IT-EXPLORE
Jurnal Penerapan Teknologi Informasi dan Komunikasi
ISSN 2828-7940, e-ISSN 2829-1727, DOI: 10.24246/itexplore.v4i2.2025.pp128-137
Volume 04 Nomor 02 Tahun 2025
Journal homepage: https://ejournal.uksw.edu/itexplore

Pengelompokan wilayah produksi tuna, cakalang, tongkol dan udang di Indonesia menggunakan algoritma K-Means

Eliyani[1], Saruni Dwiasnati[2], Sutan Mohammad Arif [3] , Reza Avrizal[4]
[1,2] Departement of Computer Science, Universitas Mercu Buana, Jakarta, Indonesia
[3,4] Departement of Mathematics and Science, Universitas Indraprasta PGRI, Jakarta, Indonesia

Informasi Artikel

Riwayat Artikel:

Dikirim:
24 Desember 2025
Direvisi:
26 Februari 2025
Diterima:
25 Juni 2025

Kata Kunci:

clustering,
K-means,
cakalang,
ikan tuna,
ikan tongkol

ABSTRAK

Penelitian ini bertujuan untuk mengelompokkan daerah produksi produk perikanan Indonesia khususnya Cakalang, Tuna, Tuna Makarel, dan Udang dengan menggunakan algoritma *K-means Clustering*. Data yang digunakan adalah data produksi tahunan tiap provinsi selama 3 tahun terakhir (2019 – 2021). Penentuan jumlah *cluster* menggunakan Metode *Elbow*. Untuk setiap komoditas diperoleh tiga klaster yaitu klaster dengan produksi rendah, produksi sedang, dan produksi tinggi. Untuk Cakalang terdapat 19 provinsi yang termasuk dalam klaster rendah, 13 provinsi tergolong sedang, dan 2 provinsi termasuk klaster tinggi. Untuk Tuna, terdapat 22 provinsi yang berada pada klaster rendah, 9 provinsi pada klaster menengah, dan 3 provinsi pada klaster tinggi. Untuk Tuna Tenggiri, terendah sebanyak 19 provinsi, sedang sebanyak 12 provinsi, dan tinggi sebanyak 3 provinsi. Untuk udang, 23 provinsi tergolong rendah, 7 provinsi sedang, dan 4 provinsi tinggi. Klaster produksi tinggi untuk Tuna Cakalang adalah Provinsi Sulawesi Utara dan Maluku Utara, Tuna adalah Provinsi Sulawesi Utara, Maluku Utara, dan Maluku, Tuna Tenggiri adalah Provinsi Aceh, Jawa Timur, dan Maluku, dan untuk udang adalah Provinsi Sumatera Utara, Kalimantan Barat, Kalimantan Selatan dan Provinsi Kalimantan Timur.

Keywords:

Flood,
Landslide,
Central Java,
K-Means,
Disaster Mitigation

ABSTRACT

The research was intended to cluster the production areas of Indonesia's fishery products especially Skipjack Tuna, Tuna, Mackarel Tuna, and shrimp using K-means Clustering. The data used was annual production data for each province for the last 3 years (2019 – 2021). Determination of the number of clusters using the Elbow Method. For each commodity, three clusters were obtained, namely clusters with low production, medium production, and high production. For Skipjack Tuna, there were 19 provinces belonging to the low cluster, 13 provinces being medium, and 2 provinces being high. For Tuna, there were 22 provinces in the low cluster, 9 provinces in the middle, and 3 provinces in the high cluster. For Mackarel Tuna, low was 19 provinces, medium was 12 provinces, and high was 3 provinces. For shrimp, 23 provinces were low, 7 provinces were medium, and 4 provinces were high. High production clusters for Skipjack Tuna were North Sulawesi and North Maluku Provinces, Tuna were North Sulawesi, North Maluku and Maluku Provinces, for Mackarel Tuna were Aceh, East Java and Maluku Provinces, and for shrimp were North Sumatra, West Kalimantan, South Kalimantan and East Kalimantan Provinces..

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

Corresponding Author:
Saruni Dwiasnati
Departement of Computer Science, Universitas Mercu Buana, Jakarta, Indonesia
Email: saruni.dwiasnati@mercubuana.ac.id

1. PENDAHULUAN

Indonesia merupakan salah satu negara yang menjadi produsen Tuna, Cakalang, dan Tongkol (TCT) terbesar di dunia, dengan hasil tangkapan mencapai 1,3 juta ton (20,06% produksi global). Sayangnya, Indonesia menempati peringkat ketiga, di belakang Thailand dan Vietnam, dalam mengekspor TCT ke Amerika Serikat. Namun penelitian [1] menunjukkan bahwa Indonesia memiliki keunggulan komparatif di pasar AS untuk komoditas TCT tersebut. Hal ini dibuktikan dengan rata-rata nilai *Revealed Comparative Advantage* (RCA) tahunan sebesar 13,68 (RCA > 1), menunjukkan pertumbuhan positif dan memposisikan Indonesia sebagai negara yang sedang naik daun. Sementara itu, udang merupakan komoditas ekspor perikanan andalan Indonesia setelah tuna, cumi-cumi, kepiting, dan rumput laut.

