

Penerapan Algoritma K-Means untuk Pengelompokan Negara di Dunia Berdasarkan Indikator Ekonomi

Ricky Anggari^{*1}, Muhammad Ifandi², Ardhifa Firdaus³, Masna Wati⁴, Haviluddin⁵
^{1,2,3,4,5}Jurusan Informatika, Fakultas Teknik, Universitas Mulawarman, Jl. Sambaliung No. 9,
Samarinda, Kalimantan Timur.
e-mail: ^{*1}rickyanggari04@gmail.com, ²mifandi68@gmail.com,
³ilhamrahmananda@gmail.com, ⁴masnawati@fkti.unmul.ac.id, ⁵Haviluddin@gmail.com

Abstrak

Perkembangan ekonomi global menuntut pemahaman mendalam tentang karakteristik ekonomi negara-negara di dunia. Penelitian ini bertujuan menerapkan algoritma K-Means Clustering untuk mengelompokkan negara berdasarkan indikator ekonomi seperti Gross Domestic Product (GDP), ekspor, impor, inflasi, dan tingkat pengangguran. Metode clustering digunakan untuk mengidentifikasi pola dan struktur ekonomi negara-negara dengan menganalisis data sekunder dari World Bank tahun 2022. Proses preprocessing meliputi pembersihan data, normalisasi menggunakan Min-Max Scaling, dan seleksi variabel ekonomi kunci. Algoritma K-Means diterapkan dengan jumlah kluster optimal sebanyak 3, yang diperoleh melalui metode Elbow. Hasil clustering menunjukkan tiga kelompok negara: negara dengan ekonomi kecil, negara berkembang, dan negara dengan ekonomi raksasa. Kluster 0 terdiri dari 52 negara yang cenderung memiliki ekonomi kecil, kluster 1 mencakup 165 negara berkembang dengan karakteristik ekonomi menengah, sedangkan kluster 2 hanya terdiri dari 2 negara yang memiliki ekonomi sangat besar. Evaluasi menggunakan Silhouette Score (0,52), Davies-Bouldin Index (0,71), dan Calinski-Harabasz Index (145,73) mengindikasikan kualitas clustering yang baik. Penelitian ini memberikan wawasan tentang klasifikasi negara berdasarkan indikator ekonomi dan dapat menjadi referensi bagi pembuat kebijakan dalam merancang strategi ekonomi yang lebih efektif.

Kata kunci—K-Means, negara, clustering, ekonomi, Silhouette.

1. PENDAHULUAN

Perkembangan ekonomi global telah mendorong berbagai negara untuk meningkatkan daya saingnya dalam berbagai sektor, termasuk perdagangan internasional, investasi, dan pengelolaan sumber daya. Pengelompokan negara berdasarkan indikator ekonomi menjadi penting untuk memahami karakteristik ekonomi suatu negara serta mengidentifikasi pola yang dapat digunakan untuk pengambilan kebijakan ekonomi. Salah satu metode yang dapat digunakan untuk pengelompokan ini adalah *K-Means* clustering, yang mampu mengelompokkan negara berdasarkan kesamaan karakteristik ekonominya secara objektif dan terstruktur.

Dalam berbagai studi sebelumnya, indikator ekonomi seperti *Gross Domestic Product* (GDP), ekspor, impor, inflasi, dan tingkat pengangguran telah digunakan untuk menilai pertumbuhan dan stabilitas ekonomi suatu negara. ekspor dan impor memiliki pengaruh signifikan terhadap pertumbuhan ekonomi suatu negara [1]. Studi ini menemukan bahwa peningkatan ekspor berkontribusi terhadap peningkatan GDP, sementara impor juga dapat berdampak positif dalam mendukung kebutuhan domestik yang tidak dapat dipenuhi oleh

produksi dalam negeri. Sementara itu, penelitian lain menunjukkan bahwa perdagangan global terus mengalami transformasi akibat digitalisasi dan dinamika rantai nilai global, yang berimplikasi pada stabilitas ekonomi negara-negara di dunia [2]. Selain perdagangan, faktor inflasi juga berperan penting dalam stabilitas ekonomi suatu negara. Inflasi tidak selalu berdampak langsung terhadap tingkat kemiskinan, tetapi dapat mempengaruhi daya beli masyarakat dan pertumbuhan ekonomi secara keseluruhan [3]. Inflasi yang terlalu tinggi dapat menyebabkan ketidakstabilan harga barang dan jasa, yang pada akhirnya dapat memperlambat laju pertumbuhan ekonomi.

