

Analisis Pola Lalu Lintas Kapal Selat Bali Berbasis AIS, K-Means, Traffic Flow

Deny Adi Setyawan¹, Agustina Purwatiningsih², Febrian Sulistyio Budi³

¹Department of Nautical Studies, Politeknik Pelayaran Banten, Indonesia

^{2,3}Department of Informatic Engineering, Universitas Duta Bangsa Surakarta, Indonesia

¹setyawandenyadi@gmail.com, ²agustinapurwa@udb.ac.id*, ³febriansulistyio12@gmail.com

Abstract

The Bali Strait is one of the busiest sea crossing routes in Indonesia, characterized by high intensity of ship movements and dynamic traffic patterns throughout the day. These conditions require comprehensive analysis to understand the characteristics of vessel movements, identify density zones, and determine peak periods that potentially increase navigation risks. This study aims to analyze ship traffic patterns in the Bali Strait using a combination of K-Means Clustering and Traffic Flow Model based on Automatic Identification System (AIS) data. The dataset consists of 790 AIS records collected during the period of 20–26 June 2025. The research stages include data preprocessing, determination of the optimal number of clusters using the Elbow method, classification of vessel movement behavior using the K-Means algorithm, and analysis of traffic parameters comprising volume, speed, density, and traffic flow. The results reveal the formation of three main clusters: low-speed vessels concentrated around port areas, medium-speed vessels operating on main trajectories, and high-speed vessels dominating the crossing lanes. Evaluation of clustering quality using the Silhouette Coefficient produced a value of 0.3040, indicating a reasonably good level of cluster separation. Furthermore, a consistent peak hour pattern was identified at 12:00, along with two high-density zones located near Ketapang Port and Gilimanuk Port. These findings demonstrate that AIS-based analysis is capable of providing measurable representation of the dynamics of ship traffic in the Bali Strait and has the potential to support operational optimization, enhancement of navigation safety, and consideration for the implementation of a Traffic Separation Scheme (TSS)

Keywords: ship traffic, AIS, k-means, clustering, bali strait.

Abstrak

Selat Bali adalah salah satu jalur penyeberangan laut tersibuk di Indonesia dengan intensitas pergerakan kapal yang tinggi dan pola lalu lintas yang dinamis sepanjang hari. Kondisi ini memerlukan analisis yang komprehensif untuk memahami karakteristik pergerakan kapal, mengidentifikasi zona kepadatan, serta menentukan periode waktu sibuk yang berpotensi meningkatkan risiko navigasi. Penelitian ini bertujuan untuk menganalisis pola lalu lintas kapal di Selat Bali menggunakan kombinasi metode *K-Means Clustering* dan *Traffic Flow Model* berbasis data *Automatic Identification System* (AIS). Data yang digunakan terdiri atas 790 rekaman AIS yang dikumpulkan selama periode 20–26 Juni 2025. Tahapan penelitian meliputi prapemrosesan data, penentuan jumlah kluster optimal menggunakan metode *Elbow*, pengelompokan perilaku pergerakan kapal dengan algoritma *K-Means*, serta analisis parameter lalu lintas yang mencakup volume, kecepatan, kepadatan, dan aliran lalu lintas. Hasil penelitian menunjukkan terbentuknya tiga kluster utama, yaitu kapal berkecepatan rendah di sekitar area pelabuhan, kapal berkecepatan sedang pada lintasan utama, dan kapal berkecepatan tinggi yang mendominasi jalur penyeberangan. Evaluasi kualitas klusterisasi menggunakan *Silhouette Coefficient* menghasilkan nilai sebesar 0,3040, yang menunjukkan tingkat pemisahan kluster yang cukup baik. Selain itu, teridentifikasi pola jam sibuk yang konsisten pada pukul 12.00 serta dua zona kepadatan tinggi yang berlokasi di sekitar Pelabuhan Ketapang dan Pelabuhan Gilimanuk. Temuan ini menunjukkan bahwa analisis berbasis AIS mampu memberikan gambaran yang terukur mengenai dinamika lalu lintas kapal di Selat Bali, serta berpotensi mendukung optimasi operasional, peningkatan keselamatan navigasi, dan pertimbangan penerapan *Traffic Separation Scheme* (TSS).

Kata kunci: lalu lintas kapal, AIS, k-means, clustering, selat bali

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

1. Pendahuluan

Selat Bali merupakan salah satu jalur penyeberangan kapal tersibuk di Indonesia, khususnya pada rute antara Pelabuhan Ketapang dan Pelabuhan Gilimanuk [1]. Jalur ini dilalui oleh berbagai jenis kapal, mulai dari kapal penumpang, kapal barang, hingga kapal *roll-on/roll-off* (Ro-Ro) dengan tingkat kepadatan lalu lintas yang tinggi mencapai 190.820 jumlah penyeberangan kapal Ro-Ro dan 1161 kpal barang [2]. Kondisi tersebut menghasilkan dinamika pergerakan kapal yang kompleks dan bersifat dinamis. Aktivitas pelayaran di

Selat Bali memiliki peran penting dalam mendukung mobilitas masyarakat serta distribusi logistik antarwilayah; namun, pada saat yang sama juga berpotensi menimbulkan kepadatan lalu lintas, antrean kapal, dan peningkatan risiko keselamatan navigasi apabila tidak dimonitor dan dikelola secara optimal [3].

