awalnya, nilai DBI untuk K=3 berada pada kisaran 0.5 sekian, yang mengindikasikan pemisahan antar *cluster* yang sudah cukup terpisah. Namun, setelah dilakukan perubahan dalam proses representasi atribut arah angin dari label kategorial seperti "Utara", "Tenggara", dan sebagainya menjadi representasi numerik biasa diubah menjadi representasi numerik berupa derajat (0° hingga 360°) terjadi peningkatan kualitas klasterisasi. Representasi arah angin sebagai angka yang bertahap memungkinkan algoritma *k-means* mengukur jarak antar data dengan lebih akurat dalam ruang fitur, yang berkontribusi terhadap penurunan nilai DBI menjadi 0.472. Artinya, klaster menjadi lebih kompak dan terpisah lebih baik. Pemilihan jumlah *cluster* tidak hanya didasarkan pada nilai DBI semata, tetapi juga mempertimbangkan hasil dari metode *elbow*, yang membantu dalam menentukan titik optimal dari penurunan *inertia*. Dengan mempertimbangkan kedua pendekatan ini, jumlah *cluster* K=3 dipilih karena memberikan keseimbangan terbaik antara kompleksitas model dan kualitas segmentasi data. Dengan demikian, pengelompokan menggunakan 3 *cluster* dinilai paling optimal dan praktis, sekaligus mencerminkan struktur alami dari data cuaca yang relatif homogen namun tetap memiliki perbedaan karakteristik yang signifikan.

III. HASIL DAN PEMBAHASAN

Tabel 3. *Centroid*

Suhu (awal)	Suhu (akhir)	Kelembapan (awal)	Kelembapan (akhir)	Kecepatan angin	Arah Angin	Karakteristik
22.9	31.9	61.2	93.0	26.4	Selatan	Panas, Kering, Tenang
23.0	32.0	67.2	95.7	26.8	Timur	Hangat, Lembap, Berangin
71.0	93.7	22.9	28.9	15.0	Barat	Dingin, Lembap, Berangin

Setelah melalui proses klasterisasi menggunakan metode *k-means* pada data cuaca wilayah Kediri, diperoleh sebanyak 3 *cluster* dengan *centroid* yang merepresentasikan kondisi rata-rata tiap *cluster*. Hasil proses klasterisasi dengan nilai-nilai yang masih dalam bentuk normalisasi, selanjutnya, akan dilakukan denormalisasi agar informasi yang diperoleh lebih mudah dianalisis. Analisis lanjutan dilakukan setelah proses denormalisasi untuk mengevaluasi karakteristik dominan dari setiap klaster berdasarkan nilai *centroid* yang ditampilkan pada tabel (3). Hasil ini memberikan gambaran mengenai pola cuaca yang terbentuk di masing-masing klaster. Klaster 1 menunjukkan kondisi panas, kering, dan tenang, dengan suhu awal sebesar 22,9°C dan suhu akhir 31,9°C, kelembapan awal 61,2% dan akhir 93,0%, serta kecepatan angin sebesar 26,4 km/jam yang berasal dari arah Selatan. Meskipun kelembapan akhir cukup tinggi,

karakteristik panas dan kering ditetapkan secara relatif berdasarkan posisi *centroid* dibandingkan klaster lainnya. Sementara itu, Klaster 2 merepresentasikan cuaca hangat, lembap, dan berangin. Hal ini tercermin dari suhu awal 23,0°C dan suhu akhir 32,0°C, kelembapan awal 67,2% dan akhir 95,7%, serta kecepatan angin mencapai 26,8 km/jam dari arah Timur. Kondisi ini mencerminkan musim penghujan yang aktif, dengan tingkat kelembapan dan pergerakan angin yang signifikan. Di sisi lain, Klaster 3 menggambarkan kondisi dingin, lembap, dan berangin, dengan suhu awal 71,0°C dan akhir 93,7°C, kelembapan awal 22,9% dan akhir 28,9%, serta kecepatan angin 15,0 km/jam dari arah Barat.

Suhu dan kelembapan pada *centroid* merupakan hasil dari proses denormalisasi setelah dilakukan standarisasi sebelumnya. Oleh karena itu, pemberian karakteristik tidak sepenuhnya didasarkan pada nilai absolut, melainkan posisi relatif antar klaster pada ruang fitur. Dengan demikian, analisis ini memungkinkan pengelompokan pola cuaca yang dominan secara lebih objektif meskipun data awal bersifat numerik dan variatif.