Udang beku merupakan bentuk ekspor udang yang paling tinggi dibandingkan bentuk lainnya, namun sayangnya tingkat pertumbuhan ekspornya masih lebih rendah dibandingkan negara pesaing [2]. Diperlukan berbagai upaya di berbagai bidang untuk meningkatkan kesejahteraan masyarakat melalui biota laut yang merupakan salah satu sektor yang menjadi sasaran Tujuan Pembangunan Berkelanjutan (SDGs) Indonesia. Konsep Ekonomi Biru memposisikan Indonesia untuk memanfaatkan potensi sumber daya kelautan untuk meningkatkan pertumbuhan ekonomi sekaligus menjaga kesehatan ekosistem laut. Salah satu upaya untuk meningkatkan produksi sumber daya kelautan adalah dengan mengelompokkan wilayah potensial Tuna, Cakalang, Makarel (TCT), dan udang di Indonesia. Pendekatan data mining dapat digunakan untuk tujuan ini.

Salah satu teknik data mining adalah clustering. Teknik ini membagi, memisahkan, atau mengelompokkan data menjadi beberapa kelompok atau cluster berdasarkan karakteristik tertentu yang diinginkan. Dengan melakukan clustering, karakteristik data dapat diidentifikasi, dan informasi yang diperoleh dapat dimanfaatkan baik untuk memahami objek yang dikelompokkan maupun untuk tujuan implementasi seperti zonasi wilayah produksi. K-Means adalah metode pengelompokan yang bersifat unsupervised learning dan non-hierarki, tidak memerlukan proses pelatihan atau luaran target, karena teknik ini mempartisi data menjadi satu atau lebih kelompok menggunakan teknik partisi berbasis centroid. Metode ini telah teruji di berbagai bidang, seperti kesehatan ([3], [4], [5], [6], [7]), telekomunikasi ([8], [9], [10]), kriminologi ([11], [12], [13], [14], [15],), pertanian ([16], [17], [18], [19], [20], [21]), perikanan ([22], [23], [24], [25]), pendidikan ([26] dan bidang lainnya.

Mengingat pentingnya pengembangan budidaya laut di Indonesia, maka penting untuk mengelompokkan provinsi-provinsi di Indonesia ke dalam beberapa kelompok dengan peluang budidaya laut terbaik. Beberapa penelitian telah memanfaatkan K-Means untuk sektor perikanan di Indonesia, antara lain penelitian yang dilakukan oleh [25] menggunakan dataset berupa persentase produksi budidaya perikanan per provinsi dan jenis budidaya periode tahun 2000-2018 yang menghasilkan tiga klaster yaitu C1 dengan peluang budidaya laut terbaik yang berjumlah 26 provinsi, C2 yaitu klaster produksi sedang yang berjumlah 3 provinsi, dan C3 wilayah produksi rendah yang berjumlah 3 provinsi. Penelitian yang dilakukan oleh [24] menggunakan dataset produksi ikan air tawar di beberapa kabupaten/kota di Sumatera Selatan pada tahun 2007-2014 yang diperoleh dari Badan Pusat Statistik (BPS). *Clustering* dilakukan dengan menggunakan metode K-Means, dengan jumlah cluster ditentukan menggunakan Metode *Elbow*, Metode *Silhouette*, dan Metode *Gap Statistics*, dan menghasilkan dua klaster yaitu klaster produksi rendah dan klaster produksi tinggi.

Selanjutnya penelitian yang dilakukan ini bertujuan untuk mengetahui *cluster* provinsi di Indonesia yang masuk kedalam kategori penghasil TCT dan udang dengan menggunakan data terkini Badan Pusat Statistik (BPS) yang diimplementasikan dengan metode K-Means, dengan jumlah cluster ditentukan menggunakan metode Elbow.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Beberapa penelitian terdahulu yang digunakan penulis sebagai pembanding dalam menerapkan algoritma K-Means untuk mengetahui Wilayah Produksi Tuna, Cakalang, Tongkol dan Udang yang ada di wilayah Indonesia. Beberapa peneliti telah melakukan penerapan algoritma K-Means ini, digunakan kumpulan data

sebanyak 30 tipe A dan B untuk tahap pelatihan, dan 10 tipe data untuk tahap pengujian yang berhasil mencapai akurasi sebesar 90% [27].

