

Penerapan Algoritma *K-Medoids Clustering* untuk Pengelompokan Bulan Rawan Bencana Kabut Asap di Kota Pekanbaru

M. Puja Alif Budiman¹, Doni Winarso²

¹Terapan Teknik Komputer, Pascasarjana, Politeknik Caltex Riau

²Sistem Informasi, Ilmu Komputer, Universitas Muhammadiyah Riau

¹puja22mtk@mahasiswa.pcr.ac.id *, ²doniwinarso@umri.ac.id

Abstract

Every year, Haze is a threat that often hits Pekanbaru City. The main source of Haze pollutants is Particulate Matter (PM10) which comes from forest and land fires. Despite the serious health impacts of haze, public awareness about haze-prone months in Pekanbaru City is still low. This leads to unawareness, which in turn leads to an increase in Acute Respiratory Tract Infection (ARI) cases during Haze. Therefore, this study aims to identify Haze-prone months in Pekanbaru City using the *K-Medoids Clustering* algorithm, with the hope that people will become more vigilant and maintain their health when these months arrive. The analysis was conducted based on the Air Pollution Standard Index (ISPU) and climate data of Pekanbaru City from 2015 to 2019. Of the 1,021 data available, grouping was done into 5 clusters. The results showed that clusters 3, 4, and 5 covering the months of August, September, and October, were identified as haze-prone months in Pekanbaru City. The validity of the cluster results was evaluated using the *Silhouette Coefficient* test, with a value of 0.638 which is included in the *Medium Structure* category, which indicates that the cluster data structure is well formed.

Keywords: clustering, k-medoids, ispu, silhouette coefficient, haze

Abstrak

Setiap tahun, Kabut Asap menjadi ancaman yang sering kali melanda Kota Pekanbaru. Sumber utama polutan Kabut Asap adalah *Particulate Matter* (PM₁₀) yang berasal dari kebakaran hutan dan lahan. Meskipun Kabut Asap membawa dampak serius bagi kesehatan, kesadaran masyarakat tentang bulan-bulan rawan Kabut Asap di Kota Pekanbaru masih rendah. Ini menyebabkan ketidakwaspadan yang berujung pada peningkatan kasus Infeksi Saluran Pernafasan Akut (ISPA) saat terjadi Kabut Asap. Oleh karena itu, studi ini bertujuan untuk mengidentifikasi bulan-bulan rawan Kabut Asap di Kota Pekanbaru dengan menggunakan algoritma *K-Medoids Clustering*, dengan harapan agar masyarakat menjadi lebih waspada dan menjaga kesehatan saat bulan-bulan tersebut tiba. Analisis dilakukan berdasarkan data Indeks Standar Pencemar Udara (ISPU) dan data iklim Kota Pekanbaru dari tahun 2015 hingga 2019. Dari 1.021 data yang tersedia, dilakukan pengelompokan menjadi 5 klaster. Hasil penelitian menunjukkan bahwa klaster 3, 4, dan 5 yang mencakup bulan Agustus, September, dan Oktober, diidentifikasi sebagai bulan-bulan rawan Kabut Asap di Kota Pekanbaru. Validitas hasil klaster dievaluasi menggunakan uji *Silhouette Coefficient*, dengan nilai 0,638 yang termasuk dalam kategori *Medium Structure*, yang mengindikasikan struktur data klaster terbentuk dengan baik.

Kata kunci: clustering, k-medoids, ispu, silhouette coefficient, kabut asap

©This work is licensed under a Creative Commons Attribution - ShareAlike 4.0 International License

1. Pendahuluan

Kabut asap merupakan bencana berulang yang menghantui Indonesia setiap tahun, dengan Kota Pekanbaru menjadi salah satu kota paling terdampak pada Agustus - September 2019. Penyebab utamanya adalah kebakaran hutan dan lahan yang terjadi selama musim kemarau, cuaca kering, struktur tanah gambut, dan pola angin. Dampaknya signifikan, terutama di Pekanbaru, yang mengalami tingkat pencemaran udara berbahaya menurut Indeks Standar Pencemar Udara (ISPU) selama bulan September 2019. Akibatnya, meningkatnya jumlah penderita penyakit Infeksi Saluran Pernafasan Akut (ISPA) terutama pada bayi, balita, ibu hamil, dan lansia.

Indeks Standar Pencemar Udara (ISPU) di Indonesia, sebagaimana diatur oleh Keputusan Menteri Negara Lingkungan Hidup Nomor KEP45/MENLH/1997, merupakan standar resmi untuk pemantauan kualitas udara. Data ISPU diperoleh melalui operasional

Stasiun Pemantauan Kualitas Udara Ambien Otomatis, dengan parameter melibatkan *Particulate Matter* (PM₁₀), SO₂ (Sulfur Dioksida), CO (Karbon Monoksida), O₃ (Ozon), dan NO₂ (Nitrogen Dioksida). PM₁₀, sebagai polutan utama yang terkait dengan asap kebakaran hutan, memiliki batas mutu sebesar 150 µgram/m³ berdasarkan Peraturan Pemerintah RI Nomor 41 Tahun 1999 tentang Pengendalian Pencemaran Udara.

Penelitian ini akan melibatkan analisis data Indeks Standar Pencemar Udara (ISPU) dan data iklim Kota Pekanbaru selama tahun 2015-2019. Melalui pengelompokan menggunakan algoritma *K-Medoids Clustering*, diharapkan dapat diidentifikasi bulan-bulan yang cenderung memiliki tingkat pencemaran udara tinggi dan risiko terjadinya bencana kabut asap. Hasil penelitian ini tidak hanya bermanfaat bagi masyarakat Pekanbaru untuk meningkatkan kewaspadaan, tetapi juga dapat menjadi dasar bagi pihak berwenang dalam

merancang kebijakan pencegahan dan mitigasi bencana.

