



Implementasi Algoritma K-Means: Analisis Klasterisasi Emisi Metana Global Berdasarkan Sektor Dan Negara Menggunakan RapidMiner

**Muhammad Faqih Alharits^{1*}, Ramdoni², Hanif Maulana Ar Rasyid³, Khalid Saefullah⁴,
Arjuna Mardika Maulana⁵**

^{1,2,3,4,5}Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email : ^{1*}alharitsfaqih@gmail.com, ²doony05122022@gmail.com, ³hanifmaulanaarasyid@gmail.com,
⁴khalidsaefullah1622@gmail.com, ⁵arjunamardika03@gmail.com

(* : coresponding author)

Abstrak – Penelitian ini bertujuan untuk menganalisis pola emisi metana global berdasarkan sektor dan negara menggunakan algoritma *K-Means clustering*. Data yang digunakan adalah dataset emisi metana global dari *Kaggle*, yang mencakup informasi tentang emisi, sektor, sub-sektor, dan negara. Melalui penerapan algoritma *K-Means*, negara-negara dikelompokkan berdasarkan kesamaan pola emisi metana mereka. Hasil analisis menunjukkan kelompok negara dengan karakteristik emisi yang berbeda, yang memungkinkan identifikasi prioritas tindakan mitigasi. Penelitian ini memberikan wawasan penting bagi pembuat kebijakan dalam upaya pengendalian emisi metana global.

Kata Kunci: Emisi Metana; Klasterisasi K-Means; Klasterisasi Emisi; Analisis Data; RapidMiner

Abstract–This study aims to analyse global methane emission patterns by sector and country using *K-Means clustering* algorithm. The data used is the global methane emission dataset from *Kaggle*, which includes information on emissions, sectors, sub-sectors, and countries. Through the application of the *K-Means* algorithm, countries were clustered based on the similarity of their methane emission patterns. The analysis results show clusters of countries with different emission characteristics, which allows the identification of prioritised mitigation actions. This research provides important insights for policy makers in the effort to control global methane emissions.

Keywords: Methane Emissions, K-Means Clustering, Emissions Clustering, Data Analysis, RapidMiner

1. PENDAHULUAN

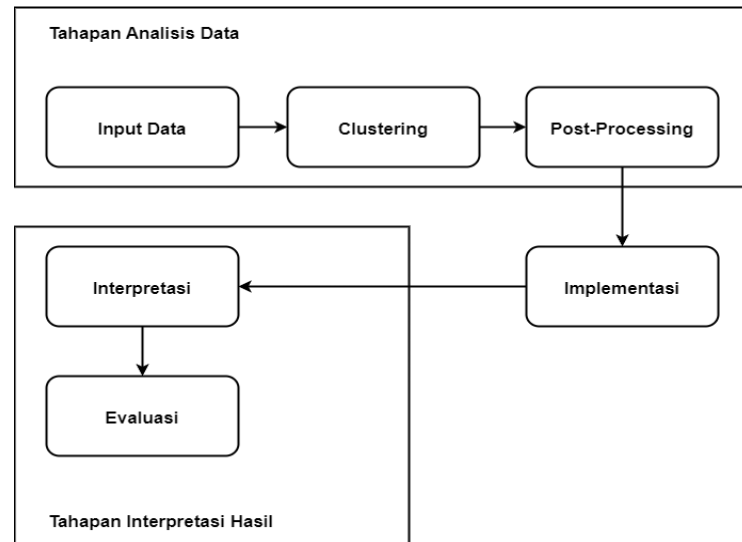
Metana merupakan salah satu gas rumah kaca yang memiliki potensi pemanasan global jauh lebih besar dibandingkan karbon dioksida dalam jangka pendek. Berdasarkan data internasional, metana bertanggung jawab atas sekitar 30% peningkatan suhu global sejak Revolusi Industri. Sumber emisi metana mencakup berbagai sektor, seperti energi (minyak, gas alam, batubara, dan bioenergi), pertanian (produksi ternak dan budidaya padi), serta limbah (pengelolaan sampah dan air limbah). Meskipun beragam sektor telah diidentifikasi, distribusi dan intensitas emisi metana sering kali bervariasi antarnegara dan wilayah, menjadikan upaya mitigasi global sebagai tantangan yang kompleks.