Sebagai respons terhadap tingginya intensitas lalu lintas dan risiko keselamatan pelayaran di perairan strategis Indonesia, pemerintah telah menetapkan sistem rute pelayaran melalui kebijakan nasional, antara lain dengan penerapan *Traffic Separation*

Scheme (TSS) pada selat Sunda dan Lombok guna meningkatkan keselamatan dan kelancaran navigasi kapal [4]. Kebijakan tersebut menegaskan pentingnya pengelolaan dan pemantauan lalu lintas kapal secara sistematis, terutama pada perairan dengan kepadatan tinggi dan karakteristik pergerakan kapal yang kompleks. Namun demikian, hingga saat ini penerapan TSS masih terbatas pada Selat Sunda dan Selat Lombok, dan belum diberlakukan di Selat Bali. Padahal, Selat Bali memiliki tingkat aktivitas pelayaran penyeberangan yang tinggi serta riwayat kecelakaan lalu lintas laut, sebagaimana setiap tahun terjadi kecelakaan dalam lima tahun terakhir. Kecelakaan kapal KMP Tunu Pratama Jaya pada tahun 2025 terdapat 65 korban. Kondisi tersebut mengindikasikan perlunya evaluasi dan pertimbangan penerapan TSS di Selat Bali sebagai upaya peningkatan keselamatan pelayaran.

Penelitian mengenai keselamatan navigasi dan risiko kecelakaan di Selat Bali telah dilakukan menggunakan data kecelakaan periode 2010-2018 serta pemodelan konsekuensi tumpahan minyak [5]. Hasil riset tersebut menunjukkan perlunya evaluasi lanjutan berbasis data yang lebih mutakhir seperti AIS untuk memperoleh gambaran terukur mengenai kepadatan dan perilaku kapal, sehingga dapat mendukung pertimbangan implementasi TSS dan strategi mitigasi keselamatan pelayaran di masa mendatang.

Di tengah tengah tantangan ini, timbul kebutuhan akan pemantauan yang lebih cerdas, Automatic Identification System (AIS). AIS muncul sebagai solusi sebagai sumber data yang sangat kaya. Tidak hanya menyimpan posisi kapal secara berkala, tetapi juga merekam kecepatan, arah dan karakteristik kapal. AIS telah digunakan secara global untuk memahami dan mengolah pola pergerakan kapal, mendeteksi jalur dominan, hingga menilai risiko navigasi [6]. Namun, besarnya volume data AIS menuntut pendekatan analisis yang mampu menyaring pola-pola penting dari data yang sangat padat dan dinamis.

Penelitian-penelitian sebelumnya menunjukkan bahwa teknik clustering, seperti *K-Means* sangat efektif dalam mengelompokkan pola operasional kapal berdasarkan kecepatan, ukuran, maupun perilaku perjalanan [6] [7]. Bukan hanya itu, clustering juga mampu mengungkap pola lintasan kapal, mendeteksi kepadatan area, dan mampu membantu proses pengambilan keputusan di sektor maritim [8] [9]. Hal ini menunjukkan adanya peluang untuk mengadaptasikan metode pendekatan clustering ke dalam konteks lalu lintas Selat Bali sebagai salah satu rute lalu lintas penyeberangan vital di Indonesia.

Selain itu, hasil penelitian lain juga memiliki temuan bahwa metode clustering berbasis kepadatan seperti DBSCAN bisa mengenali pola pergerakan kapal yang lebih kompleks dan non-linear [10]. Selain dari aspek clustering, analisa tentang AIS juga sudah diperluas ke deteksi anomali lintasan menggunakan pendekatan

deep learning seperti LSTM dan DBSCAN [11], yang menegaskan seberapa luasnya potensi analisis perilaku kapal.

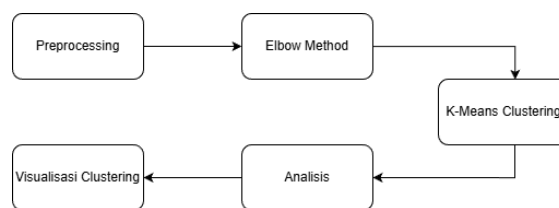
Beberapa pengembangan lain dari algoritma *K-Means* telah diadopsi ke dalam konteks big data maritim. Pendekatan seperti *MapReduce* terbukti dapat meningkatkan efisiensi dalam pengolahan data AIS berskala besar [12], sementara metode lainnya yaitu *incremental online K-Means* menawarkan pembaruan cluster secara *real-time* dan adaptif tanpa perlu melatih ulang model dari awal [13]. Ke dua pendekatan ini menunjukkan bahwa *K-Means* memiliki ruang pengembangan yang luas ketika diterapkan pada analisis pergerakan kapal modern.

Selain dengan *clustering*, analisis menggunakan konsep analisa dari dunia transportasi darat yaitu *Traffic Flow Model*. Sebuah pendekatan yang sering digunakan untuk memahami hubungan antara volume kendaraan, kecepatan, dan kepadatan jalan. Ketika model ini diterapkan pada lalu lintas laut, berpotensi menggambarkan ritme pergerakan kapal secara kuantitatif, mirip sebagaimana jalan tol dianalisis [14] [15]. Hal ini membuka sebuah peluang baru untuk memetakan waktu puncak (*peak hours*), mengidentifikasi zona padat, dan memahami bagaimana hubungan antara kecepatan dan kepadatan kapal di suatu daerah atau area.

Penelitian ini hadir sebagai jawaban dari kebutuhan praktis dan inspirasi ilmiah tersebut. Penelitian ini memiliki tujuan yang sederhana, yaitu memahami cerita di balik pergerakan kapal. Bukan hanya melihat koordinat kapal yang sedang bergerak, namun juga membaca pola, kecenderungan, dan ritme aktivitas laut. Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan, baik di sektor akademis maupun sampai di sektor operasional seperti otoritas pelabuhan dan pemangku kebijakan [16].

2. Metode Penelitian

Penelitian ini dilakukan melalui beberapa tahapan untuk menghasilkan sebuah analisis pola lalu lintas kapal pada rute pelabuhan Ketapang ke pelabuhan Gilimanuk, dimulai dari pengumpulan dataset AIS, pra-pemrosesan data, pemilihan jumlah cluster, hingga proses clustering dengan algoritma *K-Means*. Langkah penelitian dapat terlihat pada gambar 1 berikut.