Dalam penelitian yang dilakukan oleh Siti Ramadhani, dkk. tahun 2022 [1] mencoba melakukan komparasi antaran 2 metode *cluster*, yaitu *K-Medoids* dan *K-Means*. Penelitian ini mencoba mengklasifikasikan tesis mahasiswa dengan menggunakan teknik *text mining*. Hasilnya menunjukkan bahwa hasil uji *cluster K-Medoids* lebih baik dengan memperoleh nilai *Davies-Bouldin Index* (DBI) sebesar 1,56, sedangkan *K-Means* memiliki nilai DBI sebesar 2,79.

Penelitian lainnya yang dilakukan oleh Marlina, dkk. tahun 2018 [2] dan Kamila, dkk. 2019 [3] bertujuan untuk mengimplementasikan algoritma *K-Medoids* dan *K-Means* dalam melakukan pengelompokan data. Hasil penelitian tersebut membandingkan validitas *Silhouette Coefficient* dan DBI antara kedua algoritma, menunjukkan bahwa algoritma *K-Medoids* lebih baik dalam mengatasi kelemahan *K-Means* terhadap *noise* dan *outlier*. Hal ini menandakan bahwa *K-Medoids* lebih efektif dalam mengelompokkan data.

Dalam konteks ini, penerapan algoritma *K-Medoids Clustering* menjadi solusi yang menarik dan inovatif untuk memahami pola-pola temporal yang terkait dengan bencana kabut asap. *K-Medoids Clustering* dipilih karena keunggulannya dalam mengelompokkan data berdasarkan median, yang membuatnya lebih *robust* terhadap data *outlier*. Keberhasilan algoritma ini sebelumnya telah terbukti dalam konteks pengelompokan data yang kompleks, sehingga relevan untuk diterapkan dalam mengidentifikasi bulan-bulan rawan kabut asap di Kota Pekanbaru

Namun, penelitian ini akan memberikan kontribusi lebih lanjut dengan fokus pada bencana kabut asap di konteks Kota Pekanbaru. Dengan memadukan teknologi informasi dan ilmu lingkungan, harapan pada penelitian ini mampu memberikan wawasan yang lebih mendalam terhadap pola-pola terjadinya bencana kabut asap, serta mendorong langkah-langkah preventif yang lebih efektif dan responsif.

2. Metode Penelitian

Metodologi ini memberikan kerangka kerja yang jelas tentang bagaimana suatu penelitian akan dilakukan, termasuk langkah-langkah yang akan diambil, teknik yang akan digunakan, serta pendekatan analisis data yang akan diterapkan. Serangkaian langkah atau prosedur sistematis yang digunakan oleh peneliti untuk merencanakan, melaksanakan, dan menganalisis suatu penelitian dengan tujuan mencapai hasil yang valid dan dapat diandalkan ditunjukkan oleh gambar 1.



Gambar 1. Langkah-Langkah Penelitian

Kajian pustaka yang dilakukan suatu proses analisis dan sintesis informasi yang bersumber dari literatur-literatur, jurnal ilmiah, buku, artikel, atau sumber-sumber pengetahuan lainnya yang relevan dengan topik atau masalah penelitian tertentu. Tujuan utama dari kajian pustaka adalah untuk menyajikan kerangka konseptual yang mendalam, memahami perkembangan penelitian sebelumnya, mengidentifikasi kekosongan pengetahuan, dan merumuskan pertanyaan penelitian yang spesifik.

2.1. Dataset

Dataset adalah sekumpulan data yang dapat berfungsi sebagai bahan eksperimen dalam suatu penelitian [4]. *Dataset* memiliki beragam penggunaan dalam berbagai bidang dan konteks, seperti analisis data, penelitian ilmiah, pengembangan aplikasi, dan *machine learning*, dan perannya dapat berubah sesuai dengan tujuan penggunaannya.

2.2. Knowledge Discovery in Database (KDD)

KDD merupakan singkatan dari *Knowledge Discovery in Databases*, yang merujuk pada proses penggalian pengetahuan dari basis data. Proses ini mencakup tahapan seleksi, transformasi, dan interpretasi data untuk menghasilkan informasi yang berguna. Sumber sitasi untuk informasi ini tidak disediakan [5]. KDD memungkinkan untuk menemukan pengetahuan baru yang tersembunyi dalam dataset yang besar dan kompleks. Ini bisa berupa pola-pola yang tidak terduga, hubungan kausal, atau tren yang belum teridentifikasi sebelumnya.

2.3. Data Mining

Data mining merupakan suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat serta pengetahuan yang relevan dari berbagai sumber data yang terinterkoneksi [6]. *Data mining* suatu proses analisis data, terutama pada *dataset* yang besar, dengan maksud untuk mengidentifikasi hubungan yang signifikan dan menyimpulkan informasi yang sebelumnya tidak terdeteksi [7]. *Data mining* bertujuan untuk menemukan pengetahuan baru yang tidak terlihat secara langsung dalam data mentah. Ini bisa berupa pola-pola tersembunyi, hubungan kausal, atau tren yang tidak dapat dilihat dengan analisis biasa.

2.4. Clustering

Clustering adalah metode analisis data yang dipakai untuk mengelompokkan entitas yang serupa ke dalam kumpulan-kumpulan yang disebut sebagai *cluster* [8]. Fungsi *clustering* adalah untuk mengelompokkan objek-objek data yang serupa ke dalam kelompok-kelompok atau *cluster* berdasarkan kesamaan karakteristik tertentu. Tujuan utama dari *clustering* adalah untuk mengidentifikasi pola-pola yang tersembunyi dalam data, mengungkap struktur yang mungkin tidak terlihat sebelumnya, dan memahami hubungan antara objek-objek data.