Analisis pola emisi metana global menjadi langkah penting dalam memahami karakteristik sumber emisi dan prioritas tindakan mitigasi. Salah satu pendekatan yang dapat dilakukan adalah klasterisasi, di mana data dari berbagai negara dan sektor dianalisis untuk menemukan pola atau kelompok dengan karakteristik serupa. Algoritma *K-Means* merupakan metode klasterisasi yang sering digunakan karena efisiensinya dan kemampuannya untuk mengelompokkan data berdasarkan kemiripan atribut.

Penelitian ini bertujuan untuk menganalisis pola emisi metana global menggunakan algoritma *K-Means clustering*. Dengan menggunakan dataset dari *Kaggle* yang mencakup informasi tentang emisi metana, sektor, sub-sektor, dan negara, penelitian ini berupaya untuk mengidentifikasi kelompok negara berdasarkan kesamaan pola emisi mereka. Hasil penelitian diharapkan dapat memberikan wawasan yang relevan bagi pembuat kebijakan untuk merumuskan

strategi mitigasi yang lebih efektif dan terfokus, serta memberikan kontribusi terhadap literatur ilmiah di bidang data mining dan perubahan iklim.

2. METODE



Gambar 1. Tahapan Penelitian

Penelitian ini dilakukan dengan pendekatan klusterisasi untuk menganalisis pola emisi metana global berdasarkan sektor dan negara. Tahapan penelitian dirancang secara sistematis dan dijelaskan melalui diagram alur yang ditunjukkan pada Gambar 1. Diagram tersebut mencakup langkah-langkah utama, yaitu pengumpulan data, klusterisasi menggunakan algoritma *K-Means*, *post-processing*, serta interpretasi dan evaluasi hasil.

Sebagai alat bantu utama, perangkat lunak *RapidMiner* digunakan dalam seluruh proses ini, mengingat kemampuannya untuk mengintegrasikan berbagai tahapan analisis data secara efisien, mulai dari input data hingga visualisasi hasil klusterisasi. Proses ini dirancang untuk memastikan bahwa hasil penelitian memberikan wawasan yang mendalam mengenai pola emisi metana, sekaligus mendukung pengambilan keputusan yang berbasis data. Penelitian ini juga memanfaatkan fitur *RapidMiner* untuk melakukan standarisasi data, menentukan jumlah kluster optimal, serta melakukan analisis evaluasi melalui metrik seperti *silhouette score*.

Melalui metodologi ini, penelitian tidak hanya bertujuan untuk mengidentifikasi kluster negara berdasarkan pola emisi metana mereka, tetapi juga memberikan rekomendasi berbasis data yang relevan bagi pembuat kebijakan. Pendekatan yang digunakan menghubungkan setiap tahapan penelitian secara terpadu untuk memastikan hasil yang komprehensif dan akurat.

2.1. Sampel Perhitungan

Sebelum kita masuk kedalam sebuah tahap utama yaitu tahapan analisis data, akan dijelaskan terlebih dahulu mengenai sampel perhitungan yang digunakan untuk menggambarkan bagaimana proses analisis data dilakukan. Pada tahap ini, kita akan menunjukkan contoh langkah-langkah perhitungan yang diterapkan pada data yang akan dianalisis, mulai dari input data, penerapan algoritma *clustering*, hingga tahapan *post-processing*. Sampel perhitungan ini juga bertujuan untuk memberikan pemahaman yang lebih jelas mengenai alur dan teknik yang digunakan dalam penelitian ini, sehingga dapat lebih mudah mengikuti tahapan-tahapan yang dijelaskan pada bab selanjutnya.

2.1.1. Data Sampel

Berikut adalah data sampel yang diambil dari dataset emisi metana global di *Kaggle*. Data ini mencakup atribut utama: *Region, Country, Emissions, Type, Segment, Reason*, dan *BaseYear*.