Dalam penelitian lain, perbedaan tingkat ekonomi antarnegara telah dianalisis menggunakan metode statistik untuk mengelompokkan negara berdasarkan indikator ekonomi utama. Pada penelitian [4] menggunakan Analisis Diskriminan Fisher untuk mengelompokkan negara menjadi dua kategori, yaitu negara maju dan berkembang, berdasarkan pendapatan per kapita, inflasi, tingkat pengangguran, dan angka kriminalitas. Model ini berhasil menunjukkan bahwa negara maju cenderung memiliki tingkat inflasi dan pengangguran yang lebih rendah dibandingkan negara berkembang. Namun, metode analisis diskriminan memiliki keterbatasan dalam menangani data yang lebih kompleks dan multidimensional sehingga metode *K-Means* dapat menjadi alternatif yang lebih fleksibel dalam mengelompokkan negara berdasarkan indikator ekonomi tanpa memerlukan asumsi linearitas data. Analisis berbasis data mining seperti clustering dapat memberikan wawasan yang lebih dalam mengenai pola pertumbuhan ekonomi negara-negara di dunia [5]. Pendekatan ini juga telah digunakan dalam berbagai studi untuk memahami keterkaitan antara perdagangan internasional dan pertumbuhan ekonomi di negara-negara ASEAN [6]. Sebagai contoh, [7] menggunakan metode *K-Means* untuk mengelompokkan 40 negara dan wilayah di Asia berdasarkan indikator kualitas pembangunan ekonomi, yang dibagi menjadi tiga kategori kualitas (tinggi, sedang, dan rendah). Penelitian lain yang dilakukan juga menerapkan algoritma *K-Means* untuk mengklasifikasikan negara-negara maju berdasarkan indikator pendapatan negara. Hasilnya menunjukkan empat klaster dengan karakteristik ekonomi yang berbeda-beda, yang mencerminkan variasi dalam struktur pendapatan dan indikator makroekonomi antar negara [8].

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan algoritma *K-Means* dalam pengelompokan negara-negara di dunia berdasarkan GDP, ekspor, impor, inflasi, dan tingkat pengangguran. Penelitian ini diharapkan dapat memberikan wawasan mengenai klasifikasi negara berdasarkan indikator ekonomi dan memberikan rekomendasi bagi pembuat kebijakan dalam merancang strategi ekonomi yang lebih efektif.

2. METODE PENELITIAN

2.1 Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode clustering untuk mengelompokkan negara-negara berdasarkan GDP, ekspor, impor, inflasi, dan tingkat pengangguran. Metode ini memungkinkan analisis pertumbuhan ekonomi untuk mengidentifikasi hubungan antara ekspor, impor, inflasi, tingkat pengangguran, dan perkembangan ekonomi. Dengan penerapan algoritma *K-Means*, penelitian ini diharapkan dapat memberikan wawasan baru dalam mengelompokkan negara berdasarkan indikator ekonomi serta mendukung analisis kebijakan ekonomi yang lebih tepat sasaran.

Pendekatan clustering telah banyak digunakan dalam penelitian ekonomi dan sosial untuk mengidentifikasi kelompok negara dengan karakteristik serupa [9], [10]. Dalam penelitian ini, metode clustering yang digunakan adalah *K-Means*. *K-Means* sering digunakan untuk data dengan distribusi yang relatif seragam dan jumlah klaster yang dapat ditentukan sebelumnya [11]. Penggunaan metode ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai pengelompokan negara di dunia berdasarkan indikator ekonomi.

2.2 Data dan Sumber Data

Data yang digunakan dalam penelitian ini berasal dari sumber sekunder, yaitu World Bank (<https://databank.worldbank.org/source/world-development-indicators>) yang menyediakan data ekonomi global yang terpercaya. Data yang dikumpulkan mencakup periode tahun 2022 dan terdiri dari beberapa variabel ekonomi seperti nilai ekspor, nilai impor, GDP, inflasi, dan tingkat pengangguran. Secara keseluruhan, dataset yang diperoleh berjumlah 3044 baris dan terdiri dari 5 kolom variabel utama. Penggunaan data ekonomi dari World Bank ini relevan untuk mengelompokkan negara-negara berdasarkan indikator makroekonomi dan perdagangan internasional. Rincian data mentah yang digunakan dapat dilihat pada Tabel 1.