2.5. K-Medoids

K-Medoids merupakan teknik pengelompokan data yang membagi *dataset* menjadi kelompok-kelompok, dan mampu mengatasi kelemahan *K-Means* yang rentan terhadap data ekstrem [9]. Seperti *clustering* lainnya, tujuan utama dari *K-Medoids* adalah untuk mengelompokkan objek-objek data yang serupa ke dalam *cluster-cluster* yang berbeda.

Berikut ini langkah-langkah dalam algoritma *K-Medoids* :

1. Buatlah pusat-pusat *cluster* dalam jumlah *k* yang telah ditentukan.

2. Tetapkan setiap data ke dalam kluster yang paling dekat menggunakan persamaan untuk mengukur jarak *Euclidean Distance* dengan menggunakan Persamaan 2.1. berikut :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad ; 1, 2, 3, \dots, n \quad (1)$$

3. Pilihlah secara acak satu objek dari setiap *cluster* sebagai kandidat *medoids* yang baru.

4. Hitunglah jarak antara setiap objek yang tergabung dalam tiap *cluster* dengan *medoids* kandidat yang baru.

5. Simpangan dihitung dengan memperhitungkan perbedaan antara nilai total jarak baru dengan nilai total jarak sebelumnya. Jika nilai simpangan (*S*) < 0, langkah selanjutnya adalah menukar objek dengan data kluster untuk membentuk sekelompok objek baru sebanyak *k* objek, yang akan bertindak sebagai *medoids*. Namun, jika nilai simpangan (*S*) > 0, algoritma akan dihentikan.

6. Lakukan kembali langkah-langkah 3 hingga 5 secara berulang hingga tidak ada perubahan pada *medoids*, sehingga kita dapatkan *cluster* dan anggota-anggota dari setiap *cluster*.

K-Medoids cenderung lebih stabil daripada beberapa teknik *clustering* lainnya, karena *medoids* tidak bergantung pada *mean* atau rata-rata, yang dapat dipengaruhi oleh perubahan kecil dalam data.

2.6. Min-Max Normalization

Normalisasi memudahkan untuk membandingkan variabel-variabel dengan unit yang berbeda atau rentang nilai yang berbeda. Ini dapat mempermudah

interpretasi hasil analisis data atau model yang dihasilkan. Metode normalisasi *Min-Max* adalah teknik yang mengubah rentang nilai data agar berada di antara 0 dan 1 [10]. *Min-Max* normalisasi mempertahankan informasi relatif dari setiap nilai dalam *dataset*. Artinya, urutan relatif nilai-nilai tersebut tetap sama setelah normalisasi, tetapi rentang nilainya disesuaikan menjadi antara 0 dan 1.

$$v^i = \frac{v - \min_A}{\max_A - \min_A} \quad (2)$$

2.7. Silhouette Coefficient

Silhouette Coefficient merupakan alat evaluasi yang mengukur keberhasilan pembentukan *cluster* dalam analisis kluster dengan menilai kedekatan antar titik data dalam satu *cluster* dibandingkan dengan *cluster* lain, di mana nilai yang lebih tinggi menandakan pemisahan yang lebih baik dan konsistensi internal yang lebih tinggi [11].

Tabel 1. Tabel *Silhouette Coefficient*

Indeks	Kategori
0,71-1,00	<i>Strong Structure</i>
0,51-0,70	<i>Medium Structure</i>
0,26-0,50	<i>Weak Structure</i>
≤ 0,25	<i>No Structure</i>

Silhouette Coefficient metode evaluasi yang mengukur seberapa baik titik-titik dalam *cluster* sesuai dan terpisah dari *cluster* lainnya, dengan nilai positif mendekati 1 menunjukkan tingkat kesesuaian dan pemisahan yang baik [12].

2.8. Indeks Standar Pencemar Udara (ISPU)

Menurut Keputusan Menteri Negara Lingkungan Hidup Nomor KEP45/MENLH/1997 yang membahas tentang Indeks Standar Pencemar Udara (ISPU), ISPU adalah suatu angka tanpa satuan yang mencerminkan kondisi kualitas udara di suatu lokasi dan waktu tertentu, dengan mempertimbangkan dampaknya terhadap kesehatan manusia, nilai estetika, dan kehidupan lainnya. Parameter yang menjadi dasar dalam ISPU salah satunya adalah PM_{10} (*Particulate Matter*).

Tabel 2. Tabel ISPU PM_{10}

Indeks	Kategori	Warna
1-50	Baik	Hijau
51-100	Sedang	Biru
101-199	Tidak Sehat	Kuning
200-299	Sangat Tidak Sehat	Merah
≥ 300	Berbahaya	Hitam

3. Hasil dan Pembahasan

Penelitian ini mengambil sumber data yang dikumpulkan dengan mengakses data publik yang tersedia secara *online* melalui situs web <http://iku.menlhk.go.id/aqms/arsip> yang dikelola oleh Kementerian Lingkungan Hidup dan Kehutanan RI untuk memperoleh arsip data pemantauan Indeks

Standar Pencemar Udara (ISPU) Kota Pekanbaru dengan total 2.113 data yang dikumpulkan dari rentang tahun 2015 hingga 2019.

3.1. Praproses Data

Dataset mengenai Indeks Standar Pencemar Udara (ISPU) Kota Pekanbaru mencakup lima variabel yang digunakan untuk menilai tingkat kualitas udara, termasuk PM₁₀ (*Particulate Matter*), SO₂ (Sulfur Dioksida), CO (Karbon Monoksida), O₃ (Ozon), serta NO₂ (Nitrogen Dioksida).