Table 1. Sampel data dari dataset emisi metana global (*Kaggle*)

Region	Country	Emissions	Type	Segment	Reason	BaseYear
Africa	Algeria	130.79	Energy	Gas pipelines and LNG facilities	Fugitive	2022
Africa	Angola	47.64	Energy	Offshore oil	Flared	2022
Africa	Benin	0.05	Energy	Gas pipelines and LNG facilities	Vented	2022
Africa	Botswana	5.40	Energy	Bioenergy	All	2022
Africa	Cameroon	194.51	Waste	Total	All	2019-2021
Europe	Netherlands	14.42	Energy	Bioenergy	All	2022
Europe	Ukraine	335.74	Energy	Steam coal	All	2022
Europe	United Kingdom	38.53	Energy	Gas pipelines and LNG facilities	Fugitive	2022

Atribut numerik yang akan digunakan dalam perhitungan klusterisasi adalah *emissions*. Atribut kategorikal (*type, segment, reason*) akan diperhitungkan di *RapidMiner* melalui *encoding* otomatis.

2.1.2. Tahapan Perhitungan Manual *K-Means*

Centroid awal dalam proses klusterisasi *K-Means* dipilih secara manual berdasarkan nilai *emissions* dari sejumlah negara yang tersedia dalam data yang digunakan. Pemilihan ini dilakukan untuk memberikan titik awal bagi algoritma dalam menentukan klaster, sehingga distribusi data dapat dioptimalkan di setiap iterasi. Sebagai contoh, *centroid* awal yang dipilih adalah sebagai berikut:

$$C_1 = 130.79 \text{ (Algeria)}$$

$$C_2 = 47.64 \text{ (Angola)}$$

$$C_3 = 0.05 \text{ (Benin)}$$

$$C_4 = 5.40 \text{ (Botswana)}$$

$$C_5 = 194.51 \text{ (Cameroon)}$$

Setelah *centroid* awal ditentukan, langkah berikutnya adalah menghitung jarak *Euclidean* antara setiap negara dengan masing-masing *centroid*. Jarak *Euclidean* digunakan karena metrik ini mampu memberikan ukuran jarak linier dalam ruang multidimensi yang sesuai dengan data numerik. Rumus umum yang digunakan untuk menghitung jarak *Euclidean* dalam satu dimensi yang dituliskan sebagai berikut:

$$d(i, k) = \sqrt{(x_i - C_k)^2}$$

Untuk memberikan gambaran lebih jelas mengenai perhitungan ini, misalkan dilakukan perhitungan untuk negara Algeria, yang memiliki nilai *emission* sebesar 130.79. Langkah-langkah perhitungan jarak antara Algeria dengan semua *centroid* dapat diuraikan sebagai berikut:

$$d(1,1) = \sqrt{(130.79 - 130.79)^2} = 0.00$$

$$d(1,2) = \sqrt{(130.79 - 47.64)^2} = 83.15$$

$$d(1,3) = \sqrt{(130.79 - 0.05)^2} = 130.74$$

$$d(1,4) = \sqrt{(130.79 - 5.40)^2} = 125.39$$

$$d(1,5) = \sqrt{(130.79 - 194.51)^2} = 63.72$$

Maka didapatkan hasil dari perhitungan jarak untuk negara Algeria dan untuk semua negara dengan cara yang sama:

Table 2. Hasil Perhitungan Jarak *Euclidean* untuk Seluruh Negara

Country	$d(i, 1)$	$d(i, 2)$	$d(i, 3)$	$d(i, 4)$	$d(i, 5)$	Klaster
Algeria	0.00	83.15	130.74	125.39	63.72	1
Angola	83.15	0.00	47.59	42.24	146.87	2
Benin	130.74	47.59	0.05	5.35	194.46	3
Botswana	125.39	42.24	5.35	0.00	189.11	4
Cameroon	63.72	146.87	194.46	189.11	0.00	5
Netherlands	116.37	33.22	14.37	9.02	180.09	4
Ukraine	204.95	288.10	335.69	330.34	141.23	5
United Kingdom	92.26	9.11	38.48	33.13	155.98	2

Proses perhitungan ini diulangi secara iteratif hingga *centroid* stabil, yaitu ketika posisi *centroid* tidak lagi mengalami perubahan signifikan di antara iterasi. Pada setiap iterasi, negara-negara dalam dataset akan dievaluasi kembali, dan masing-masing negara akan ditugaskan ulang ke klaster yang memiliki jarak paling dekat dengan *centroid* baru yang telah diperbarui.