Table 1. Data Mentah Indikator Ekonomi Negara-Negara di Dunia Tahun 2022

Country Name	Country Code	Series Name	Series Code	2022 [YR2022]
Afghanistan	AFG	GDP (current US\$)	NY.GDP.MKTP.CD	14497243872.13374
Afghanistan	AFG	GDP per capita (current US\$)	NY.GDP.PCAP.CD	357.2611527981439
Afghanistan	AFG	GDP per capita, PPP (current international \$)	NY.GDP.PCAP.PP.CD	2122.9958151266474
Afghanistan	AFG	Inflation, consumer prices (annual %)	FP.CPI.TOTL.ZG	..
Afghanistan	AFG	GNI (current US\$)	NY.GNP.MKTP.CD	14540297385.723482
Afghanistan	AFG	GNI per capita (constant 2015 US\$)	NY.GNP.PCAP.KD	..
Afghanistan	AFG	Life expectancy at birth, total (years)	SP.DYN.LE00.IN	62.879
Afghanistan	AFG	Literacy rate, adult total (% of people ages 15+)	SE.ADT.LITR.ZS	..
Afghanistan	AFG	Poverty headcount ratio at \$2.15 a day (2017 P...)	SI.POV.DDAY	..
Afghanistan	AFG	School enrollment, tertiary (% gross)	SE.TER.ENRR	..
Afghanistan	AFG	Unemployment, youth total (% of total labor force)	SL.UEM.1524.ZS	17.365
Afghanistan	AFG	Exports of goods and services (current US\$)	NE.EXP.GNFS.CD	2664599570.7669716
Afghanistan	AFG	Imports of goods and services (current US\$)	NE.IMP.GNFS.CD	7901784705.749463
Afghanistan	AFG	GNI per capita, PPP (current international \$)	NY.GNP.PCAP.PP.CD	2130
:				
Zimbabwe	ZWE	GNI per capita, PPP (current international \$)	NY.GNP.PCAP.PP.CD	3480

2.3 Preprocessing Data

Sebelum diterapkan ke algoritma clustering, data diproses melalui beberapa tahap preprocessing, yaitu:

- Menghapus data duplikat agar setiap negara hanya memiliki satu indikator yang unik.
- Melakukan pivoting data agar semua negara hanya memiliki satu baris, dengan indikator sebagai kolom.
- Melakukan reset index agar dataset lebih rapi setelah pivoting.
- Mengonversi nilai ".." yang sering muncul dalam dataset menjadi **NaN** agar dapat dianalisis lebih lanjut.
- Menghapus kolom tanpa nama (Unnamed) dan kolom "NaN" yang mungkin muncul setelah pivoting untuk memastikan data lebih bersih dan terstruktur.
- Mengubah Semua kolom yang seharusnya numerik ke float.

- Pembersihan Data: Pada proses pembersihan data, dilakukan penghapusan kolom yang seluruh nilainya kosong, kemudian missing value pada kolom numerik diisi dengan nilai median.
- Normalisasi Data: Menggunakan metode Min-Max Scaling agar setiap variabel memiliki skala yang seragam [14].
- Tahap terakhir dalam preprocessing adalah memilih kolom-kolom yang akan digunakan untuk proses clustering, yaitu "Country Name", "Country Code", "Export", "GDP", "Import", "Inflation", dan "Unemployment". Selain itu, nama kolom juga akan disesuaikan agar lebih mudah diproses.

2.4 Metode K-Means Clustering

Algoritma K-Means adalah salah satu metode clustering berbasis partisi yang membagi data menjadi k kluster berdasarkan kedekatan antara data dan centroid kluster [14]. Langkah-langkah utama dalam penerapan K-Means:

1. Menentukan Jumlah Kluster (K) yang akan digunakan.
2. Setelah jumlah kluster ditentukan, langkah berikutnya adalah memilih K titik centroid secara acak dari dataset. Centroid ini akan menjadi pusat awal dari setiap kluster. Pemilihan centroid awal dapat dilakukan secara acak atau menggunakan teknik seperti K-Means++, yang memilih centroid awal dengan cara yang lebih terkontrol untuk mengurangi kemungkinan konvergensi ke solusi lokal yang buruk.
3. Menghitung Jarak Setiap Data ke Centroid Terdekat.
Setiap data X_i dihitung jaraknya terhadap masing-masing centroid C_j . Jarak yang paling umum digunakan adalah Euclidean Distance, yang ditunjukkan pada persamaan (1):

$$d(X_i, C_j) = \sqrt{\sum_{k=1}^n (X_{ik} - C_{jk})^2} \quad (1)$$

Di mana X_{ik} adalah nilai fitur ke- K dari data X_i dan C_{jk} adalah nilai fitur ke- K dari centroid C_j . Semakin kecil nilai $d(X_i, C_j)$, semakin dekat data dengan centroid tersebut.