Tabel 3. Tabel *Dataset* ISPU

Tanggal	PM ₁₀	SO ₂	CO	O ₃	NO ₂
19/08/2015	149	0	0	0	0
22/08/2015	36	0	0	0	0
28/08/2015	91	0	0	0	0
04/09/2015	512	0	0	0	0
06/09/2015	81	0	0	0	0
07/09/2015	790	0	0	0	0
07/09/2015	174	0	0	0	0
07/09/2015	782	0	0	0	0
07/09/2015	790	0	0	0	0
07/09/2015	241	0	0	0	0
...
30/09/2019	78	0	0	0	0
03/10/2019	102	0	0	0	0
04/10/2019	65	0	0	0	0

Diantara 5 parameter yang tersedia, maka dipilihlah parameter PM₁₀ (*Particulate Matter*) sebagai *input*.

Setelah melakukan seleksi data awal, langkah selanjutnya adalah membersihkan data PM₁₀ (*Particulate Matter*) yang mengandung *missing value*. Dari total 2.113 entri data, sebanyak 1.092 di antaranya memiliki nilai yang hilang. Selain itu, terdapat redundansi data di mana terdapat beberapa entri pada tanggal yang sama namun dengan waktu yang berbeda.

Tabel 4. Tabel *Dataset* Sebelum Normalisasi

Tanggal	PM ₁₀
19/08/2015	149
22/08/2015	36
28/08/2015	91
04/09/2015	512
06/09/2015	81
07/09/2015	790
07/09/2015	174
07/09/2015	782
07/09/2015	790
07/09/2015	241
...	...
30/09/2019	78
03/10/2019	102
04/10/2019	65

Untuk membentuk data yang konsisten per tanggal, maka dipilih satu entri data pada setiap tanggal yang mengalami redundansi data. Dengan demikian, setelah proses pembersihan data, tersisa 1.021 entri data yang akan digunakan sebagai *dataset*.

Setelah tahap pembersihan data selesai, langkah berikutnya adalah normalisasi data pada *dataset* menggunakan *Min-Max Normalization*, yang bertujuan untuk mengurangi risiko *overflows* dan *underflows* dengan menyesuaikan skala data ke dalam rentang 0-1.

Tabel 5. Tabel *Dataset* Setelah Normalisasi

Tanggal	Normalisasi PM ₁₀
19/08/2015	0,141
22/08/2015	0,033
28/08/2015	0,085
04/09/2015	0,485
06/09/2015	0,076
07/09/2015	0,749
08/09/2015	0,636
09/09/2015	0,104
10/09/2015	0,152
11/09/2015	0,716
...	...
30/09/2019	0,073
03/10/2019	0,096
04/10/2019	0,061

3.2. Pengklasteran Data

Setelah melalui normalisasi data, langkah berikutnya adalah menerapkan proses *data mining* dengan memanfaatkan teknik pengelompokan *K-Medoids Clustering* sebanyak 5 *cluster*. Makna dari 5 *cluster* yang akan dibentuk dapat dilihat pada tabel 6.

Tabel 6. Tabel *Cluster*

Cluster	Indikator PM ₁₀
Cluster 1	Baik
Cluster 2	Sedang
Cluster 3	Tidak Sehat
Cluster 4	Sangat Tidak Sehat
Cluster 5	Berbahaya

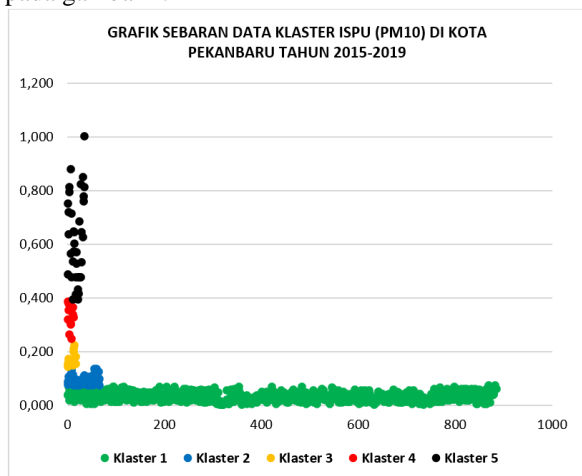
Setelah menentukan makna *cluster*, maka proses *clustering* dilakukan agar setiap data menempati klasternya masing-masing. Adapun hasil pengklasteran data yang didapatkan dapat dilihat pada tabel 7.

Tabel 7. Tabel Hasil *Clustering*

Tanggal	PM ₁₀	Normalisasi PM ₁₀	Cluster
19/08/2015	149	0,141	3
22/08/2015	36	0,033	1
28/08/2015	91	0,085	2
04/09/2015	512	0,485	5
06/09/2015	81	0,076	2
07/09/2015	790	0,749	5
08/09/2015	671	0,636	5
09/09/2015	110	0,104	2
10/09/2015	161	0,152	3
11/09/2015	755	0,716	5
...
30/09/2019	78	0,073	2
03/10/2019	102	0,096	2
04/10/2019	65	0,061	1

Setelah melakukan *clustering*, maka data hasil klaster dibentuk dalam grafik *scatter* untuk melihat pola

sebaran data setelah dilakukan proses pengklasteran. Untuk melihat grafik sebaran data *cluster* dapat dilihat pada gambar 2.



Gambar 2. Sebaran Data Hasil *Cluster*

Adapun jumlah anggota atau data yang tergabung pada tiap *cluster* dapat dilihat pada tabel 8.