Penelitian sebelumnya [28] K-Means, pengelompokan dan penggunaan data dukungan perangkat lunak RapidMiner dapat ditingkatkan agar nelayan dapat menggunakan sumber daya perikanan menurut provinsi lagi. Penelitian menunjukkan bahwa Jawa Tengah merupakan salah satu dari 21 provinsi di Indonesia yang mencapai tingkat klasterisasi tinggi (C1). Sementara itu, Provinsi Kalimantan Tengah, Jawa Timur, Sulawesi Tenggara, Sumatera Barat, Jawa Barat, Gorontalo, Nusa Tenggara Barat, Bangka Belitung, Sulawesi Utara, Bengkulu, Kalimantan Barat, DI Yogyakarta, Sumatera Utara, Sulawesi Tengah, Kepulauan Seribu, Bali, Lampung, Sulawesi Selatan, DKI Jakarta, dan Banten tergolong dalam klaster tingkat rendah (C2).

2.2. Landasan Teori

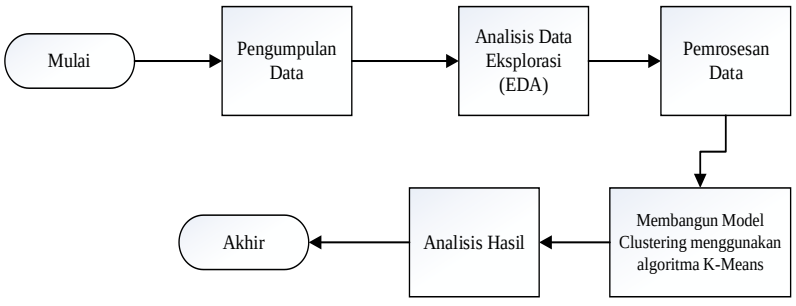
Data mining adalah suatu proses yang melibatkan pengumpulan dan pemanfaatan data historis untuk menemukan pola, keteraturan, serta hubungan dalam kumpulan data berukuran besar. Salah satu aspek utama dalam data mining adalah clustering, yaitu teknik pengelompokan data di mana kategori atau kelompoknya belum ditentukan sebelumnya [29]. Data mining memainkan peran penting dalam berbagai bidang, termasuk industri, keuangan, meteorologi, sains, dan teknologi, karena kemampuannya dalam mengolah data dalam skala besar. Proses ini mencakup berbagai metode, seperti klasifikasi, clustering, regresi, pemilihan variabel, serta analisis pola pembelian, yang sangat bermanfaat dalam mengekstraksi informasi dari kumpulan data yang luas [30].

Algoritma K-Means, yang merupakan bagian dari metode pembelajaran tanpa supervisi, digunakan untuk membagi data ke dalam beberapa klaster berdasarkan pola tertentu. Pembelajaran tanpa supervisi dalam data mining bertujuan untuk mengenali pola dari berbagai variabel atau atribut tanpa menetapkan variabel tertentu sebagai label utama [31]. K-Means merupakan metode klasterisasi data yang bersifat *unsupervised* dan *non-hierarchical clustering*. Teknik ini tidak membutuhkan proses training dan target luaran di mana data dipisahkan menjadi satu atau beberapa kelompok menggunakan teknik partisi yang didasarkan pada titik pusat (centroid).

Clustering adalah metode yang digunakan untuk mengidentifikasi dan mengelompokkan data berdasarkan karakteristik yang serupa. Tujuan utama dari teknik ini adalah mengorganisir data atau objek ke dalam kelompok (cluster) yang berisi elemen-elemen dengan kesamaan tertentu. Dalam prosesnya, *clustering* berusaha menggabungkan objek yang memiliki kemiripan dalam satu kelompok, sekaligus memastikan adanya perbedaan yang jelas antar cluster. Sebagai hasilnya, objek dalam satu cluster memiliki karakteristik yang lebih mirip satu sama lain dibandingkan dengan objek di cluster lain [32].

3. METODE PENELITIAN

Metodologi yang diterapkan dalam penelitian ini terdiri dari beberapa tahap, dimulai dengan pengumpulan data, dilanjutkan dengan analisis deskriptif terhadap data awal. Selanjutnya dilakukan prapemrosesan data, kemudian dibangun model *clustering* menggunakan algoritma K-Means. Hasil *clustering* dari algoritma ini kemudian dianalisis. Tahapan penelitian ini disajikan pada Gambar 1.



Gambar 1. Metodologi Penelitian.

Dataset diperoleh dari Badan Pusat Statistik tahun 2023 dan mencakup data produksi tahunan Cakalang, Tuna, Tenggiri, dan Udang dari masing-masing provinsi di Indonesia periode tahun 2019 hingga 2021. Analisis deskriptif kemudian dilakukan untuk memahami kondisi data yang menginformasikan langkah-langkah pemrosesan data yang diperlukan selama tahap prapemrosesan data. Tahap prapemrosesan data terdiri dari reduksi data, pembersihan data, transformasi data, dan penskalaan data. Reduksi data bertujuan untuk memperkecil volume data dengan tetap menjaga keutuhan informasi penting, dan dalam penelitian ini dilakukan pemilihan fitur. Pembersihan data adalah proses mengoreksi atau menghapus data yang salah, tidak lengkap, atau tidak relevan dari kumpulan data dengan menangani nilai yang hilang dan menghilangkan *outlier*. Transformasi data mengubah data menjadi format atau struktur yang lebih sesuai untuk analisis; misalnya, data kategorikal dapat diubah menjadi data numerik. Penskalaan data adalah proses penyesuaian skala fitur sehingga setiap fitur berkontribusi secara proporsional terhadap analisis.