4. Mengelompokkan Data ke Centroid Terdekat.
Setelah jarak antara setiap data dan setiap centroid dihitung, setiap data akan dikelompokkan ke dalam kluster yang memiliki centroid terdekat. Data X_i akan dimasukkan ke kluster C_j jika memiliki jarak minimum ke centroid tersebut, ditunjukkan pada persamaan (2):

$$X_i \in C_j \text{ Jika } d(X_i, C_j) = \min_j d(X_i, C_j) \quad (2)$$

Dengan demikian, setiap iterasi akan memperbarui keanggotaan data dalam kluster berdasarkan centroid yang paling dekat.

5. Memperbarui Posisi Centroid.
Setelah semua data dikelompokkan ke kluster masing-masing, centroid diperbarui dengan menghitung rata-rata dari semua data dalam kluster tersebut. Posisi centroid baru dihitung dengan persamaan yang ditunjukkan pada persamaan (3):

$$C_j = \frac{1}{|C_j|} \sum_{X_i \in C_j} X_i \quad (3)$$

Di mana C_j adalah centroid baru kluster j , $|C_j|$ adalah jumlah data dalam kluster j , dan X_i adalah semua data yang termasuk dalam kluster tersebut. Proses ini memastikan bahwa centroid berpindah ke posisi yang lebih representatif dari data dalam kluster.

6. Mengecek Konvergensi
Algoritma K-Means akan berhenti jika posisi centroid tidak mengalami perubahan signifikan atau jika jumlah iterasi maksimum telah tercapai. Konvergensi dapat diperiksa dengan

menghitung perubahan posisi centroid antara iterasi sebelumnya dan iterasi saat ini dengan persamaan (4):

$$\sum_{j=1}^K \|C_j^{(t+1)} - C_j^{(t)}\| < \epsilon \quad (4)$$

Di mana $C_j^{(t)}$ adalah posisi centroid pada iterasi t , $C_j^{(t+1)}$ adalah posisi centroid pada iterasi berikutnya, dan ϵ adalah nilai toleransi yang ditetapkan. Jika perubahan posisi centroid lebih kecil dari ϵ , maka algoritma dianggap telah mencapai konvergensi.

Beberapa penelitian menunjukkan bahwa *K-Means* bekerja dengan baik pada data yang memiliki struktur kluster yang jelas [10]. Namun, algoritma ini sensitif terhadap pemilihan jumlah kluster dan dapat mengalami kesulitan dalam menangani data dengan distribusi yang tidak teratur.

2.5 Evaluasi dan Validasi Hasil Clustering

Evaluasi dan validasi hasil clustering sangat penting untuk menilai kualitas dan efektivitas pengelompokan yang dilakukan. Pemilihan metrik Silhouette Score, Davies-Bouldin Index (DBI), Calinski-Harabasz Index, dan Elbow Method didasarkan pada kemampuan mereka dalam mengukur kohesi dan pemisahan kluster dengan baik. Silhouette Score digunakan karena memberikan gambaran seberapa baik setiap titik data berada dalam klusternya dibandingkan dengan kluster lain. DBI dipilih karena mampu mengevaluasi tingkat kesamaan antar kluster berdasarkan variabilitas dalam kluster. Calinski-Harabasz Index digunakan karena dapat mengukur kepadatan dan pemisahan kluster, yang berguna dalam menilai struktur hasil clustering dengan lebih komprehensif. Dalam penelitian ini, tiga metrik evaluasi digunakan untuk menganalisis performa clustering: Silhouette Score, yang mengukur kedekatan titik data dalam satu kluster dibandingkan dengan kluster lain [10]; Davies-Bouldin Index (DBI), yang mengukur tingkat kesamaan antar kluster, di mana semakin kecil nilainya, semakin baik hasil clustering [11]; dan Calinski-Harabasz Index, yang mengukur seberapa padat dan terpisahnya kluster yang terbentuk, dengan nilai yang lebih tinggi menunjukkan struktur kluster yang lebih baik [13]. Selain itu, penelitian sebelumnya juga menunjukkan bahwa pemanfaatan metrik ini telah diterapkan dalam berbagai studi clustering, termasuk dalam analisis literasi masyarakat dan performa akademik siswa [15], [16].