Tabel 8. Tabel Distribusi Hasil *Clustering*

Cluster	Jumlah
1	885
2	68
3	18
4	13
5	37

Pada tabel 8, proses *clustering* menghasilkan dengan *cluster* 1 memiliki 885 data, *cluster* 2 memiliki 68 data, *cluster* 3 memiliki 18 data, *cluster* 4 memiliki 13 data, dan *cluster* 5 memiliki 37 data. Untuk melihat *medoids* atau titik pusat *cluster*, dapat dilihat pada tabel 9.

Tabel 9. Tabel *Medoids* Hasil *Clustering*

PM ₁₀	Normaliasi PM ₁₀	Cluster
50	0,047	Cluster 1
100	0,094	Cluster 2
189	0,179	Cluster 3
314	0,297	Cluster 4
500	0,474	Cluster 5

Berdasarkan informasi dalam tabel 9, dapat disimpulkan bahwa nilai *medoids* akhir untuk masing-masing klaster adalah sebagai berikut: klaster 1 memiliki *medoids* 0,047, klaster 2 memiliki *medoids* 0,094, klaster 3 memiliki *medoids* 0,179, klaster 4 memiliki *medoids* 0,297, dan klaster 5 memiliki *medoids* 0,474.

Tabel 10. Tabel Distribusi Hasil *Clustering*

Cluster	Jumlah Data	Bulan Tahun (Jumlah Hari)
1 Baik	885	Agustus 2015 (1), Oktober 2015 (1), November 2015 (19), Desember 2015 (4), Januari 2016 (3), Februari 2016 (11), Maret 2016 (12), April 2016 (14), Mei 2016 (5), Juni 2016 (9), Juli 2016 (1), Agustus 2016 (6), September 2016 (26), Oktober 2016

		(18), November 2016 (20), Desember 2016 (24), Januari 2017 (29), Februari 2017 (24), Maret 2017 (20), April 2017 (22), Mei 2017 (26), Juni 2017 (27), Juli 2017 (26), Agustus 2017 (28), September 2017 (29), Oktober 2017 (28), November 2017 (27), Desember 2017 (26), Januari 2018 (26), Februari 2018 (25), Maret 2018 (23), April 2018 (27), Mei 2018 (26), Juni 2018 (25), Juli 2018 (25), Agustus 2018 (29), September 2018 (25), Oktober 2018 (26), November 2018 (16), Desember 2018 (22), Januari 2019 (8), Februari 2019 (4), Maret 2019 (14), April 2019 (25), Mei 2019 (22), Juni 2019 (14), Juli 2019 (10), Agustus 2019 (3), September 2019 (3), Oktober 2019 (1)
2 Sedang	68	Agustus 2015 (1), September 2015 (6) Oktober 2015 (6), November 2015 (1), Agustus 2016 (3), Februari 2017 (1), Maret 2017 (2), Juli 2017 (1), Oktober 2017 (2), September 2018 (1), November 2018 (1), Desember 2018 (1), Februari 2019 (1), Maret 2019 (14), Juni 2019 (2), Juli 2019 (9), Agustus 2019 (11), September 2019 (4), Oktober 2019 (1)
3 Tidak Sehat	18	Agustus 2015 (1), September 2015 (2) Oktober 2015 (4), Agustus 2019 (7), September 2019 (4)
4 Sangat Tidak Sehat	13	September 2015 (2) Oktober 2015 (6), Agustus 2016 (1), September 2019 (4)
5 Berbahaya	37	September 2015 (2) Oktober 2015 (6), Agustus 2016 (1), September 2019 (4)

Setelah proses *clustering* dilakukan, dengan mengacu pada tabel 10, bulan-bulan rawan bencana kabut asap dianalisis menggunakan Indeks Standar Pencemar Udara (ISPU) dengan kategori Tidak Sehat, Sangat Tidak Sehat, dan Berbahaya, sesuai dengan Peraturan Pemerintah (PP) RI Nomor 41 Tahun 1999 Tentang Pengendalian Pencemaran Udara, di mana baku mutu untuk polusi Particulate Matter (PM₁₀) adalah 150 µgram/m³, yang dalam Tabel 2 ditempatkan dalam kategori Tidak Sehat.

Cluster 3 mencakup bulan-bulan di mana terjadi 4 kejadian bencana kabut asap, termasuk pada bulan Agustus 2015, September 2015, Oktober 2015, dan September 2019, dengan total 18 hari di antaranya terjadi dengan kategori ISPU "Tidak Sehat".

Sementara itu, *cluster* 4 mencakup 4 bulan dengan 13 hari bencana kabut asap, termasuk pada bulan September 2015, Oktober 2015, Agustus 2016, dan September 2019, dengan mayoritas hari terjadi dengan kategori ISPU "Sangat Tidak Sehat".

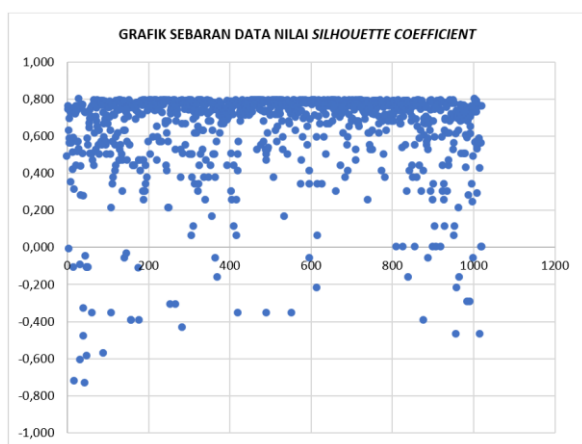
Cluster 5 juga mengalami kejadian bencana kabut asap pada 3 bulan, yaitu September 2015, Oktober 2015, dan September 2019, dengan total 37 hari, mayoritas di antaranya terjadi dengan kategori ISPU "Berbahaya".