Langkah-langkah dalam proses clustering menggunakan algoritma K-Means adalah sebagai berikut:

1. Inisialisasi *centroid*, yang melibatkan penentuan nilai K sebagai jumlah cluster yang diinginkan. Ada banyak cara untuk menentukan jumlah cluster K yang optimal, dan dalam penelitian ini digunakan metode *Elbow Criterion* dan *Silhouette Score*.
2. Memilih K dari dataset X sebagai centroid.
3. Menghitung jarak setiap objek terhadap setiap *centroid* pada setiap *cluster*, menggunakan rumus *Euclidean Distance* seperti yang tertera pada persamaan (1) sebagai berikut:

$$D(X_1, X_1) = \|x1 - x2\| = \sqrt{\sum_{j=1}^{p} \left|X_{2j} - X_{1j}\right|^2} \quad (1)$$

4. Mengalokasikan semua data ke *centroid* terdekat.
5. Melakukan iterasi dan menentukan posisi baru *centroid* menggunakan persamaan (1).
6. Mengulangi langkah ketiga jika posisi *centroid* baru tidak sama.

Keterangan:

- X_1 dan X_2 adalah dua vektor atau titik dalam ruang berdimensi p (dengan p adalah banyaknya variabel yang dimiliki vektor tersebut).
- $|X_{2j}-X_{1j}|$ adalah selisih antara nilai variabel ke-j X_2 dengan nilai variabel ke-j X_1.
- $\sum_{j=1}^{p} \left|X_{2j} - X_{1j}\right|^2$ adalah jumlah selisih kuadrat antara nilai variabel ke-j X_2 dan X_1 untuk semua variabel (j) dari 1 sampai p.

Metode Elbow merupakan sebuah teknik umum yang digunakan dalam analisis cluster untuk menentukan jumlah cluster yang optimal dalam suatu dataset. Jumlah cluster yang optimal adalah titik dimana reduksi inersia mulai melambat secara signifikan, ditunjukkan dengan titik "siku" pada grafik. Siku ini diperoleh dengan memplot nilai SSE (*Sum of Squares*), dimana penurunan SSE mulai melambat secara signifikan dengan bertambahnya jumlah cluster [33]. Penelitian oleh [34] menunjukkan bahwa dengan menggunakan Metode Elbow menghasilkan Davies-Bouldin Index (DBI) yang lebih rendah dibandingkan dengan menggunakan algoritma K-Means saja.

Penentuan jumlah *cluster* optimal menggunakan Metode Elbow selanjutnya dikonfirmasi dengan menggunakan metode *Silhouette Score*. Pendekatan ini mempertimbangkan Skor Silhouette untuk setiap titik data, yang mengukur kedekatan suatu titik data dengan titik lain dalam klaster yang sama dibandingkan dengan titik di klaster lain. Kisaran Skor Silhouette adalah antara -1 dan 1. Skor yang mendekati 1 menunjukkan bahwa suatu titik data memiliki kedekatan yang lebih tinggi dengan titik-titik dalam cluster yang sama dibandingkan dengan titik-titik di cluster lain, yang menunjukkan pengelompokan yang baik. Sebaliknya, nilai yang mendekati -1 menunjukkan bahwa titik data lebih dekat ke titik di cluster lain dibandingkan dengan titik di dalam cluster yang sama. Persamaan (2) merupakan rumus untuk memperoleh *Silhouette Score*.

$$\frac{(b-a)}{\mathrm{Max}(a,b)} \quad (2)$$

Keterangan:

- a = kohesi, yaitu jarak antara suatu titik data dengan titik data lain dalam cluster yang sama, kemudian dirata-ratakan.
- b = separasi, yaitu jarak antara suatu titik data dengan titik data lain dalam cluster yang berbeda.