Gambar 1. Langkah Penelitian

2.1 Dataset

Dataset yang digunakan dalam penelitian ini berisi 790 rekaman pergerakan kapal selama tujuh hari di wilayah

Selat Bali, khususnya rute pelabuhan Ketapang ke pelabuhan Gilimanuk. Data diambil pada periode 20 - 26 Juni 2025 yang merepresentasikan kondisi lalu lintas kapal dalam rentang waktu satu minggu pada musim kemarau. Pemilihan periode ini bertujuan untuk memperoleh gambaran pola lalu lintas kapal pada kondisi operasional yang relatif stabil. Meskipun demikian, pada tanggal 25 Juni 2025 tercatat adanya gangguan operasional berupa antrean kapal yang cukup tinggi karena adanya ombak tinggi, sehingga variasi kepadatan lalu lintas tetap dapat diamati dalam dataset yang digunakan.

Data AIS yang digunakan mencakup beberapa atribut utama yang berkaitan dengan pergerakan dan karakteristik kapal. Atribut kecepatan kapal (*Speed Over Ground/SOG*) dinyatakan dalam satuan knot, sedangkan arah pergerakan kapal (*Course Over Ground/COG*) dinyatakan dalam satuan derajat dengan rentang nilai 0–360°. Posisi kapal direpresentasikan oleh koordinat *latitude* dan *longitude* dalam satuan derajat desimal, sementara panjang kapal (*Vessel Length*) dinyatakan dalam satuan meter. Kejelasan tipe data dan satuan pada setiap atribut ini menjadi dasar penting dalam tahap pra-pemrosesan, khususnya pada proses normalisasi data sebelum dilakukan analisis cluster.

Secara keseluruhan, jumlah data mentah yang diperoleh pada tahap awal pengumpulan data adalah sebanyak 791 *record*. Setelah melalui proses *filtering* dan pembersihan data, termasuk penghapusan data duplikat dan data yang tidak lengkap, jumlah data yang digunakan dalam tahap analisis selanjutnya menjadi 790 *record*. Jumlah data tersebut dinilai cukup representatif untuk menggambarkan pola lalu lintas kapal di wilayah Selat Bali selama periode pengamatan.

Data AIS yang digunakan juga memiliki karakteristik serupa dengan yang umum dipakai dalam penelitian analisis lintasan dan perilaku kapal di seluruh dunia [6] [17]. Dengan cakupan area yang relatif kecil namun memiliki intensitas pergerakan yang tinggi, *dataset* ini ideal untuk digunakan dalam menganalisis kepadatan, pola waktu sibuk, serta clustering perilaku kapal.

Selain itu, AIS juga dimanfaatkan untuk menilai performa pelabuhan yang telah dilakukan dalam beberapa penelitian *modern*, contohnya melalui konstruksi indikator port performance berbasis segmentasi trajektori AIS [18]. Hal ini membuktikan bahwa data AIS memiliki cakupan analisis yang sangat fleksibel, mulai dari perilaku kapal hingga performa operasional pelabuhan.

2.2 Preprocessing Data

Tahapan *preprocessing* data dilakukan untuk memastikan kualitas data AIS yang digunakan dalam penelitian berada pada kondisi yang layak untuk dianalisis. Proses ini mencakup pemeriksaan kelengkapan data, validasi rentang koordinat, penanganan nilai tidak valid, serta normalisasi data

numerik. Kualitas data pada tahap ini sangat berpengaruh terhadap akurasi metode klasterisasi, karena data yang mengandung noise atau ketidakwajaran dapat menurunkan performa algoritma *K-Means*, sebagaimana ditekankan dalam penelitian sejenis berbasis data AIS [7] [17].

Adapun tahapan *preprocessing* yang dilakukan dalam penelitian ini, meliputi:

a. *Filtering area*, *filtering area* dilakukan untuk membatasi data AIS agar hanya mencakup wilayah perairan Selat Bali. Data kapal diseleksi berdasarkan rentang koordinat geografis, yaitu *latitude* -8,164 hingga -8,131 dan *longitude* 114,402 hingga 114,438. Pembatasan wilayah ini bertujuan untuk memastikan bahwa data yang dianalisis benar-benar merepresentasikan pergerakan kapal di Selat Bali, sehingga data dari perairan di luar wilayah penelitian tidak ikut terolah dan memengaruhi hasil analisis.

b. *Handling missing value*, penanganan missing values dilakukan dengan memeriksa keberadaan nilai kosong maupun nilai tidak *valid* pada data AIS. Baris data yang memiliki nilai kritis tidak lengkap atau tidak wajar, terutama pada kolom *latitude*, *longitude*, *speed*, dan MMSI, dihapus dari dataset. Selain itu, data dengan koordinat tidak valid, seperti nilai 0,0, juga dianggap sebagai *error* dan dikeluarkan. Langkah ini penting untuk mencegah *distorsi* hasil perhitungan jarak, kecepatan, dan proses klasterisasi.

c. *Data redundancy*, data AIS pada dataset ini direkam dalam interval waktu setiap 3 jam, sehingga frekuensi pengiriman data relatif jarang dibandingkan data AIS *real-time* yang umumnya dikirim setiap beberapa detik. Oleh karena itu, pada penelitian ini tidak dilakukan proses *resampling* waktu secara khusus untuk mengurangi beban komputasi. Data duplikat yang muncul pada waktu dan identitas kapal yang sama tetap diperiksa dan dihapus, sehingga setiap *record* merepresentasikan satu kondisi pergerakan kapal pada waktu tertentu.

d. Penyaringan nilai ekstrim (*outliers*), penyaringan nilai ekstrem dilakukan pada atribut kecepatan kapal untuk menghilangkan nilai yang tidak realistis. Contohnya adalah data dengan kecepatan 0 knot di tengah laut atau nilai kecepatan yang tidak sesuai dengan karakteristik operasional kapal. Namun demikian, data kapal dengan kecepatan sangat rendah atau 0 knot yang terjadi di sekitar area pelabuhan tetap dipertahankan pada tahap klasterisasi untuk menjaga kelengkapan pola pergerakan. Pendekatan ini mengikuti metode *preprocessing* yang digunakan pada penelitian AIS terdahulu [19] [9].

e. Normalisasi data numerik, tahap akhir *preprocessing* adalah normalisasi data numerik untuk menyamakan skala antar fitur sebelum dilakukan klasterisasi menggunakan algoritma *K-Means*. Normalisasi dilakukan agar atribut dengan rentang nilai besar tidak mendominasi perhitungan jarak, sehingga hasil cluster yang diperoleh lebih seimbang dan representatif terhadap karakteristik data AIS.