2.6 Implementasi dan Tools yang Digunakan

Implementasi metode clustering dalam penelitian ini dilakukan menggunakan berbagai alat dan teknologi untuk memastikan analisis yang efisien dan akurat. Data diproses dan dianalisis menggunakan Python, dengan pustaka utama seperti Scikit-Learn, Pandas, dan Matplotlib. Dalam proses clustering, metode K-Means digunakan dengan jumlah kluster yang telah ditentukan menggunakan Elbow Method. Setelah clustering selesai, hasilnya divisualisasikan dalam bentuk grafik 3D scatter plot dan 2D scatter plot untuk menunjukkan distribusi kluster yang terbentuk.

Proses implementasi mencakup preprocessing data dengan pembersihan menggunakan `fillna()` untuk mengisi nilai yang hilang dan normalisasi menggunakan Min-Max Scaling, penerapan clustering dengan K-Means, evaluasi model menggunakan Silhouette Score, Davies-Bouldin Index, dan Calinski-Harabasz Index, serta visualisasi hasil clustering dengan scatter plot 3D dan 2D untuk memperjelas distribusi kluster. Metode implementasi ini juga telah digunakan dalam penelitian sebelumnya untuk pengelompokan penyebab kematian, pembangunan literasi masyarakat, serta segmentasi berbasis indikator sosial-ekonomi [15], [17]. Dengan pendekatan ini, penelitian dapat memberikan pemahaman yang lebih mendalam mengenai pengelompokan negara di dunia berdasarkan indikator ekonomi yang dianalisis menggunakan algoritma K-Means.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Preprocessing Data

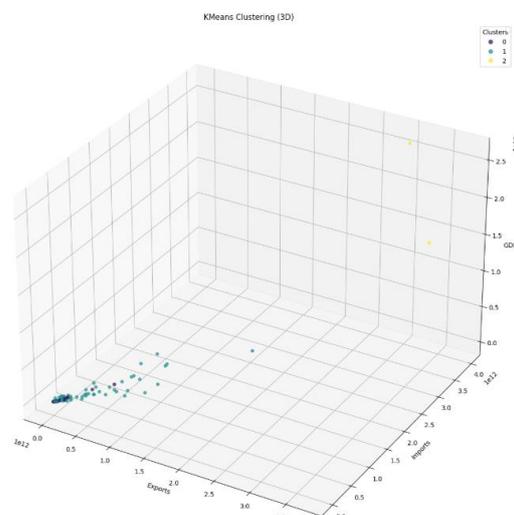
Setelah melalui tahapan preprocessing yang telah dijelaskan pada Bab 2, diperoleh data yang siap digunakan dalam proses clustering menggunakan algoritma K-Means. Hasil dari tahap preprocessing ini menghasilkan 219 baris data. Untuk lebih jelasnya, data yang telah diproses dapat dilihat pada Tabel 2.

Table 2. Data Setelah Proses Preprocessing

Country Name	Country Code	Exports	GDP	Imports	Inflation	Unemployment
Afghanistan	AFG	2.664600e+09	1.449724e+10	7.901785e+09	7.930929	17.3650
Albania	ALB	7.073860e+09	1.901724e+10	9.033376e+09	6.725203	24.7350
Algeria	DZA	6.951690e+10	2.256385e+11	4.601535e+10	9.265516	31.6620
American Samoa	ASM	4.090000e+08	8.710000e+08	6.770000e+08	7.930929	12.9845
Andorra	AND	1.766543e+10	3.380613e+09	2.167707e+10	7.930929	12.9845
Angola	AGO	4.634476e+10	1.043997e+11	2.641254e+10	21.355290	28.2880
Antigua and Barbuda	ATG	1.021456e+09	1.867733e+09	1.175559e+09	7.531078	12.9845
Argentina	ARG	1.028158e+11	6.327901e+11	9.681494e+10	7.930929	19.0040
Armenia	ARM	9.800655e+09	1.951351e+10	9.988144e+09	8.640911	26.3540
:						
Zimbabwe	ZWE	9.139216e+09	3.278975e+10	1.209661e+10	104.7051	16.984

3.2 Hasil Penerapan Algoritma K-Means Clustering

Dalam penelitian ini, metode clustering diterapkan untuk mengelompokkan negara-negara berdasarkan indikator ekonomi seperti ekspor, impor, GDP, inflasi, dan tingkat pengangguran menggunakan algoritma K-Means. Data yang digunakan dalam proses clustering ini merupakan hasil preprocessing yang telah disajikan pada Tabel 2. Hasil pengelompokan kemudian divisualisasikan menggunakan scatter plot 3D yang menggambarkan perbedaan kluster dalam ruang fitur ekspor, impor, GDP, inflasi, dan pengangguran, seperti terlihat pada Gambar 1.