Tabel 4. Hasil Akhir

Wilayah	Suhu	Kelembapan	Kecepatan Angin
Kota	Panas	Kering	Tenang
Mojoroto	Panas	Kering	Tenang
Pesantren	Panas	Kering	Tenang
Badas	Panas	Kering	Tenang
Banyakan	Panas	Kering	Tenang
Gampengrejo	Panas	Kering	Tenang
Grogol	Panas	Kering	Tenang
Gurah	Panas	Kering	Tenang
Kandangan	Panas	Kering	Tenang
Kandat	Panas	Kering	Tenang
Kayenkidul	Panas	Kering	Tenang
Kepung	Panas	Kering	Tenang
Kras	Panas	Kering	Tenang
Kunjang	Panas	Kering	Tenang
Mojo	Panas	Kering	Tenang
Ngadiluwih	Panas	Kering	Tenang
Ngancar	Panas	Kering	Tenang
Ngasem	Panas	Kering	Tenang
Pagu	Panas	Kering	Tenang
Papar	Panas	Kering	Tenang
Pare	Panas	Kering	Tenang
Plemahan	Panas	Kering	Tenang
Plosoklaten	Panas	Kering	Tenang
Puncu	Panas	Kering	Tenang
Porwoasri	Panas	Kering	Tenang
Ringinrejo	Panas	Kering	Tenang
Semen	Panas	Kering	Tenang
Tarokan	Panas	Kering	Tenang
Wates	Panas	Kering	Tenang

Pada tabel (4) terlihat hasil dari pemanfaatan *k-means clustering* yang berhasil mengidentifikasi tiga *cluster* berbeda berdasarkan karakteristik cuaca di berbagai area. Total 29 area di Kediri cenderung memiliki karakter berdasar pada *cluster* 0 yang menunjukkan kondisi suhu yang panas, kelembapan yang relatif kering, dan kecepatan angin yang tenang. Distribusi

dari *cluster* sendiri menunjukkan bahwa terdapat ketidakseimbangan di antara *cluster* dengan jumlah *cluster* 0 berjumlah 6080 data, *cluster* 1 4218 data, dan *cluster* 3 3.945 data. Meskipun demikian, kehadiran dua *cluster* lain dengan jumlah data yang cukup signifikan menunjukkan bahwa terdapat variasi pola cuaca yang layak untuk dikaji lebih lanjut.

Pada tabel (4) juga menunjukkan bahwa distribusi *cluster* pada sebagian besar wilayah di Kediri mengalami cuaca yang serupa, dengan karakteristik yang konsisten. Penulis telah melakukan beberapa penyesuaian dan pengujian untuk memperbaiki distribusi klaster, seperti menghapus data *outlier* dan menerapkan metode normalisasi yang berbeda. Namun, meskipun berbagai pendekatan ini telah dicoba, dominasi klaster tetap terjadi. Analisa yang telah dilakukan oleh penulis menemukan penyebab dari distribusi *cluster* yang lebih dominan pada suatu *cluster* adalah data yang digunakan cenderung memiliki kemiripan karakter cuaca yang tinggi, yang menandakan bahwa variasi cuaca dalam data relatif homogen. Serta rentang waktu data yang tidak mencakup perubahan musim secara kontras, seperti musim hujan dan kemarau, maka variasi cuaca cenderung kecil.

Dengan demikian, analisis ini menunjukkan bahwa wilayah di Kediri memiliki variasi cuaca yang relatif terbatas, dengan dominasi pola cuaca tertentu yang tersebar secara luas di berbagai area. Kondisi ini juga diperkuat oleh nilai indeks *davies bouldin index* (DBI) yang diperoleh sebesar 0.47 menunjukkan bahwa pemisahan antar klaster cukup baik semakin kecil nilai DBI, semakin jelas pemisahan antar klaster. Namun, nilai ini juga mengindikasikan bahwa sebagian klaster memiliki kemiripan internal yang cukup tinggi, sehingga batas antar kelompok tidak terlalu ekstrem. Artinya, meskipun klasterisasi berhasil memisahkan data ke dalam tiga kelompok yang bermakna, terdapat kedekatan karakteristik di antara sebagian klaster.