Melalui rangkaian tahapan *preprocessing* tersebut, dataset yang digunakan telah disesuaikan untuk mendukung proses klusterisasi dan analisis lalu lintas kapal secara optimal.

2.3 Penentuan Jumlah Cluster dengan Metode Elbow

Penentuan jumlah cluster yang optimal merupakan tahap penting dalam proses klusterisasi, karena jumlah cluster akan memengaruhi kemampuan model dalam merepresentasikan pola pergerakan kapal di wilayah penelitian. Pada penelitian ini, metode *Elbow* digunakan untuk menentukan jumlah cluster yang paling sesuai berdasarkan karakteristik data AIS di Selat Bali.

Metode Elbow mengevaluasi nilai *Within-Cluster Sum of Squares* (WCSS) untuk berbagai jumlah cluster (k). Nilai WCSS merepresentasikan total kuadrat jarak antara setiap titik data dengan pusat cluster (*centroid*) yang menaunginya, sehingga mencerminkan tingkat homogenitas data dalam setiap cluster. Secara matematis, WCSS dapat dirumuskan pada persamaan 1 berikut:

$$WCSS = \sum_{j=1}^k \sum_{x_i \in C_j} |x_i - \mu_j|^2 \quad (1)$$

Dimana k adalah jumlah cluster, C_j adalah cluster ke- j , X_i adalah titik data dalam cluster, dan μ_j adalah centroid dari cluster tersebut.

Dalam konteks lalu lintas laut, klusterisasi tidak merepresentasikan jalur yang bersifat statis seperti pada jaringan jalan darat, melainkan mengelompokkan pola pergerakan kapal berdasarkan kemiripan karakteristik spasial dan kinematik, seperti posisi geografis dan kecepatan. Oleh karena itu, nilai WCSS pada penelitian ini dihitung berdasarkan jarak antar fitur yang telah dinormalisasi, sehingga mencerminkan kedekatan pola pergerakan kapal dalam suatu area pengamatan di Selat Bali, tanpa mendefinisikan lebar jalur atau penampang lalu lintas secara eksplisit.

Seiring bertambahnya jumlah cluster, nilai WCSS akan mengalami penurunan karena data terbagi ke dalam kelompok yang semakin spesifik. Namun, pada titik tertentu, laju penurunan WCSS mulai melambat dan membentuk pola menyerupai siku (*elbow*). Titik ini menunjukkan jumlah cluster yang paling optimal karena mampu merepresentasikan struktur alami data tanpa menghasilkan pembagian cluster yang berlebihan (*overfitting*). Pendekatan ini telah banyak digunakan dalam penelitian klusterisasi pergerakan kapal berbasis AIS [7] [8].

Jumlah cluster yang diperoleh melalui metode Elbow selanjutnya digunakan sebagai dasar dalam proses klusterisasi data AIS. Hasil cluster ini kemudian diintegrasikan ke dalam analisis lanjutan, termasuk pemodelan *traffic flow*, di mana setiap cluster merepresentasikan segmen pola pergerakan kapal yang memiliki karakteristik lalu lintas berbeda di wilayah Selat Bali.

2.4 Clustering K-Means

Tahapan *clustering* dengan metode *K-Means* adalah tahapan penerapan algoritma untuk mengelompokkan data berdasarkan kedekatan jarak terhadap pusat cluster (*centroid*). Pengukuran kedekatan jarak ini dilakukan menggunakan *Euclidean Distance* antara data kapal (x) dengan *centroid* (c), seperti ditunjukkan pada persamaan 2 berikut:

$$d(x, c) = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \quad (2)$$

Algoritma *K-Means* dipilih dalam penelitian ini atas dasar beberapa pertimbangan metodologis. Metode ini terbukti efektif untuk menangani dataset berskala relatif besar seperti data *Automatic Identification System* (AIS), karena mampu mengidentifikasi struktur pola secara efisien pada kumpulan data multidimensional. Selain itu, proses iteratif *K-Means* memiliki kinerja komputasi yang cepat sehingga sesuai untuk analisis lalu lintas kapal yang memerlukan pengolahan dalam waktu singkat. Dari sisi keilmiahannya, algoritma ini juga mudah direproduksi dan diimplementasikan kembali oleh peneliti lain karena prosedurnya sederhana dan berbasis parameter yang jelas. Karakteristik tersebut menjadikan *K-Means* cocok diterapkan pada data numerik seperti parameter kecepatan kapal, dimensi panjang kapal, koordinat geografis (*latitude-longitude*), serta arah pelayaran (*course*).