3.3. Evaluasi Model

Setelah data telah di-cluster dalam tabel 5, langkah berikutnya adalah melakukan analisis menggunakan metode *Silhouette Coefficient* untuk mengevaluasi seberapa baik objek ditempatkan dalam sebuah cluster.

Tabel 11. Tabel Hasil Uji *Silhouette Coefficient*

Tanggal	Normalisasi PM ₁₀	Cluster	s(i)	Kategori
19/08/2015	0,141	3	0,490	Weak Structure
22/08/2015	0,033	1	0,762	Strong Structure
28/08/2015	0,085	2	0,748	Strong Structure
04/09/2015	0,485	5	-0,008	No Structure
06/09/2015	0,076	2	0,634	Medium Structure
07/09/2015	0,749	5	0,586	Medium Structure
08/09/2015	0,636	5	0,562	Medium Structure
09/09/2015	0,104	2	0,698	Medium Structure
10/09/2015	0,152	3	0,724	Strong Structure
11/09/2015	0,716	5	0,350	Weak Structure
...
30/09/2019	0,073	2	0,698	Medium Structure
03/10/2019	0,096	2	0,724	Strong Structure
04/10/2019	0,061	1	0,350	No Structure
Rata-rata <i>Silhouette Coefficient</i>			0,638	Medium Structure



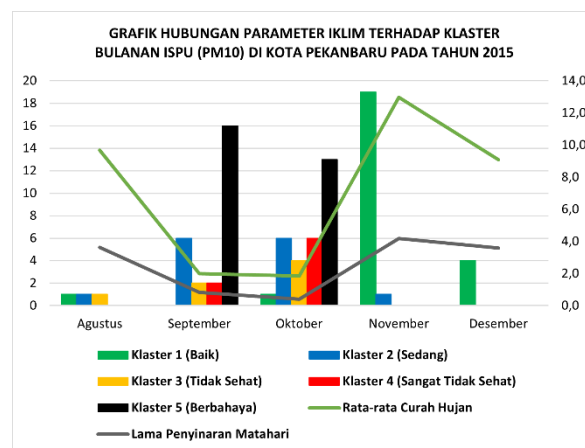
Gambar 3. Sebaran Data Hasil Uji *Silhouette Coefficient*

Setelah melakukan evaluasi menggunakan *Silhouette Coefficient* pada seluruh data, hasilnya menunjukkan bahwa dari total 1.021 data, didapatkan 645 data berada pada kategori *Strong Structure*, 215 data kategori *Medium Structure*, 95 data kategori *Weak Structure*, dan 66 data kategori *No Structure*. Jika nilai *Silhouette Coefficient* dari seluruh data dihitung rata-rata, diperoleh nilai sebesar 0,638 yang mengindikasikan kategori struktur cluster yang sedang. Dengan

demikian, kesimpulan dapat diambil bahwa seluruh data yang dikelompokkan memiliki kualitas cluster yang sebagian besar termasuk dalam kategori struktur cluster yang *Medium Structure*.

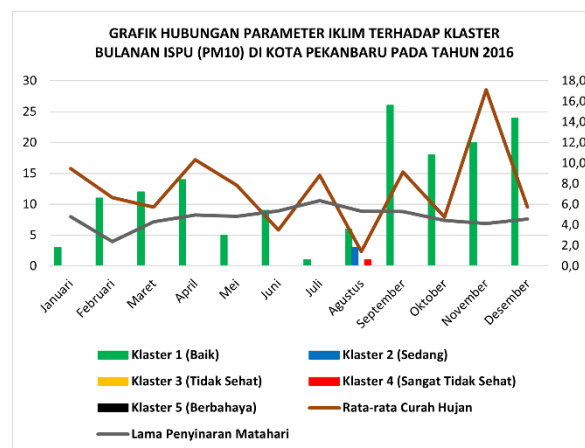
3.4. Analisa Clustering

Bencana kabut asap sangat terkait dengan kondisi cuaca. Oleh karena itu, berdasarkan hasil cluster kejadian bencana kabut asap yang telah dilakukan, penting untuk mempertimbangkan faktor-faktor iklim.



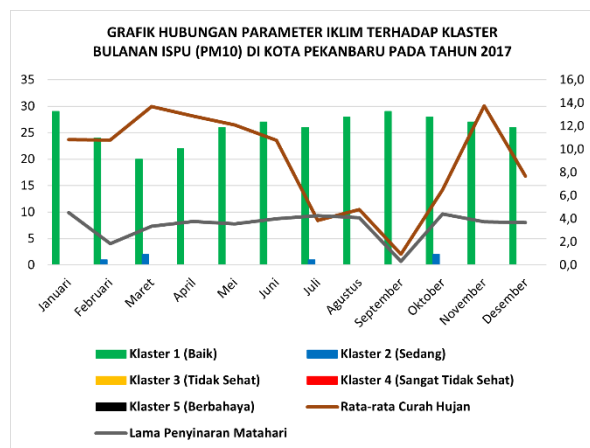
Gambar 4. Hubungan Cluster Dengan Iklim Tahun 2015

Pada gambar 4, terlihat bahwa tingkat polusi udara di Kota Pekanbaru pada tahun 2015 mulai menurun menjadi kategori "Tidak Sehat" pada bulan Agustus, mencapai puncaknya pada bulan September dan Oktober dengan dominan dalam kategori "Berbahaya". Kualitas udara mulai membaik kembali pada bulan November. Kurva rata-rata Curah Hujan bulanan menunjukkan penurunan dari bulan Agustus, mencapai titik terendah pada bulan September dan Oktober dengan 1,8-2,0 mm/bulan, kemudian meningkat kembali pada bulan November menjadi 13,0 mm/bulan, yang berkontribusi pada perbaikan kualitas udara. Pola yang serupa terlihat pada kurva Lama Penyinaran Matahari, di mana pada puncak kabut asap di bulan September dan Oktober, jumlah jam penyinaran matahari menurun menjadi 0,4/jam per hari.