4. HASIL DAN PEMBAHASAN

4.1. Pemrosesan awal

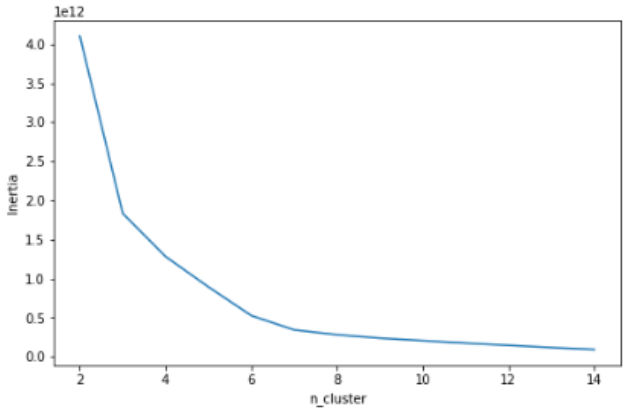
Hasil analisis deskripsi data seperti disajikan pada Gambar 3 menunjukkan bahwa tidak ada data yang hilang sehingga tidak diperlukan penanganan terhadap data yang hilang. Tipe datanya sesuai untuk diproses karena hampir semua tipe datanya bilangan bulat dan tidak perlu dilakukan reduksi data. Data atribut yang siap diolah terdiri dari Nama Provinsi, produksi tahun 2019, produksi tahun 2020, dan produksi tahun 2021 untuk masing-masing ikan cakalang, ikan tongkol, ikan tuna, dan udang. Penskalaan data tidak diperlukan karena skala data tidak berbeda nyata.



Gambar 2. Tidak ada data yang hilang.

4.2. Jumlah Cluster

Jumlah cluster optimal yang diperoleh dengan menggunakan metode Elbow adalah tiga *cluster* untuk seluruh dataset seperti disajikan pada Gambar 3. Titik siku, seperti yang ditunjukkan pada Gambar 3, terjadi pada tiga cluster.



Gambar 3. Jumlah cluster yang optimal berdasarkan metode Elbow

4.3. Pemodelan Klaster

Hasil pengolahan data menggunakan algoritma clustering K-Means menunjukkan bahwa dataset ikan cakalang, ikan tongkol, tuna, dan udang dapat dibagi menjadi tiga klaster, yaitu klaster Produksi Tinggi, Klaster Produksi Sedang, dan Klaster Produksi Rendah. Jumlah masing-masing provinsi untuk masing-masing klaster hasil laut bernilai ekonomi tinggi tersebut disajikan pada Tabel 1.

Tabel 1. Jumlah Wilayah untuk Tiap Klaster

Klaster	Cakalang	Tuna	Tongkol	Udang
Tinggi	2	3	3	4
Sedang	13	9	12	7
Rendah	19	22	19	23

Tabel 2 menunjukkan wilayah provinsi di Indonesia yang tergolong dalam klaster produksi tinggi untuk empat jenis komoditas perikanan, yaitu cakalang, tuna, tongkol, dan udang. Untuk komoditas cakalang, provinsi Sulawesi Utara dan Maluku Utara tercatat sebagai wilayah dengan produksi tinggi. Pada komoditas tuna, provinsi Sulawesi Utara, Maluku Utara, dan Maluku masuk dalam klaster produksi tinggi. Sementara itu, produksi tongkol yang tinggi ditemukan di provinsi Aceh, Jawa Timur, dan Maluku. Adapun komoditas udang memiliki wilayah produksi tinggi yang tersebar di Sumatera Utara, Kalimantan Barat, Kalimantan Selatan, dan Kalimantan Timur.

Tabel 2. Wilayah Klaster Produksi Tinggi

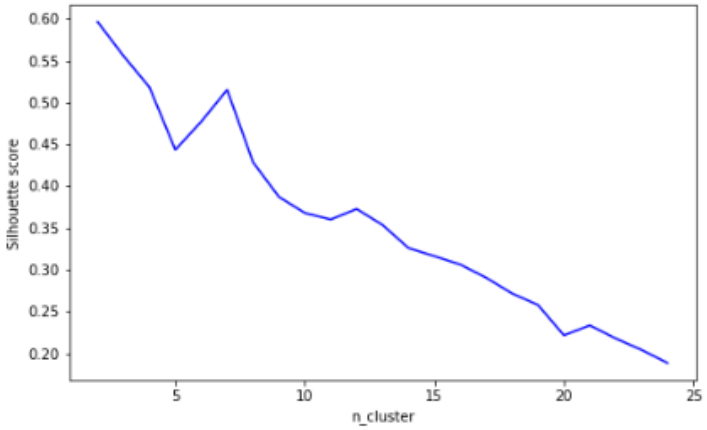
Klaster Provinsi dengan Produksi Tinggi			
Cakalang	Tuna	Tongkol	Udang
Sulawesi Utara	Sulawesi Utara	Aceh	Sumatera Utara
Maluku Utara	Maluku Utara	Jawa Timur	Kalimantan Barat
	Maluku	Maluku	Kalimantan Selatan
			Kalimantan Timur

Data ini mengindikasikan bahwa wilayah Indonesia timur dan sebagian Sumatera serta Kalimantan memiliki potensi besar dalam produksi perikanan tertentu. Selanjutnya daerah yang berproduksi sedang disajikan pada Tabel 3, sementara daerah lainnya tergolong daerah dengan produksi rendah.