Penerapan *K-Means* diarahkan untuk mengelompokkan kapal berdasarkan atribut utama meliputi *speed*, *vessel length*, *latitude*, *longitude*, dan *course*, sehingga dapat merepresentasikan perbedaan perilaku pergerakan antar-vessel di Selat Bali. Secara prosedural, algoritma bekerja dengan menginisialisasi sejumlah centroid awal sesuai jumlah kluster yang telah ditetapkan, kemudian dilanjutkan dengan serangkaian iterasi penentuan keanggotaan kluster berdasarkan jarak terdekat terhadap *centroid*. Pada setiap iterasi, posisi *centroid* baru μ_j diperbarui dengan menghitung rata-rata (*mean*) dari seluruh anggota cluster tersebut menggunakan persamaan 3 berikut:

$$\mu_j = \frac{1}{N_j} \sum_{x \in C_j} x \quad (3)$$

Dimana N_j adalah jumlah total data dalam cluster j . Proses ini berulang hingga mencapai konvergensi (posisi *centroid* tidak lagi berubah). Pendekatan *K-Means* sering digunakan dalam penelitian maritim untuk memodelkan perilaku kapal dan mengidentifikasi kelompok-kelompok operasional [7] [9] [17].

3. Hasil dan Pembahasan

3.1. Hasil

Sub bab ini menyajikan rangkaian dari hasil penelitian berdasarkan urutan proses analisis, mulai dari pembersihan data sampai identifikasi pola lalu lintas kapal di Selat Bali. Semua hasil yang diperoleh

dipaparkan secara objektif tanpa interpretasi berlebihan, dan didukung oleh tabel dan gambar untuk memperkuat pemahaman.

3.1.1. Hasil *Preprocessing* Data

Setelah dilakukan proses pembersihan data, dataset AIS berhasil disusun menjadi format yang lebih rapi dan siap diolah. Nilai *attribute* yang kosong dan anomali seperti kecepatan 0 di tengah laut berhasil dihilangkan, menghasilkan *dataset* akhir yang lebih representatif. Tahapan ini serupa dengan standar *preprocessing* yang diterapkan pada penelitian AIS sebelumnya *Cluster* untuk memastikan kualitas data sebelum analisis lanjutan [6].

No	Date	Day	Time	Vessel_Name	MMSI	Lat	Long	Vessel_Type	Vessel_Age	Vessel_Length	Speed	Course
1	2025-06-20	Friday	00:00:00	DHARMA RUCITRA	525005324	-8.160231	114.436351	Penumpang	61	48.60	5.6	139
2	2025-06-20	Friday	00:00:00	KMP SAMUDERA UTAMA	525005290	-8.159460	114.437185	Penumpang	12	65.00	1.7	134
3	2025-06-20	Friday	00:00:00	KMP KARYA MARITIM 3	525024403	-8.153642	114.432325	Penumpang	14	61.20	3.9	158
4	2025-06-20	Friday	00:00:00	KMP NUSA DUA	525004268	-8.150364	114.432216	Penumpang	53	47.90	4.4	54
5	2025-06-20	Friday	00:00:00	KMP RENY II	525002080	-8.142887	114.428441	Penumpang	57	41.44	6.0	340
6	2025-06-20	Friday	00:00:00	KMP TRISAKTI ELFINA	525010181	-8.140975	114.434789	Penumpang	18	61.10	6.4	295
7	2025-06-20	Friday	00:00:00	MARINA PRATAMA	525002076	-8.139998	114.422093	Penumpang	53	54.50	5.0	268
8	2025-06-20	Friday	00:00:00	NUSA MAKMUR	525000514	-8.144182	114.412620	Penumpang	35	47.90	3.5	123
9	2025-06-20	Friday	00:00:00	TUNJU PRATAMA J 5888	525022387	-8.145064	114.413307	Kargo	19	58.00	5.4	277
10	2025-06-20	Friday	00:00:00	SWARNA CAKRA	525002127	-8.148869	114.402571	Penumpang	37	72.00	3.7	312

Gambar 2. Sample Data Setelah *Preprocessing* Data

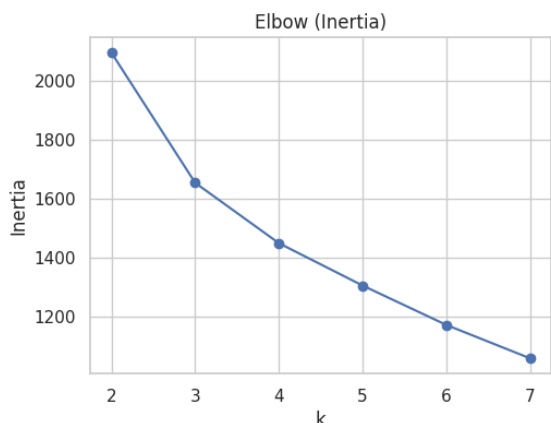
Dataset final yang digunakan mencakup:

- Data 7 hari pergerakan kapal.
- Interval 3 jam.
- 1.300 lebih titik pergerakan.
- 60 lebih kapal unik berdasarkan MMSI.

Struktur final yang bersih ini menjadi fondasi untuk analisis *clustering* dan *traffic flow*.

3.1.2. Penentuan Jumlah Cluster (*Elbow Method*)

Pendekatan dengan metode *Elbow* diterapkan pada atribut numerik (*speed*, *vessel length*, *latitude*, *longitude*). Hasil plot ini menunjukkan bahwa titik siku pada nilai $k = 3$, yang berarti tiga cluster dianggap paling representatif tanpa membuat kluster terlalu terfragmentasi. Pendekatan ini sesuai dengan pendekatan *Elbow* pada studi *clustering* maritim [6] [8].



Gambar 3. Grafik *Elbow Method*

3.1.3. Hasil *Clustering* dengan *K-Means*

Algoritma *K-Means* berhasil membentuk tiga cluster utama, yang masing-masing menggambarkan

kelompok perilaku kapal berbeda. Hasil pengelompokan ini dihasilkan dari proses iteratif hingga *centroid* konvergen.

Untuk memahami karakteristik dari setiap cluster yang terbentuk, dilakukan analisis terhadap pusat cluster (*centroid*) dengan menggunakan pendekatan nilai rata-rata dari masing-masing variabel pada setiap cluster. Pendekatan ini dipilih karena dapat memberikan gambaran yang lebih representatif terhadap kondisi umum lalu lintas kapal pada setiap kelompok cluster, khususnya dalam menggambarkan perbedaan pola kecepatan, ukuran kapal, dan persebaran posisi geografis di Selat Bali.