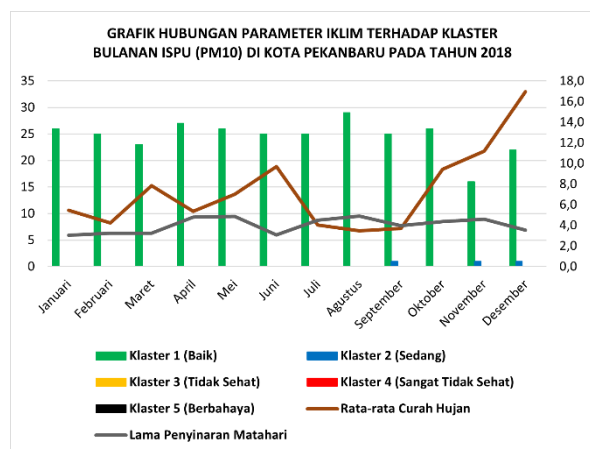
Gambar 5. Hubungan Cluster Dengan Iklim Tahun 2016

Gambar 5, terlihat bahwa secara umum kualitas udara di Kota Pekanbaru pada tahun 2016 berada dalam kategori "Baik". Namun, terdapat satu hari pada bulan Agustus di mana kualitas udara mencapai kategori "Sangat Tidak Sehat", yang disebabkan oleh penurunan Curah Hujan (RR) menjadi 1,4 mm/bulan. Kualitas udara kembali membaik menjadi kategori "Baik" pada bulan September, karena Curah Hujan (RR) meningkat kembali menjadi 9,1 mm/bulan. Perubahan pada kurva Lama Penyinaran Matahari (ss) tidak signifikan karena kejadian kabut asap hanya terjadi satu hari pada bulan Agustus saja.



Gambar 6. Hubungan Cluster Dengan Iklim Tahun 2017

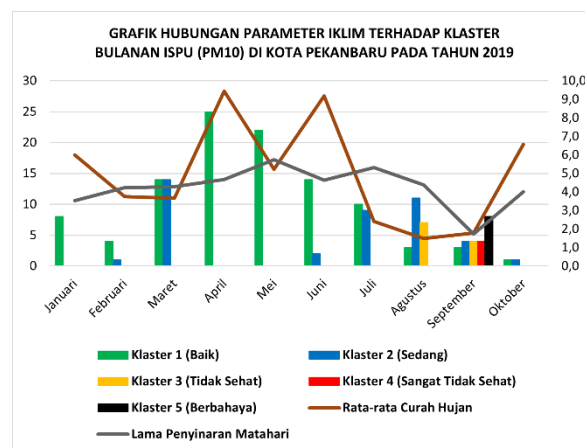
Pada gambar 6, terlihat bahwa secara keseluruhan kualitas udara di Kota Pekanbaru pada tahun 2017 umumnya berada dalam kategori "Baik", meskipun ada beberapa hari di bulan-bulan tertentu di mana kualitas udara menurun menjadi kategori "Sedang". Pada grafik parameter iklim, terlihat bahwa pada bulan September terjadi penurunan signifikan dalam Curah Hujan (RR) hingga mencapai 0,9 mm/bulan. Hal serupa terjadi pada grafik Lamanya Penyinaran Matahari (ss), yang mengalami penurunan signifikan hingga mencapai 0,3/jam per harinya. Meskipun terjadi penurunan yang signifikan pada bulan September, tidak terjadi penurunan kualitas udara atau kejadian kabut asap.



Gambar 7. Hubungan Cluster Dengan Iklim Tahun 2018

Dari gambar 7, terlihat bahwa sebagian besar waktu pada tahun 2018, kualitas udara di Kota Pekanbaru berada dalam kategori "Baik", meskipun ada beberapa

hari di bulan-bulan tertentu di mana kualitas udara turun menjadi "Sedang". Namun, ada penurunan dalam Curah Hujan (RR) pada bulan Juli, Agustus, dan September, meskipun tidak berdampak pada penurunan kualitas udara.



Gambar 8. Hubungan Cluster Dengan Iklim Tahun 2019

Gambar 8 menunjukkan bahwa sebagian besar waktu pada tahun 2019, kualitas udara di Kota Pekanbaru umumnya berada dalam kategori "Baik", namun terdapat beberapa hari di bulan-bulan tertentu di mana terjadi penurunan kualitas udara ke tingkat "Sedang", seperti yang terjadi pada bulan Maret dan Juli. Pada bulan Agustus, terjadi penurunan kualitas udara ke tingkat "Sedang" dan "Tidak Sehat", dan puncak penurunan tersebut terjadi pada bulan September dengan mencapai kategori "Berbahaya". Namun, pada bulan Oktober, kualitas udara mulai membaik kembali.

Pada 2019 grafik rata-rata Curah Hujan (RR) bulanan menunjukkan penurunan dari bulan Juli hingga puncak kabut asap pada bulan September, dengan Curah Hujan (RR) turun mencapai 1,5-1,7 mm/bulan. Namun, pada bulan Oktober, curah hujan mulai meningkat menjadi 6,6 mm/bulan, yang menyebabkan perbaikan kualitas udara. Pola yang serupa juga terlihat pada grafik Lama Penyinaran Matahari (ss), dimulai dengan penurunan pada bulan Agustus hingga mencapai puncaknya pada bulan September, di mana Lama Penyinaran Matahari (ss) mengalami penurunan hingga 1,7/jam per harinya.