Dengan cara ini, distribusi anggota klaster secara bertahap akan semakin sesuai dengan pola data yang sebenarnya. Proses ini terus berlanjut hingga konvergensi tercapai, yaitu ketika tidak ada lagi perubahan pada komposisi klaster atau posisi *centroid*, sehingga hasil klasterisasi dianggap optimal dan representatif untuk analisis lebih lanjut.

2.2. Tahapan Analisis Data

2.2.1. Input Data

Tahap pertama dalam analisis data adalah pengumpulan data emisi metana global, yang diambil dari dataset yang tersedia di *Kaggle*. Data ini mencakup informasi terkait emisi, sektor, sub-sektor, dan negara. Proses input data dilakukan dengan mengikuti beberapa langkah spesifik untuk memastikan format dan kualitas data yang tepat. Dataset ini memiliki beberapa kolom penting, antara lain:

- Region*: Menunjukkan wilayah geografis tempat negara berada, misalnya Afrika, Asia, Eropa, dan lain-lain.
- Country*: Nama negara yang bersangkutan, seperti Algeria, Angola, dan lain-lain.

- c. *Emissions*: Jumlah emisi metana dalam satuan yang relevan, yang mencerminkan kontribusi masing-masing negara terhadap total emisi metana.
- d. *Type*: Menunjukkan kategori emisi, seperti energi, pertanian, dan limbah.
- e. *Segment*: Merinci jenis segmen dari sumber emisi.
- f. *Reason*: Menggambarkan alasan di balik emisi yang tercatat,
- g. *BaseYear*: Tahun dasar yang digunakan untuk estimasi data emisi, yang mencakup periode dari 2019 hingga 2022.
- h. *Notes*: Menyediakan informasi tambahan terkait data, seperti estimasi atau catatan penting lainnya.

Proses input data dilakukan dengan menggunakan perangkat lunak *RapidMiner*, di mana format data diatur dengan opsi berikut:

- a. *Header Row*: Baris pertama data diidentifikasi sebagai *header* untuk mendefinisikan nama kolom.
- b. *File Encoding*: Menggunakan *encoding* “*windows-1252*” untuk memastikan bahwa karakter dalam data ditampilkan dengan benar.
- c. *Column Separator*: Menggunakan koma (“,”) sebagai pemisah antar kolom.
- d. *Skip Comments*: Mengabaikan komentar yang tidak relevan dalam dataset./

1		region	country	emissions	type	segment	reason	baseYear	notes
2	0	Africa	Algeria	257.6112060...	Agriculture	Total	All	2019-2021	Average base...
3	1	Africa	Algeria	0.052000001...	Energy	Bioenergy	All	2022	Estimates fro...
4	2	Africa	Algeria	130.7989959...	Energy	Gas pipelines...	Fugitive	2022	Not available
5	3	Africa	Algeria	69.74189758...	Energy	Gas pipelines...	Vented	2022	Not available
6	4	Africa	Algeria	213.9869995...	Energy	Onshore gas	Fugitive	2022	Not available
7	5	Africa	Algeria	464.3080139...	Energy	Onshore gas	Vented	2022	Not available
8	6	Africa	Algeria	469.7869873...	Energy	Onshore oil	Flared	2022	Not available
9	7	Africa	Algeria	85.18720245...	Energy	Onshore oil	Fugitive	2022	Not available
10	8	Africa	Algeria	1154.119995...	Energy	Onshore oil	Vented	2022	Not available
11	9	Africa	Algeria	0.004000000...	Energy	Other from coal	All	2022	Estimates fro...
12	10	Africa	Algeria	14.20855045...	Energy	Other from oil ...	All	2022	Estimates fro...
13	11	Africa	Algeria	67.0	Energy	Satellite-detec...	All	2022	Not available
14	12	Africa	Algeria	2669.194580...	Energy	Total	All	2022	Estimates fro...
15	13	Africa	Algeria	6.475720882...	Other	Total	All	2019-2021	Average base...
16	14	Africa	Algeria	510.9656066...	Waste	Total	All	2019-2021	Average base...
17	15	Africa	Angola	400.9695739...	Agriculture	Total	All	2019-2021	Average base...