Tabel 1. Tabel Nilai Rata-Rata Variabel *Cluster*

Cluster	Speed	Vessel Length	Latitude	Longitude
0	2.99	56.22	-8.1574	114.4320
1	1.64	59.34	-8.1442	114.4073
2	5.16	56.99	-8.1465	114.4190

Berdasarkan Tabel 1 diatas, terlihat bahwa masing-masing cluster memiliki karakteristik lalu lintas kapal yang berbeda. 1 menunjukkan nilai kecepatan rata-rata paling rendah, yang mengindikasikan bahwa area dengan pergerakan kapal yang relatif lambat dan berpotensi merupakan zona antrean atau manuver di dekat pelabuhan. Sebaliknya, *Cluster 2* memiliki kecepatan rata-rata tertinggi, yang mencerminkan jalur pelayaran dengan arus lalu lintas lebih lancar dan memiliki aktivitas kapal yang lebih dinamis. Sementara itu, *Cluster 0* berada pada kondisi transisi dengan karakteristik kecepatan sedang, yang dapat diinterpretasikan sebagai area pergerakan kapal yang stabil namun tetap padat.

Penggunaan algoritma *K-Means* berhasil membentuk tiga cluster utama, yang masing-masing menggambarkan kelompok perilaku kapal yang berbeda. Hasil dari pengelompokan ini dihasilkan dari proses iteratif hingga *centroid* konvergen. Secara umum, karakteristik ketiga cluster ini menunjukkan pola yang distingtif. *Cluster 0* didominasi oleh kapal-kapal yang memiliki ukuran lebih besar dan beroperasi dengan kecepatan sedang. Sementara itu, *Cluster 1* merepresentasikan kelompok kapal dengan kecepatan rendah yang posisi geografisnya cenderung sering berada di dekat area pelabuhan. Sebaliknya, *Cluster 2* terdiri dari kapal-kapal yang melakukan pergerakan cepat, terutama mendominasi pada jalur penyeberangan utama di selat tersebut.

Selain dari perbedaan berdasarkan kecepatan rata-rata, masing-masing *cluster* juga menunjukkan memiliki kecenderungan unik yang berbeda. Kapal yang tergolong dalam *cluster* berkecepatan rendah lebih sering terdistribusi di sekitar area pelabuhan, yang mengindikasikan adanya aktivitas manuver, antrean, atau persiapan untuk bersandar. Sebaliknya, pada kapal di cluster kecepatan sedang dan tinggi cenderung membentuk lintasan yang lebih memanjang dan

konsisten mengikuti jalur penyeberangan utama Selat Bali.

Pola ini menunjukkan bahwa hasil *clustering* tidak hanya memisahkan kapal berdasarkan atribut numerik semata, tetapi juga secara implisit merepresentasikan fungsi operasional kapal di lapangan. Dengan demikian, pengelompokan yang dihasilkan mampu menggambarkan perilaku lalu lintas kapal secara lebih kontekstual dan mendekati kondisi nyata di perairan Selat Bali.

Pertama, pada *cluster* 0 kelompok ini merepresentasikan kapal-kapal yang sedang berada dalam fase pelayaran yang tenang dan stabil. Seperti yang terlihat pada sampel kapal Dharma Rucitra, KMP Jambo X, dan KM. Gerbang Samudra 5, kecepatan mereka terjaga konsisten di kisaran 6,5 hingga 6,8 knot. Angka ini bukan sekadar statistik, melainkan indikasi bahwa kapal-kapal berukuran besar ini sedang melintas dengan nyaman di jalur pelayaran—tidak sedang terburu-buru, namun tetap menjaga efisiensi momentumnya. Rincian sampel kapal yang mewakili ritme tenang pada cluster ini dapat dilihat pada Tabel 1 berikut.

Tabel 2. Tabel Cluster 0

Cluster	Vassel Name	Vassel Type	Speed
0	Dharma Rucitra	Penumpang	6.5
0	KMP Jambo X	Penumpang	6.5
0	KM Gerbang Samudra 5	Penumpang	6.8
...
0	KMP Trima Jaya 9	Penumpang	6.8

Berbeda dengan kelompok cluster sebelumnya, Cluster 1 menangkap "denyut nadi" kesibukan di area pelabuhan. pada cluster ini, kecepatan kapal menurun drastis menjadi hanya sekitar 4,0 hingga 4,1 knot, seperti yang ditunjukkan oleh kapal Dharma Swarna Cakra, Pancar Indah, dan KMP Trisakti Elfina. Lambatnya laju ini melukiskan kehati-hatian kapal saat harus bermanuver di ruang sempit, antre untuk bersandar, atau baru saja lepas tali. Ini adalah potret dari fase kritis di mana kewaspadaan navigasi jauh lebih diutamakan dibandingkan kecepatan. Data sampel yang menggambarkan aktivitas ini disajikan dalam Tabel 2 berikut.