Berdasarkan data grafik dan penjelasan yang telah diuraikan, dapat disimpulkan bahwa bulan-bulan yang memiliki tingkat kerawanan tinggi untuk terjadinya bencana kabut asap di Kota Pekanbaru adalah Agustus, September, dan Oktober. Selama tiga bulan tersebut, kabut asap umumnya mencapai kategori "Tidak Sehat" hingga bahkan "Berbahaya". Penurunan Curah Hujan (RR) yang terjadi pada periode tersebut, disebabkan oleh cuaca ekstrem saat transisi dari akhir musim kemarau ke awal musim hujan, menjadi faktor utama yang mempengaruhi kejadian kabut asap. Penurunan ini memicu peningkatan kebakaran hutan dan lahan karena kekeringan, terutama di lahan gambut yang rentan terbakar dengan mudah.

4. Kesimpulan

Berdasarkan diskusi dalam Bab 4, dapat disimpulkan bahwa:

1. Algoritma *K-Medoids Clustering* berhasil diterapkan secara efektif untuk mengelompokkan bulan-bulan rawan bencana kabut asap di Kota Pekanbaru. Dari 1021 data yang dikelompokkan menggunakan algoritma ini, terdapat 5 kelompok, yaitu: *cluster 1* (baik) dengan 885 data, *cluster 2* (sedang) dengan 68 data, *cluster 3* (tidak sehat) dengan 18 data, *cluster 4* (sangat tidak sehat) dengan 13 data, dan *cluster 5* (berbahaya) dengan 37 data. Uji *Silhouette Coefficient* menghasilkan nilai 0,638 dengan kategori *Medium Structure*, menandakan bahwa semua data yang dikelompokkan memiliki kualitas kelompok Struktur Medium.

2. Bulan-bulan yang termasuk dalam kategori rawan bencana kabut asap adalah Agustus, September, dan Oktober. Umumnya, pada ketiga bulan tersebut, kejadian bencana kabut asap dapat mencapai kategori "Tidak Sehat" hingga "Berbahaya".

Daftar Rujukan

- [1] S. Ramadhani, D. Azzahra, and Z. Tomi, "Comparison of K-Means and K-Medoids Algorithms in Text Mining based on Davies Bouldin Index Testing for Classification of Student's Thesis," *Digital Zone*, vol. 13, no. 1, pp. 24–33, May 2022. doi: 10.31849/digitalzone.v13i1.9292.
- [2] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, vol. 4, no. 2, pp. 64, Dec. 2018. doi: 10.24014/coreit.v4i2.4498.
- [3] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119, Feb. 2019. doi: 10.24014/rmsi.v5i1.7381.
- [4] Y. Yuliska and K. U. Syaliman, "Literatur Review Terhadap Metode, Aplikasi dan Dataset Peringkasan Dokumen Teks Otomatis untuk Teks Berbahasa Indonesia," *IT Journal Research and Development*, vol. 5, no. 1, pp. 19–31, Jul. 2020. doi: 10.25299/itjrd.2020.vol5(1).4688.
- [5] D. S. O. Panggabean, E. Buulolo, and N. Silalahi, "Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda," *JURIKOM (Jurnal Riset Komputer)*, vol. 7, no. 1, pp. 56, Feb. 2020. doi: 10.30865/jurikom.v7i1.1947.
- [6] Novi Karolina Surbakti, "Data Mining Pengelompokan Pasien Rawat Inap Peserta BPJS Menggunakan Metode Clustering (Studi Kasus: RSU.Bangkalan)," *Journal of Information and Technology*, vol. 1, no. 2, pp. 47–53, Aug. 2021. doi: 10.32938/jitu.v1i2.1470.
- [7] T. D. Prastiwi, J. Pricilia, dan E. Rasywir, "Prediksi Tingkat Kelulusan Menggunakan K-Means Pada Program Studi Informatika Unismuh Makassar," *Jurnal Fasilitas*, vol. 13, no. 3, pp. 6061, Dec. 2023. doi: 10.37859/jf.v13i3.6061.
- [8] H. Prastiwi, J. Pricilia, and E. Rasywir, "Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering," *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)*, vol. 2, no. 1, pp. 141–148, Apr. 2022. doi: 10.33998/jakakom.2022.2.1.34.
- [9] A. D. Andini and T. Arifin, "IMPLEMENTASI ALGORITMA K-MEDOIDS UNTUK KLASIFIKASI DATA PENYAKIT PASIEN DI RSUD KOTA BANDUNG," *Jurnal Responsif: Riset Sains dan Informatika*, vol. 2, no. 2, pp. 128–138, Aug. 2020. doi: 10.51977/jti.v2i2.247.
- [10] I. Permana and F. N. Salisah, "Pengaruh Normalisasi Data Terhadap Performa Hasil Klasifikasi Algoritma Backpropagation: The Effect of Data Normalization on the Performance of the Classification Results of the Backpropagation Algorithm," *Indonesian Journal of Informatic Research and Software Engineering (IJIRSE)*, vol. 2, no. 1, pp. 67–72, Mar. 2022. doi: 10.57152/ijirse.v2i1.311.
- [11] K. P. Simanjuntak and U. Khaira, "Pengelompokan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 1, no. 1, pp. 7–16, Mar. 2021. doi: 10.57152/malcom.v1i1.6.
- [12] Nurmilayanti and N. Hartono, "Implementasi Metode Clustering Sebagai Penunjang Strategi dalam Manajemen Pelanggan," *ejurnal.umri.ac.id*, Dec. 2023, doi: 10.37859/jf.v13i3.5617.