Gambar 2. Proses Input Data Kedalam Perangkat Lunak RapidMiner

2.2.2. Clustering

Tahap klasterisasi merupakan langkah penting dalam penelitian ini, di mana algoritma *K-Means* diterapkan untuk mengelompokkan negara-negara berdasarkan pola emisi metana mereka. Beberapa parameter yang digunakan dalam proses klasterisasi adalah sebagai berikut:

The image shows a 'Parameters' window for the 'Clustering (k-Means)' operator. It contains several settings:

- add cluster attribute:** Checked (checkbox).
- add as label:** Unchecked (checkbox).
- remove unlabeled:** Unchecked (checkbox).
- k:** Set to 5 (text input).
- max runs:** Set to 10 (text input).
- determine good start values:** Checked (checkbox).
- measure types:** Set to 'Mixed Measures' (dropdown menu).
- mixed measure:** Set to 'MixedEuclideanDistance' (dropdown menu).
- max optimization steps:** Set to 100 (text input).
- use local random seed:** Unchecked (checkbox).

Gambar 3. Pengaturan Parameter Pada Operator Clustering (K-Means)

- K (Jumlah kluster):** Parameter ini ditetapkan sebesar 5, yang menunjukkan bahwa data akan dikelompokkan menjadi lima kluster. Pemilihan jumlah kluster yang tepat sangat penting untuk memastikan bahwa setiap kelompok mencerminkan pola emisi yang relevan.
- Max Runs:** Parameter ini diatur menjadi 10, yang menunjukkan jumlah maksimum iterasi yang akan dilakukan untuk mencari kluster yang optimal.
- Determine Good Start Values:** Opsi ini diaktifkan untuk memungkinkan algoritma menentukan nilai awal yang baik untuk kluster. Hal ini penting untuk meningkatkan konvergensi dan mengurangi waktu komputasi.
- Measure Types:** Jenis pengukuran yang digunakan dalam proses klusterisasi adalah *Mixed Measures*, yang memungkinkan penggunaan berbagai jenis data dalam analisis.
- Mixed Measure:** Parameter ini diatur pada *MixedEuclideanDistance*, yang memungkinkan pengukuran jarak antara data baik dalam bentuk numerik maupun kategorikal.
- Use Local Random Seed:** Opsi ini diaktifkan untuk memastikan bahwa hasil yang diperoleh konsisten setiap kali algoritma dijalankan, dengan menggunakan *seed* acak lokal.

Setelah parameter-parameter ini ditetapkan, algoritma *K-Means* dijalankan untuk mengelompokkan data.

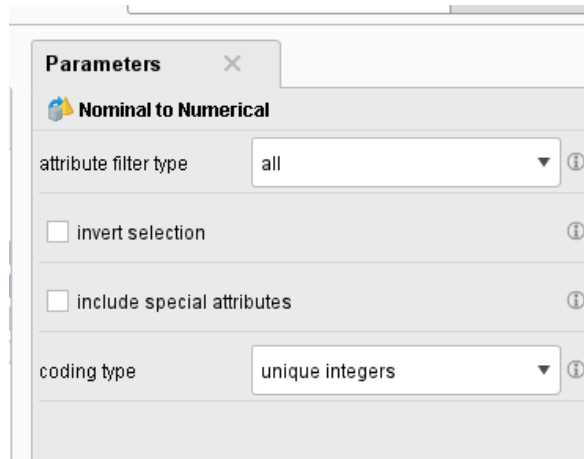
2.2.3. Post-Processing

Tahap *post-processing* merujuk pada langkah-langkah yang dilakukan setelah algoritma klusterisasi *K-Means*. Tujuan dari tahap ini adalah untuk menganalisis dan mengevaluasi hasil kluster yang telah dihasilkan, serta untuk menyiapkan data untuk interpretasi lebih lanjut. Dalam konteks ini, meskipun operator "*Nominal to Numerical*" dapat terlihat seperti alat *pra-processing* (pra-pemrosesan), penggunaannya setelah klusterisasi menunjukkan bahwa data yang dihasilkan dari kluster perlu dipersiapkan untuk evaluasi dan analisis lebih lanjut.