Tabel 3. Tabel Cluster 1

Cluster	Vassel Name	Vassel Type	Speed
1	Dharma Swarna Cakra	Penumpang	4.0
1	Pancar Indah	Penumpang	4.0
1	Dharma Rucitra	Penumpang	4.1
...
1	KMP Trisakti Elfina	Penumpang	4.1

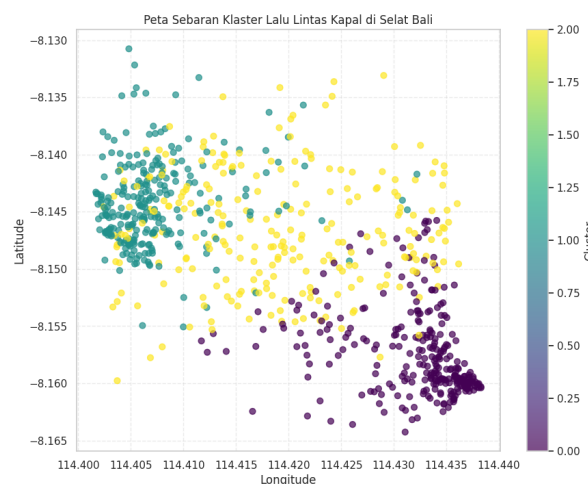
Terakhir, Cluster 2 menyajikan dinamika yang jauh lebih energik dan bebas. Kelompok ini dihuni oleh kapal-kapal yang sedang memacu mesinnya di perairan terbuka, seperti Agung Samudra XVIII, Trans Jawa 9, dan Gilimanuk. Dengan kecepatan menembus angka

8,1 hingga 9,1 knot, kapal-kapal ini merepresentasikan semangat efisiensi waktu di jalur penyeberangan utama (*crossing*). Tingginya kecepatan rata-rata ini menandakan bahwa kapal telah lepas dari zona manuver pelabuhan dan kini fokus mencapai tujuan secepat mungkin. Karakteristik kapal-kapal cepat ini terangkum dalam Tabel 3 berikut.

Tabel 4. Tabel Cluster 2

Cluster	Vassel Name	Vassel Type	Speed
2	Agung Samudra XVIII	Penumpang	8.1
2	Trans Jawa 9	Kargo	8.1
2	KMP Prathita IV	Penumpang	8.3
...
2	Gilimanuk	Penumpang	9.1

Visualisasi ini selaras dengan pendekatan *clustering* perilaku kapal yang dilakukan pada penelitian AIS global [9] [13].



Gambar 4. Visualisasi Cluster Berdasarkan Lokasi Kapal

Gambar 4 menunjukkan peta persebaran hasil klasterisasi lalu lintas kapal di Selat Bali, di mana setiap titik merepresentasikan posisi kapal yang diwarnai sesuai dengan *cluster* yang terbentuk. Terlihat bahwa *cluster* dengan kecepatan rendah cenderung terkonsentrasi pada wilayah tertentu di bagian barat dan tenggara area pengamatan, yang mengindikasikan zona dengan aktivitas manuver kapal yang tinggi, seperti area mendekati pelabuhan atau jalur antrean. Sebaliknya, cluster dengan kecepatan lebih tinggi menunjukkan sebaran yang lebih luas dan relatif memanjang mengikuti arah jalur pelayaran utama, mencerminkan kondisi arus lalu lintas yang lebih lancar. Sementara itu, cluster dengan karakteristik kecepatan sedang tampak berada di antara kedua pola tersebut dan berperan sebagai zona transisi pergerakan kapal. Pola spasial ini menegaskan bahwa hasil klasterisasi tidak hanya membedakan kapal berdasarkan atribut numerik, tetapi juga mampu merepresentasikan dinamika lalu lintas kapal secara geografis di Selat Bali.

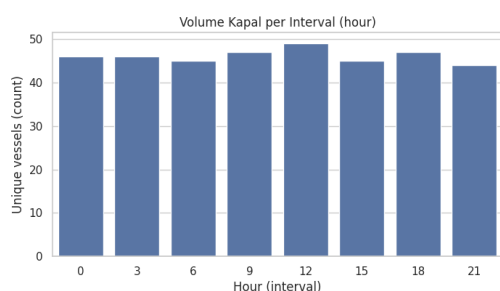
Berdasarkan hasil statistik kecepatan pada masing-masing *cluster*, dapat terlihat perbedaan pola

pergerakan kapal di Selat Bali. *Cluster 2* memiliki kecepatan rata-rata tertinggi sebesar 5,16 knot, dengan rentang kecepatan 2,2 hingga 9,1 knot, yang menunjukkan karakteristik jalur pelayaran dengan arus lalu lintas yang relatif lancar. Sebaliknya, *Cluster 1* memiliki kecepatan rata-rata terendah sebesar 1,64 knot, dengan kecepatan minimum mencapai 0,1 knot, yang mengindikasikan area dengan dominasi pergerakan lambat dan aktivitas manuver kapal yang tinggi, khususnya di sekitar pelabuhan. Sementara itu, *Cluster 0* menunjukkan kecepatan rata-rata sebesar 2,99 knot, dengan rentang kecepatan 0,2 hingga 6,8 knot, yang berada di antara kedua cluster tersebut dan merepresentasikan zona transisi pergerakan kapal.

Sebagai penegasan dan evaluasi terhadap hasil analisis karakteristik *cluster* dan pola persebaran yang telah dibahas sebelumnya, proses evaluasi kualitas klusterisasi dilakukan dengan menggunakan *Silhouette Coefficient*. Nilai *Silhouette* yang diperoleh sebesar 0,3040 menunjukkan bahwa pembentukan cluster dengan metode *K-Means* pada penelitian ini berada pada tingkat pemisahan yang cukup baik. Nilai tersebut mengindikasikan bahwa data kapal secara umum lebih dekat dengan *cluster* tempatnya berada dibandingkan dengan *cluster* lainnya, meskipun masih terdapat tumpang tindih antar cluster. Kondisi ini sejalan dengan karakteristik lalu lintas kapal di Selat Bali yang bersifat dinamis dan memiliki area pergerakan yang saling beririsan, sebagaimana tercermin pada hasil visualisasi persebaran *cluster*. Dengan demikian, hasil klusterisasi yang diperoleh dapat dinilai mampu merepresentasikan pola lalu lintas kapal di Selat Bali secara realistis dan dapat digunakan sebagai dasar analisis lanjutan.

3.1.4. Hasil Analisis Traffic Volume dan Peak Hours

Volume lalu lintas kapal dihitung berdasarkan jumlah kapal setiap interval 3 jam. Hasil pengolahan menunjukkan adanya fluktuasi signifikan pada jam-jam tertentu.