IV. KESIMPULAN

Penelitian ini berhasil menerapkan metode klasterisasi *k-means* pada data cuaca wilayah Kediri dengan menghasilkan tiga *cluster* utama yang mencerminkan karakteristik cuaca berbeda secara relatif. Pemilihan jumlah *cluster* didukung oleh evaluasi menggunakan *davies bouldin index* (DBI) yang menunjukkan kualitas pemisahan *cluster* yang optimal pada $K=3$. Hasil klasterisasi memberikan gambaran pola cuaca yang dominan, mulai dari kondisi panas dan kering hingga dingin dan berangin, yang secara umum merefleksikan homogenitas variasi cuaca di wilayah tersebut. Kelebihan dari penelitian ini adalah penggunaan representasi arah angin dalam bentuk numerik derajat yang meningkatkan akurasi pengukuran jarak antar data sehingga memperbaiki kualitas klasterisasi. Selain itu, proses denormalisasi memungkinkan interpretasi hasil klaster menjadi lebih mudah dan bermakna secara aplikatif. Namun, keterbatasan utama terletak pada homogenitas data cuaca yang relatif rendah dan cakupan waktu pengambilan data yang tidak meliputi variasi musim secara ekstrem, sehingga variasi *cluster* cenderung terbatas dan distribusi data antar *cluster* tidak merata. Oleh karena itu, penting untuk melakukan analisis lebih lanjut dan mempertimbangkan penggunaan metode lain atau data yang lebih bervariasi untuk menentukan jumlah *cluster* yang lebih optimal. Penelitian ini membuka peluang untuk pengembangan lebih lanjut, seperti penggabungan *k-means* dengan algoritma lain atau penerapan pada data cuaca jangka panjang untuk mendapatkan pemahaman yang lebih mendalam tentang cuaca di wilayah Kediri.

DAFTAR PUSTAKA

- [1] E. T. Rahardjo, *CUACA DAN IKLIM*. Jakarta: PT Gramedia Widiasarana Indonesia, 1998.
- [2] N. Kurniawan, "Analisis Dampak Perubahan Iklim terhadap Sektor Pertanian," *Literacy Notes*, vol. 1, no. 2, 2023.
- [3] A. Chusyairi, "Clustering Data Cuaca Ekstrem Indonesia dengan K-Means dan Entropi," *Journal of Informatics and Communication Technology (JICT)*, vol. 5, no. 1, pp. 1-10, 2023.
- [4] F. D. Rahman, M. I. Z. Mulki, and A. Taryana, "Clustering dan klasifikasi data cuaca Cilacap dengan menggunakan metode K-Means dan Random Forest," *Jurnal SINTA: Sistem Informasi dan Teknologi Komputasi*, vol. 1, no. 2, pp. 90-97, 2024.
- [5] S. M. M. K. M. Cluster, "Pemanfaatan Data Radar Cuaca Untuk Membuat Peringatan Dini Cuaca Secara," *Newton-Maxwell Journal of Physics*, vol. 4, no. 2, Okt. 2023.
- [6] R. Muliono and Z. Sembiring, "Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen," *CESS (Journal of Computer Engineering, System and Science)*, vol. 4, no. 2, pp. 2502-2714, 2019.
- [7] M. Orisa, "Optimasi Cluster pada Algoritma K-Means," *Prosiding SENIATI*, vol. 6, no. 2, pp. 430-437, 2022.
- [8] N. A. Maori and E. Evanita, "Metode elbow dalam optimasi jumlah cluster pada K-Means clustering," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 14, no. 2, pp. 277-288, 2023.
- [9] I. T. Umagapi, B. Umaternate, H. Hazriani, and Y. Yuyun, "Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa," *Prosiding SISFOTEK*, vol. 7, no. 1, pp. 303-308, 2023.
- [10] I. Rinaldi and H. Harmayani, "Penggunaan algoritma clustering K-Means untuk mengelompokkan pola cuaca," *Journal of Science and Social Research*, vol. 8, no. 1, pp. 343-348, 2025.
- [11] M. . Ulfah, A. S. Irtawaty, D. R. Sari, Z. Zulkifli, and Y. Kurniawan, "Analisa Davies Bouldin Index (DBI) dalam pengelompokan daerah rawan longsor," *Prosiding SNITT Poltekba*, vol. 6, pp. 135-140, 2023.
- [12] D. N. Alfiansyah, V. R. S. Nastiti, and N. Hayatin, "Penerapan metode K-Means pada data penduduk miskin per kecamatan Kabupaten Blitar," *Jurnal Repositor*, vol. 4, no. 1, 2022.
- [13] A. Winarta and W. J. Kurniawan, "Optimasi cluster k-means menggunakan metode elbow pada data pengguna narkoba dengan pemrograman python," *JTIK (Jurnal Teknik Informatika Kaputama)*, vol. 5, no. 1, pp. 113-119, 2021.
- [14] M. Sholeh and K. Aeni, "Perbandingan evaluasi metode Davies Bouldin, Elbow dan Silhouette pada model clustering dengan menggunakan algoritma K-Means," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 8, no. 1, p. 56, 2023.
- [15] R. Gustrianda and D. I. Mulyana, "Penerapan data mining dalam pemilihan produk unggulan dengan metode algoritma K-Means dan K-Medoids," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 27, 2022.