- Urutan Proses: Post-processing** dilakukan setelah kluster terbentuk. Ini berbeda dengan *pre-processing*, yang mencakup langkah-langkah pengolahan data sebelum analisis dimulai. Dalam

hal ini, operator “*Nominal to Numerical*” digunakan untuk mengonversi atribut nominal yang ada dalam hasil kluster menjadi format numerik agar dapat dievaluasi dengan lebih baik.

- b. Fokus pada Hasil: *Post-processing* berfokus pada analisis hasil kluster yang sudah ada, termasuk perhitungan kinerja kluster dan penyiapan data untuk interpretasi.



Gambar 4. Pengaturan Parameter Pada Operator Nominal to Numerical

Dalam tahap *post-processing* ini, beberapa parameter yang digunakan dalam *RapidMiner* adalah sebagai berikut:

- a. *Attribute Filter Type*: Diatur untuk memilih semua atribut yang ada dalam dataset untuk dianalisis lebih lanjut.
- b. *Invert Selection*: Opsi ini diaktifkan, yang berarti bahwa semua atribut yang dipilih akan digunakan dalam analisis.
- c. *Include Special Attributes*: Opsi ini diaktifkan untuk memastikan bahwa atribut khusus juga dimasukkan dalam analisis.
- d. *Coding Type*: Diatur pada *unique integers*, yang berarti bahwa nilai-nilai dapat digunakan dalam analisis kuantitatif.

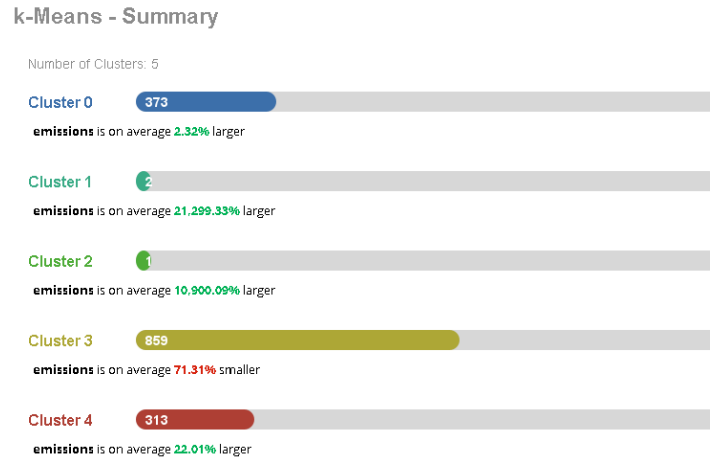
Setelah tahap ini selesai, hasil kluster akan dievaluasi menggunakan operator *Performance* untuk menghitung metrik yang menunjukkan seberapa baik kluster terbentuk. Di dalam parameter *Performance*, berikut parameter yang digunakan dalam operator *Performance*:

- a. *Avg. Within Centroid Distance*: Parameter ini diatur untuk mengukur jarak rata-rata antara setiap titik data dalam kluster ke *centroid* kluster tersebut.
- b. *Maximize*: Opsi ini diaktifkan untuk memastikan bahwa tujuan evaluasi adalah memaksimalkan kriteria evaluasi yang ditetapkan.

Dengan menggunakan parameter-parameter ini, evaluasi kinerja kluster memberikan wawasan mengenai efektivitas algoritma *K-Means* dalam mengelompokkan data.