Tabel 3. Wilayah Klaster Produksi Sedang

Klaster Provinsi dengan Produksi Sedang			
Cakalang	Tuna	Tongkol	Udang
Aceh	Aceh	Sumatera Utara	Jambi
Maluku Utara	DKI Jakarta	Sumatera Barat	Bengkulu
Sumatera Barat	Bali	Riau Kepulauan	Jawa Barat
DKI Jakarta	Nusa Tenggara Barat	Jawa Barat	Jawa Timur
Jawa Timur	Sulawesi Selatan	Bali	Kalimantan Tengah
Nusa Tenggara Barat	Sulawesi Tenggara	Nusa Tenggara Barat	Maluku
Nusa Tenggara Timur	Gorontalo	Nusa Tenggara Timur	Papua barat
Sulawesi Tengah	Papua Barat	Sulawesi Utara	
Sulawesi Selatan	Papua	Sulawesi Tengah	
Sulawesi Tenggara		Sulawesi Selatan	
Gorontalo		Sulawesi Tenggara	
Maluku		Maluku Utara	
Papua			

Berdasarkan data, wilayah di Indonesia yang berpotensi tinggi menghasilkan berbagai jenis ikan tuna adalah Sulawesi Utara, Maluku Utara, Maluku, Aceh, dan Jawa Timur. Untuk produksi udang, daerah yang potensial adalah provinsi Sumatera Utara, Kalimantan Barat, Kalimantan Selatan, dan Kalimantan Timur.



Gambar 4. Nilai *Silhouette Score*

Metode evaluasi yang digunakan adalah metode evaluasi internal menggunakan nilai *Silhouette Score*. Pada penelitian ini, nilai Silhouette Score relatif masih tinggi pada jumlah klaster sama dengan tiga seperti ditunjukkan pada Gambar 4, dimana kemudian nilai tersebut cenderung menurun seiring dengan pertambahan jumlah klaster.

5. SIMPULAN DAN SARAN

Penelitian yang dilakukan ini menunjukkan efektivitas metode *clustering* K-Means dikombinasikan dengan metode *Elbow* dan Skor *Silhouette* dalam mengelompokkan wilayah produksi hasil laut bernilai ekonomi tinggi di Indonesia yaitu cakalang, tuna, tongkol dan udang dengan jumlah klaster optimal yaitu tiga. Klaster-klaster ini diberi label wilayah Produksi Tinggi, Produksi Sedang, dan Produksi Rendah.

Selanjutnya untuk ikan cakalang, ikan tenggiri, tuna, dan udang, hasil pengelompokan menunjukkan adanya pengelompokan berbeda yang menyoroti wilayah dengan kapasitas produksi yang bervariasi. Daerah dengan produksi tuna yang tinggi antara lain Sulawesi Utara, Maluku Utara, Maluku, Aceh, dan Jawa Timur, sedangkan provinsi penghasil udang terkemuka adalah Sumatera Utara, Kalimantan Barat, Kalimantan Selatan, dan Kalimantan Timur.

Penelitian ke depan disarankan dapat menggabungkan data spasial (GIS) dan faktor lingkungan seperti suhu laut, salinitas, dan kedalaman perairan untuk memahami alasan di balik perbedaan produksi antar wilayah. Dengan pendekatan ini, hasil klasterisasi bisa lebih kontekstual dan aplikatif untuk kebijakan zonasi produksi perikanan. Setelah klasterisasi, dapat dilakukan prediksi tren produksi menggunakan metode seperti ARIMA, LSTM, atau *Random Forest Regression* untuk tiap klaster. Ini membantu pemerintah atau stakeholder merencanakan strategi jangka panjang berdasarkan pola yang telah terbentuk.

DAFTAR PUSTAKA

[1] T. R. Hartanto, Suharno and Burhanuddin, "Daya Saing Ekspor Ikan Tuna-Cakalang-Tongkol Indonesia Di Pasar Amerika Serikat," *JPHPI*, vol. 24, no. 2, p. 227 – 235, 2021.

[2] C. Yolandika, T. Sofyani, H. Mursyid, N. Anggraini and D. Berliana, "The competitiveness of Indonesia's frozen shrimp exports in international market," *in 11th International and National Seminar on Fisheries and Marine Science, IOP Conf. Series: Earth and Environmental Science* 1118 (2022) 012074, pp. 1-6. doi:10.1088/1755-1315/1118/1/012074., 2022.

[3] S.-L. Hung, C.-Y. Kao and J.-W. Huang, "Constrained K-means and Genetic Algorithm-based Approaches for Optimal Placement of Wireless Structural Health Monitoring Sensors," *Civil Engineering Journal*, vol. 8, no. 12, pp. 2675-2692. doi: 10.28991/CEJ-2022, 2022.