Gambar 1. Hasil Cluster dengan K-Means

Metode *K-Means* diterapkan dengan jumlah kluster optimal sebanyak 3 kluster berdasarkan *Elbow Method*. Distribusi negara dalam masing-masing kluster menunjukkan pola berdasarkan nilai Ekspor, Impor, GDP, Inflasi, dan Pengangguran.

- Cluster 0 (Ungu) → Kelompok ini mencakup beberapa negara dengan tingkat ekonomi relatif kecil dibandingkan dengan kelompok lainnya, dan terdiri atas 52 negara.
- Cluster 1 (Biru Muda) → Kelompok terbesar terdiri atas 165 negara dengan perekonomian tingkat menengah, yang mengindikasikan bahwa sebagian besar negara di dunia masih tergolong sebagai negara berkembang.
- Cluster 2 (Kuning) → Terdapat dua titik yang jauh terpisah dari kluster lainnya, menunjukkan bahwa negara-negara dalam kelompok ini memiliki ekonomi yang jauh lebih besar dibandingkan yang lain. Kedua negara tersebut adalah Amerika Serikat dan China, yang dikenal sebagai negara dengan perekonomian terbesar di dunia.

3.3 Evaluasi Hasil Clustering

Penilaian kualitas hasil clustering dilakukan dengan cara evaluasi menggunakan tiga metrik yang komprehensif yaitu Silhouette Score untuk mengukur kedekatan data dalam kluster, Davies-Bouldin Index untuk mengevaluasi tingkat kesamaan antar kluster, dan Calinski-Harabasz Index untuk mengukur kepadatan dan pemisahan kluster sebagaimana ditunjukkan pada hasil visualisasi di Gambar 2.

```
KMeans_Cluster:
Silhouette Score: 0.52
Davies-Bouldin Index: 0.71
Calinski-Harabasz Index: 145.73
```

Gambar 2. Hasil Evaluasi

1. Silhouette Score: Untuk menilai seberapa baik data dikelompokkan dalam kluster tertentu.
2. Davies-Bouldin Index (DBI): Menunjukkan tingkat kesamaan antar kluster, di mana nilai lebih rendah menunjukkan pemisahan kluster yang lebih baik.
3. Calinski-Harabasz Index: Mengukur kepadatan dan pemisahan kluster, di mana nilai lebih tinggi menunjukkan struktur kluster yang lebih baik.

Evaluasi dilakukan terhadap berbagai jumlah kluster, mulai dari 2 hingga 9, dengan menggunakan *Elbow Method* untuk menentukan jumlah kluster yang optimal. Hasil evaluasi menunjukkan bahwa *K-Means* dengan 3 kluster memberikan hasil terbaik, ditunjukkan oleh nilai Silhouette Score sebesar 0,52. Nilai ini menunjukkan bahwa pengelompokan memiliki kualitas yang cukup baik. Selain itu, nilai Davies-Bouldin Index yang rendah (0,71) dan Calinski-Harabasz Index yang tinggi (145,73) semakin memperkuat bahwa struktur kluster pada $k = 3$ memiliki pemisahan yang jelas dan distribusi yang baik dibandingkan dengan jumlah kluster lainnya.

3.4 Analisis Pola Perdagangan Global Berdasarkan Hasil Clustering

Inflasi dan tingkat pengangguran merupakan indikator penting dalam menilai kestabilan dan tingkat kesejahteraan ekonomi suatu negara. Berdasarkan hasil clustering menggunakan algoritma *K-Means*, negara-negara dikelompokkan ke dalam tiga kluster yang merepresentasikan tingkat perkembangan ekonomi yang berbeda. Pengelompokan ini dilakukan berdasarkan indikator GDP, ekspor, impor, inflasi, dan pengangguran.

1. Kluster 0 – Negara Kurang Berkembang

Kluster ini terdiri dari 52 negara yang umumnya memiliki GDP rendah, tingkat inflasi yang tinggi, serta pengangguran yang cukup signifikan. Negara-negara dalam kelompok ini juga menunjukkan keterbatasan dalam pembangunan ekonomi dan infrastruktur. Karakteristik ini

menunjukkan bahwa klaster 0 merepresentasikan negara-negara dengan tingkat perkembangan ekonomi yang masih tertinggal.