3. ANALISA DAN PEMBAHASAN

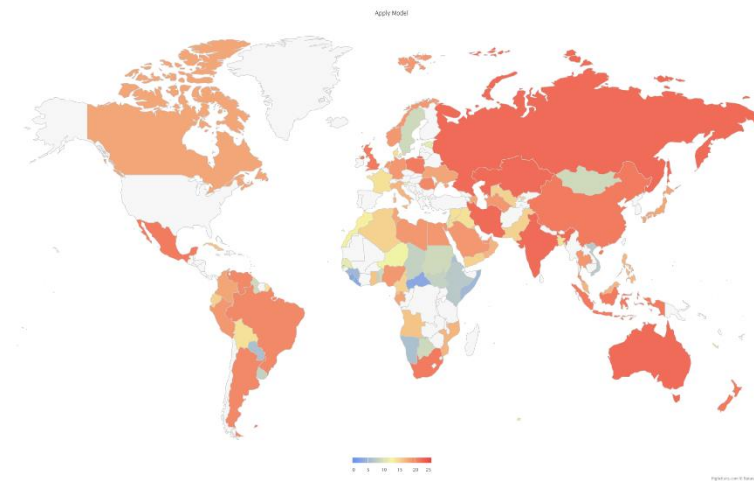
3.1. Interpretasi



Gambar 5. Ringkasan Kluster

Dari analisis *K-Means Summary* (Gambar 5) menunjukkan empat kluster yang berbeda:

- Cluster 0*, diperoleh sebanyak 373 entri dengan emisi rata-rata yang hanya sedikit lebih tinggi dibandingkan nilai acuan. Hal ini menunjukkan bahwa meskipun kelompok ini memiliki banyak entri, kontribusinya terhadap total emisi relatif moderat. Langkah mitigasi di kelompok ini dapat difokuskan pada peningkatan efisiensi teknologi dan penerapan praktik ramah lingkungan untuk mempertahankan tingkat emisi pada level yang rendah.
- Cluster 1*, klaster ini menarik perhatian karena meskipun hanya terdiri dari 3 entri, tingkat emisi yang dihasilkan sangat signifikan, yaitu 21.29% lebih tinggi dari nilai acuan. Fakta ini mengindikasikan adanya entitas atau wilayah dengan tingkat emisi yang sangat tinggi namun terbatas secara jumlah. Perlu dilakukan analisis lebih lanjut untuk mengidentifikasi penyebab spesifik tingginya emisi pada klaster ini, seperti sektor industri berat, penggunaan energi yang tidak efisien, atau sumber daya alam yang dipergunakan secara intensif.
- Cluster 2*, dengan hanya satu entri namun emisi rata-rata 10.90% lebih tinggi dari acuan, klaster ini menunjukkan pentingnya fokus pada entitas tunggal yang berperan besar terhadap emisi secara keseluruhan. Pemerintah atau pengambil kebijakan dapat menargetkan entitas ini dengan pendekatan khusus seperti investasi dalam energi terbarukan atau pemberlakuan regulasi ketat terkait emisi.
- Cluster 3*, menjadi klaster dengan jumlah entri terbanyak, yaitu 859 entri, namun dengan emisi rata-rata yang lebih rendah dibandingkan nilai acuan. Klaster ini merepresentasikan wilayah atau entitas yang telah berhasil mengendalikan tingkat emisinya dengan baik. Analisis terhadap faktor-faktor keberhasilan di klaster ini dapat memberikan best practices yang berguna untuk diterapkan di klaster lainnya.
- Cluster 4*, Terdiri dari 313 entri dengan tingkat emisi yang signifikan, yakni 22.01% lebih tinggi dari acuan. Klaster ini perlu menjadi fokus utama dalam strategi mitigasi karena jumlah entri yang besar dan tingkat emisi yang tinggi menunjukkan adanya potensi besar untuk mengurangi emisi secara signifikan. Strategi mitigasi yang dapat diterapkan meliputi insentif bagi penggunaan teknologi hijau, edukasi masyarakat tentang pengelolaan sumber daya, serta peningkatan kerja sama dengan sektor swasta.



Gambar 6. Peta Emisi Metana Global

Gambar 6 menampilkan peta emisi metana global, di mana warna yang lebih gelap menunjukkan negara dengan kontribusi emisi yang lebih tinggi. Peta ini memberikan gambaran visual yang jelas tentang bagaimana emisi metana terdistribusi di berbagai belahan dunia. Negara-negara seperti Kazakhstan, Iran, dan European Union muncul sebagai penyumbang utama emisi, mencolok dengan warna yang lebih gelap, yang mengindikasikan tingkat emisi yang signifikan.