[4] Y. Liu, "Analysis and Prediction of College Students' Mental Health Based on K-means Clustering Algorithm," *Applied Mathematics and Nonlinear Sciences*, vol. 7, no. 1, p. 01–512. doi : https://doi.org/10.2478/amns.2021.1.00099., 2022.

[5] W. Wahidin, M. Mugihartadi, T. H. B. Aviani, H. Pratiwi, Y. I. Wijaya, A. Andie, A. P. Windarto and A. Waluyo, "Application of data mining techniques using the K-Means Method on Unmet Need of Health Services by Province in In in Indonesia," *Journal of Physics: Conference Series*, vol. 1783, pp. n. pag. DOI 10.1088/1742-6596/1783/1/012012., 2021.

[6] J. Singh, A. K. Darpe and S. P. Singh, "Bearing remaining useful life estimation using an adaptive data-driven model based on health state change point identification and K-means clustering," *Measurement Science and Technology*, vol. 31, no. 8, pp. n.pag. DOI 10.1088/1361-6501/ab6671, 2020.

[7] M. B. Gesicho, M. C. Were and A. Babic, "Evaluating performance of health care facilities at meeting HIV-indicator reporting requirements in Kenya: an application of K-means clustering algorithm," *BMC Med Inform Decis Mak*, vol. 21, no. 6, pp. n.pag. DOI:10.1186/s12911-020-01367-9., 2021.

[8] D. S. Rachmahwati, R. Andreswari and F. Hamami, "Customer Segmentation of Cellular Telecommunication Company Using K-Means Algorithm (Case Study of PT Indosat)," in *International Conference of Science and Information Technology in Smart Administration* (ICSINTESA) (2022): 45-50. DOI:10.1109/ICSINTESA56431.2022.10041532., 2022.

[9] A. Arifa, C. Rindu, R. Astriratma and A. Zaidiah, "K-Means Algorithm Implementation for Project Health Clustering," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, p. n. pag. DOI:10.29207/resti.v7i5.5181., 2023.

[10] S. Alshammari, M. B. Zolkepli and R. Abdullah, "Genetic Algorithm Based Parallel K-Means Data Clustering Algorithm Using MapReduce Programming Paradigm on Hadoop Environment (GAPKCA)," in *International Conference on Soft Computing and Data Mining* (2020). DOI:10.1007/978-3-030-36056-6_10, 2020.

[11] [. K. Appiah, K. Wirekoh, E. N. Aidoo, S. D. Oduro and Y. D. Arthur, "A model-based clustering of expectation–maximization and K-means algorithms in crime hotspot analysis," *Research in Mathematics*, p. n. pag. DOI:10.1080/27684830.2022.2073662., 2022.

[12] R. Mohemad, N. N. M. Muhait, N. M. M. Noor and Z. A. Othman, "Performance analysis in text clustering using k-means and k-medoids algorithms for Malay crime documents," *International Journal of Electrical and Computer Engineering* (IJECE) (2022): n. pag. DOI:10.11591/ijece.v12i5.pp5014-5026., 2022.

[13] A. H. A. A. Bolkiah, H. H. Hamzah, Z. Ibrahim, N. M. Diah, A. M. Sapawi and H. F. M. Hanum, "Crime Scene Prediction Using the Integration of K-Means Clustering and Support Vector Machine," in *IEEE 10th Conference on Systems, Process & Control (ICSPC)* : 242-246. DOI:10.1109/ICSPC55597.2022.10001768., 2022.

[14] V. Gendre, "Efficient Crime Analysis Based on Hybrid Approach by Combining Dynamic Time Wrapping Algorithm with K-Means Clustering Approach," *International Journal for Research in Applied Science and Engineering Technology*, p. n. pag. DOI:10.222, 2022.

[15] B. M. V. Costa, R. d. O. Moraes and H. M. Watanuki, "Estimating the Number of Clusters for the Algorithm of K-Means in a Context of Big Data," A*nais do Encontro Nacional de Engenharia de Produção*, p. n. pag. DOI:10.1448, 2023.

[16] L. Yang, X. Wang, Y. Li, Z. Xie, Y. Xu, R. Han and C. Wu, "Identifying Working Trajectories of the Wheat Harvester In-Field Based on K-Means Algorithm," *Agriculture*, vol. 12, no. 11, p. 1837. DOI:10.3390/agriculture12111837, 2022.

[17] Y. Si, J. Gao, D. Li and T. Song, "Study on Farmland Moisture Assessment Based on Improved K-means,," in *2nd International Conference on Electronic Information Technology and Smart Agriculture* (ICEITSA) (2022): 235-239. DOI:10.1109/ICEITSA57468.2022.00048, 2022.

[18] M. Wu, Z. Li, J. Chen, Q. Min and T. Lu., "A Dual Cluster-Head Energy-Efficient Routing Algorithm Based on Canopy Optimization and K-Means for WSN," *Sensors*, vol. 22, p. 9731. DOI:10.3390/s22249731., 2022.