2. Klaster 1 – Negara Berkembang

Klaster ini merupakan kelompok terbesar dalam hasil clustering, dengan total 165 negara. Negara-negara dalam klaster ini memiliki GDP dan volume perdagangan internasional yang berada di tingkat menengah, namun masih menghadapi tantangan dalam pengendalian inflasi dan penciptaan lapangan kerja. Indonesia, Brasil, dan India merupakan contoh negara berkembang yang termasuk dalam kelompok ini.

3. Klaster 2 – Negara Maju

Klaster ini terdiri dari dua negara dengan GDP yang sangat tinggi, inflasi yang stabil, dan tingkat pengangguran yang rendah. Berdasarkan hasil clustering, kedua negara dalam klaster ini adalah Amerika Serikat dan China, yang mewakili negara-negara dengan struktur ekonomi kuat dan daya beli masyarakat yang tinggi.

Hasil clustering ini menunjukkan hubungan yang konsisten antara tingkat GDP suatu negara dengan kestabilan inflasi dan pengangguran. Semakin tinggi GDP suatu negara, semakin cenderung negara tersebut memiliki kestabilan ekonomi yang lebih baik.

4. KESIMPULAN

Penelitian ini menghasilkan pengelompokan negara-negara berdasarkan indikator ekonomi menggunakan algoritma *K-Means*. Berdasarkan analisis clustering yang komprehensif, ditemukan bahwa negara-negara dapat dibagi menjadi tiga kelompok yaitu kelompok negara dengan ekonomi kecil sebanyak 52 negara, negara dengan ekonomi menengah sebanyak 165 negara, serta negara dengan ekonomi raksasa terdiri dari negara Amerika Serikat dan China. Hasil evaluasi clustering menunjukkan kualitas pengelompokan dengan *Silhouette Score* sebesar 0.52, *Davies-Bouldin Index* sebesar 0.71, dan *Calinski-Harabasz Index* sebesar 145.73.

5. SARAN

Penelitian selanjutnya disarankan untuk memperluas lingkup analisis dengan menambahkan variabel ekonomi tambahan seperti angka kemiskinan dan tingkat literasi warga. Selain itu disarankan menggunakan algoritma genetika dan DBSCAN.

DAFTAR PUSTAKA

- [1] N. A. Mandai, M. Amelia, V. E. Kamila, N. H. Bramantya, and D. Desamawan, “Analisis Pengaruh Ekspor dan Impor Terhadap Pertumbuhan Ekonomi di Indonesia Periode 2019-2023,” vol. 12, no. 2, pp. 248–256, May 2024.
- [2] M. I. Fahmi, A. A. Zuheri, and N. Kholis, “Transformasi Perdagangan Global: Pengaruh Perdagangan Digital, Dinamika Rantai Nilai Global (GVC), dan Geopolitik,” *AL-IQTISHOD: Jurnal Pemikiran dan Penelitian Ekonomi Islam*, vol. 12, no. 2, pp. 237–252, Jul. 2024, doi: 10.37812/aliqtishod.v12i2.1581.
- [3] M. Kolibu, V. A. Rimate, and D. S. M. Engka, “Pengaruh Tingkat Inflasi, Investasi, Pertumbuhan Ekonomi Dan Tingkat Pengangguran Terhadap Tingkat Kemiskinan Di Provinsi Sulawesi Utara,” vol. 12, pp. 55–77, 2024.
- [4] U. A. Gani, Salasi R, R.M.Bambang, and K. Umam, “Analisis Diskriminan Untuk Mengelompokkan Negara Maju Dan Negara Berkembang Dengan Metode Fishers,” *Jurnal*