Visualisasi ini tidak hanya mempermudah identifikasi negara-negara dengan kontribusi emisi tinggi, tetapi juga memungkinkan analisis pola geografis yang lebih mendalam. Misalnya, kawasan yang berwarna lebih gelap mungkin mencerminkan aktivitas industri yang lebih intensif, praktik pertanian yang kurang ramah lingkungan, atau kebijakan energi yang kurang berkelanjutan.

3.2. Evaluasi

Dalam tahap evaluasi ini, kualitas kluster yang dihasilkan dari analisis *K-Means* akan ditinjau untuk menentukan seberapa baik algoritma ini dalam mengelompokkan data emisi metana. Evaluasi awal menunjukkan bahwa kluster dengan jumlah entri yang lebih tinggi, seperti *Cluster 3*, memiliki emisi rata-rata yang lebih kecil. Hal ini mengindikasikan konsistensi dalam pengelompokan data, di mana negara-negara dalam kluster ini cenderung memiliki karakteristik emisi yang rendah. Keberhasilan pengelompokan ini bisa jadi diakibatkan oleh kebijakan lingkungan yang lebih baik atau penerapan teknologi yang lebih bersih di negara-negara tersebut.

Selain itu, visualisasi peta emisi metana global memberikan gambaran yang jelas mengenai distribusi emisi di berbagai negara. Dari peta ini, terlihat bahwa negara-negara dengan kontribusi emisi tinggi, seperti Kazakhstan dan Iran, memiliki pola yang konsisten dalam hasil kluster. Evaluasi terhadap distribusi emisi ini menjadi penting untuk memahami tantangan yang dihadapi dalam pengendalian emisi.

Hasil evaluasi juga membawa implikasi bagi kebijakan mitigasi. Negara-negara yang teridentifikasi dalam kluster dengan emisi tinggi mungkin perlu meninjau dan memperbarui kebijakan lingkungan mereka untuk mengurangi emisi metana secara efektif. Dengan memahami kelebihan dan kekurangan dari kluster yang terbentuk, strategi mitigasi yang tepat dapat dirumuskan. Evaluasi ini pada akhirnya memberikan wawasan berharga tentang kualitas analisis serta potensi langkah-langkah mitigasi yang dapat diambil untuk mengatasi masalah emisi metana di tingkat global.



4. KESIMPULAN

Dari analisis kluster emisi metana global yang dilakukan menggunakan algoritma *K-Means*, ditemukan bahwa terdapat variasi signifikan dalam kontribusi emisi di antara berbagai negara. Hasil kluster menunjukkan bahwa negara-negara seperti Kazakhstan, Iran, dan negara-negara dalam European Union merupakan penyumbang utama emisi metana. Melalui visualisasi yang dihasilkan, dapat diidentifikasi pola distribusi emisi yang berkaitan dengan aktivitas industri, praktik pertanian, dan kebijakan energi.

Evaluasi kluster menunjukkan bahwa negara-negara dalam kluster dengan emisi rendah, meskipun memiliki jumlah entri yang lebih tinggi, cenderung menerapkan kebijakan lingkungan yang lebih baik. Hal ini memberikan indikasi bahwa intervensi kebijakan dapat berkontribusi pada pengurangan emisi metana secara signifikan.

REFERENCES

- Raut, A. (2022). Global methane emissions [Data set]. Kaggle. <https://www.kaggle.com/datasets/ashishraut64/global-methane-emissions>
- Junjie Wu. (2012). Advances in K-means Clustering: a Data Mining Thinking, 1-5.
- George Glocker. (2009). A Critical Potential of Methane and Its Absorption in the Ultra- Violet, 74-77.
- Markus Hoffman., & Ralf Klittenberg. (2014). RapidMiner Data Mining Use Cases and Business Analytics Applications, 157-163.
- Jiawei Han., & Micheline Kamber. (2000). Data Mining: Concepts and Techniques, 6.
- Charles Zai. (2022). Implementasi Data Mining Sebagai Pengolahan Data, 3.
- Pandey, A. (2014). Study and analysis of K-Means clustering algorithm using Rapidminer: A case study on students' exam result. International Journal of Engineering Research and Applications, 4 (12), 60-64.
- Dwi, N. (2015). Metode Nilai Jarak Guna Kesamaan Atau Kemiripan Ciri Suatu Citra, 22-23.