[19] S. Sunori, J. Pant, H. Pant, M. Manu and P. Juneja, "Design of K-Means Clustering & SVM based Models for Soil Fertility Classification," in *3rd International Conference on Smart Electronics and Communication (ICOSEC)*, 2, 2022.

[20] A. A. Aldino, D. Darwis, A. Prastowo and C. S. 2. I. o. K.-M. A. f. C. C. P. F. A. i. S. L. R. J. P. C. S. 1. 0. DOI:10.1088/1742-6596/1751/1/012038., "Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency," *J. Phys.: Conf. Ser.* 1751 012038. DOI:10.1088/1742-6596/1751/1/012038, 2021.

[21] Eliyani, S. Dwiasnati, S. M. Arif, R. Avrizal and N. Fatimah, "Application of Machine Learning in Clustering Maize Producing Regions in Indonesia," *Computer Engineering and Applications*, vol. 13, no. 2, pp. 30-36, 2024.

[22] S. M. Rasdas, J. A. Costales, A. C. Fajardo and J. S. Limbago, "Abundance Mapping of Commercial Fisheries Production using K-Means and Forecasting Algorithm in Manila Bay, Philippines," in *IEEE 8th International Conference for Convergence in Technology (I2CT)* (2023): 1-5. DOI:10.1109/I2CT57861.2023.10126360., 2023.

[23] B. Nurdin, R. Meiyanti, A. Fahada and Marleni, "Application of the K-Means Method for Clustering Capture Fisheries Products in North Aceh with A Data Mining Approach," *Journal of Advanced Zoology*, vol. 2023, p. n. pag. DOI:10.17762/jaz.v44i4., 2023.

[24] Y. Febriani, Y. P. Sari and D. Octaria, "Metode K-Means Cluster Untuk Mengelompokkan Kota/Kabupaten di Sumatera Selatan Berdasarkan Produksi Ikan Air Tawar," *Sainmatika: Jurnal Ilmiah Matematika dan Ilmu Pengetahuan Alam*, vol. 18, no. 2, pp. 175-182, 2021.

[25] A. Firlansyah, A. A. N. Risal, F. Adiba and A. B. Kaswar, "Clustering Produksi Perikanan Budidaya Laut berdasarkan Provinsi Menggunakan Algoritma K-means," *JESSI*, vol. 2, no. 1, pp. 55-60, 2021.

[26] D. W. Sari and K. G. Ayu, "Penentuan Kriteria Dalam Memilih Sekolah Dasar Dengan Menerapkan K-Means Clustering (Studi Kasus : Wilayah Kecamatan Mampang)," *Sainstech*, vol. 29, no. 2, pp. 24-28, 2019.

[27] A. Firlansyah, A. A. N. Risal, F. Adiba, and A. B. Kaswar, "Clustering Produksi Perikanan Budidaya Laut Berdasarkan Provinsi Menggunakan Algoritma K-Means," *J. Embed. Syst. Secur. Intell. Syst.*, vol. 2, no. 1, p. 55, 2021, doi: 10.26858/jessi.v2i1.20378.

[28] H. U. Sari, A. P. Windarto, and D. Hartama, "Analisa Metode Data Mining Pada Produksi Perikanan Laut Yang Dijual Di Tempat Perikanan Ikan (Tpi)," KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer), vol. 3, no. 1, pp 630–636, 2019, doi: 10.30865/komik.v3i1.1671.

[29] R. Hablum, A. Khairan, and R. Rosihan, "Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means," JIKO (Jurnal Inform. dan Komputer), vol. 2, no. 1, pp. 26–33, 2019, doi: 10.33387/jiko.v2i1.1053.

[30] M. Syakur, B. Khotimah, E. Rochman and B. Satoto, "Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer"

[31] U. Ma'rifatin, "Implementasi Algoritma K- Means untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng," *Semin. Nas. Inov. Teknol.*, no. 2549–7952, pp. 285–291, 2020.

[32] T. M. M. Tyas and A. I. Purnamasari, "Penerapan Algoritma K-means dalam Mengelompokkan Demam Berdarah Dengue Berdasarkan Kabupaten," *Blend Sains J. Tek.*, vol. 1, no. 4, pp. 277–283, 2023, doi: 10.56211/blendsains.v1i4.231.

[33] M. Anjelita, A. P. Windarto, A. Wanto, and I. Sudahri, “Pengembangan Datamining Klastering Pada Kasus Pencemaran Lingkungan Hidup”, Seminar Nasional Teknologi Komputer Sains , pp. 309–313, 2020.

[34] I. Wahyudi, M. Sulthan and L. Suhartini, "Analisa Penentuan Cluster Terbaik Pada Metode K-Means Menggunakan Elbow Terhadap Sentra Industri Produksi Di Pamekasan," J. Apl. Teknol. Inf. dan Manaj., vol. 2, no. 2, p. 72–81, 2021..