- Geuthèè: Penelitian Multidisiplin*, vol. 01, no. 01, pp. 1–12, Mar. 2018, [Online]. Available: <http://www.journal.geutheeinstitute.com>.
- [5] I. S. Arfiani, “Analisis Empiris Hubungan antara Ekspor, Impor, Nilai Tukar dan Pertumbuhan Ekonomi di Indonesia,” *Jurnal Ekonomi Pembangunan*, vol. 17, no. 2, pp. 81–98, Dec. 2019, doi: 10.29259/jep.v17i2.9485.
- [6] Rika Kurnia, Zuha Lazuardi Muhammad Nafaris Al-Fath, Melita Sari, and Muhammad Kurniawan, “Pengaruh Perdagangan Internasional Ekspor Dan Impor Terhadap Pertumbuhan Ekonomi Di Negara ASEAN,” *Anggaran : Jurnal Publikasi Ekonomi dan Akuntansi*, vol. 2, no. 2, pp. 280–301, Jun. 2024, doi: 10.61132/anggaran.v2i2.617.
- [7] A. Wang, “Research on the Economic Development Quality of 40 Asian Countries and Regions Based on Entropy Weighted TOPSIS and K-means Clustering,” *Academic Journal of Mathematical Sciences*, vol. 5, no. 1, pp. 64–69, 2024, doi: 10.25236/ajms.2024.050110.
- [8] I. Prabaswara, D. A. A. Pertiwi, and Jumanto, “Analysis of K-means clustering algorithm in advance country clustering using rapid miner,” *Journal of Studen Research Exploration*, vol. 2, no. 2, pp. 101–110, Jun. 2024.
- [9] S. Mutiah, Y. Hasnataeni, A. Fitrianto, E. Erfiani, and L. M. R. D. Jumansyah, “Perbandingan Metode Klastering K-Means dan DBSCAN dalam Identifikasi Kelompok Rumah Tangga Berdasarkan Fasilitas Sosial Ekonomi di Jawa Barat,” *Teorema: Teori dan Riset Matematika*, vol. 9, no. 2, p. 247, Sep. 2024, doi: 10.25157/teorema.v9i2.16290.
- [10] S. D. K. Wardani, A. S. Ariyanto, M. Umroh, and D. Rolliawati, “Perbandingan Hasil Metode Clustering K-Means, Db Scanner & Hierarchical Untuk Analisa Segmentasi Pasar,” *JIKO (Jurnal Informatika dan Komputer)*, vol. 7, no. 2, pp. 191–201, Sep. 2023, doi: 10.26798/jiko.v7i2.796.
- [11] Risman, Syaripuddin, and Suyitno, “Implementasi Metode Dbscan Pada Pengelompokan Kabupaten/Kota Di Pulau Kalimantan Berdasarkan Indikator Kesejahteraan Rakyat,” *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, pp. 22–28, Oct. 2019.
- [12] Ramadhana, Islamiyah, and A. P. A. Masa, “Penerapan Data Mining Menggunakan Metode K-Means Clustering Pada Data Ekspor Batubara,” *Adopsi Teknologi dan Sistem Informasi (ATASI)*, vol. 2, no. 1, pp. 35–42, Jun. 2023, doi: 10.30872/atasi.v2i1.595.
- [13] M. Ardana Wijaya, D. Satria Prayoga, A. Karunia Rahman, and A. Puspita Sari, “Perbandingan Algoritma K-Means dan DBSCAN dalam Metode Clustering dengan PCA untuk Analisis Data Statistik Negara Dunia,” *Seminar Nasional Informatika Bela Negara (SANTIKA)*, vol. 3, pp. 63–70, 2023, [Online]. Available: <https://ourworldindata.org/>
- [14] F. M. Pranata, S. H. Wijoyo, and N. Y. Setiawan, “Analisis Performa Algoritma K-Means dan DBSCAN Dalam Segmentasi Pelanggan Dengan Pendekatan Model RFM,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 1, Jan. 2017, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [15] M. Jelita, “Analisis Clustering Menggunakan Metode K-Means untuk Mengelompokkan Kabupaten/Kota di Indonesia berdasarkan Unsur- Unsur Pembangun Literasi Masyarakat

(UPLM),” pp. 701–710, 2023, [Online]. Available:
<https://docs.rapidminer.com/latest/studio/guided/index.html>

- [16] N. Bili, R. T. Abineno, and A. A. Pekuwali, “PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PENGELOMPOKAN PERFORMA SISWA PADA PEMBELAJARAN BAHASA INDONESIA (Studi Kasus: SD INPRES WAINGAPU 3),” *SATI: Sustainable Agricultural Technology Innovation*, pp. 523–537, Aug. 2024.
- [17] A. Nursikuwagus, Suherman, and I. Alamsyah, “KLUSTERISASI PENYEBAB KEMATIAN DI INDONESIA DENGAN PENERAPAN ALGORITMA K-MEANS,” *JSiI (Jurnal Sistem Informasi)*, vol. 11, no. 2, pp. 56–63, Sep. 2024, doi: 10.30656/jsii.v11i2.9276.