Gambar 5. Volume Kapal per Interval Waktu

Hasil dari pengolahan menunjukkan:

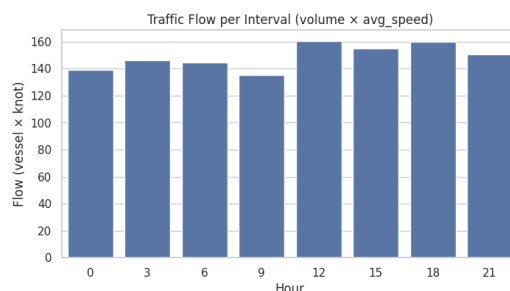
- Volume tertinggi terjadi pada pukul 12.00.
- Volume terendah pada pukul 06.00.
- Polanya konsisten selama 7 hari.

Pola ini seperti temuan trafik harian pada studi densitas kapal berskala besar [14], di mana pergerakan cenderung meningkat ketika pagi sampai siang hari.

Peningkatan *volume* kapal pada jam-jam tertentu menunjukkan adanya pola operasional yang terstruktur dan berulang, bukan bersifat acak. Kondisi ini mengindikasikan bahwa pergerakan kapal di Selat Bali sangat dipengaruhi oleh jadwal penyeberangan dan aktivitas pelabuhan. Dengan demikian, *volume* lalu lintas dapat dipandang sebagai representasi langsung dari intensitas aktivitas transportasi laut yang terjadi pada interval waktu tertentu.

3.1.5. Hasil Analisis Kecepatan Rata-rata Kapal

Kecepatan rata-rata setiap interval waktu memperlihatkan dinamika yang menarik, terutama pada jam-jam dengan *volume* tinggi.



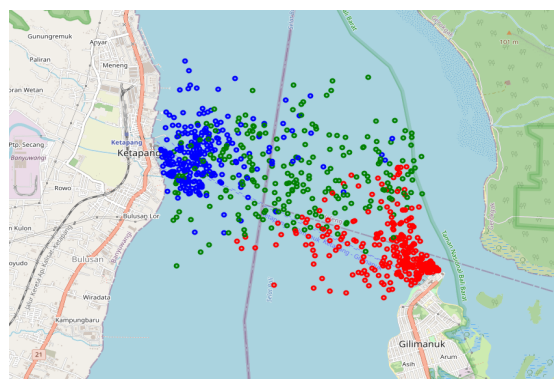
Gambar 6. Kecepatan Rata-rata Kapal per Interval Waktu

Meskipun *volume* meningkat pada jam sibuk, kecepatan tidak menurun drastis, menandakan bahwa rute Ketapang–Gilimanuk memiliki arus lalu lintas yang relatif stabil. Temuan serupa juga muncul pada studi trafik laut di Selat Malaka [14].

Tidak terjadinya penurunan kecepatan secara drastis meskipun *volume* kapal meningkat menunjukkan bahwa Selat Bali masih memiliki kapasitas perairan yang memadai untuk menampung kepadatan lalu lintas. Hal ini menjadi pembeda utama antara lalu lintas laut dan lalu lintas darat, di mana keterbatasan ruang sering kali menyebabkan penurunan kecepatan yang signifikan pada kondisi padat.

3.1.6. Hasil Analisis Kepadatan Kapal

Kepadatan dihitung berdasarkan jumlah kapal yang berada dalam area tertentu (*grid bounding box*). Visualisasi koordinat menunjukkan titik-titik konsentrasi pergerakan kapal.



Gambar 7. Heatmap Kepadatan Kapal di Selat Bali

Hasil memperlihatkan dua zona padat:

- a. Dekat Pelabuhan Ketapang,
- b. Dekat Pelabuhan Gilimanuk.

Zona kepadatan yang teridentifikasi di sekitar Pelabuhan Ketapang dan Gilimanuk mencerminkan titik konsentrasi aktivitas keluar-masuk kapal. Kondisi ini berpotensi meningkatkan kompleksitas navigasi, terutama ketika kapal dengan karakteristik dan kecepatan yang berbeda berada pada ruang perairan yang relatif terbatas. Oleh karena itu, informasi kepadatan berbasis AIS memiliki peran penting sebagai dasar pengambilan keputusan dalam pengelolaan lalu lintas dan keselamatan pelayaran.

Zona padat ini mencerminkan pola yang ditemukan juga pada penelitian densitas ruang-waktu [15].

3.2. Pembahasan

Sub Bab ini membahas tentang hubungan antara hasil yang diperoleh, dan menjawab pertanyaan penelitian: bagaimana pola lalu lintas kapal di Selat Bali dan bagaimana clustering serta traffic flow membantu memahami hal tersebut.

3.2.1. Interpretasi *Clustering* Perilaku Kapal

Ketiga cluster yang terbentuk dari algoritma K-Means menggambarkan tiga kelompok perilaku dominan:

- a. Kapal bergerak lambat di dekat pelabuhan, cenderung melakukan manuver, bongkar muat, atau persiapan sandar.
- b. Kapal bergerak sedang dalam perjalanan, kecepatan stabil di jalur utama.
- c. Kapal bergerak cepat, kemungkinan kapal yang mengejar jadwal atau memiliki spesifikasi mesin lebih besar.

Keberadaan ketiga cluster tersebut menegaskan bahwa lalu lintas kapal di Selat Bali bersifat heterogen dan dinamis. Interaksi antara kapal berkecepatan rendah dan tinggi pada jalur yang sama menuntut pengelolaan lalu lintas yang adaptif, terutama pada periode dengan volume tinggi. Dalam konteks ini, hasil clustering dapat digunakan sebagai dasar awal untuk memahami potensi konflik lintasan dan kebutuhan pengaturan navigasi yang lebih terarah.

Kelompok cluster ini sangat sejalan dengan klasifikasi perilaku kapal pada penelitian *K-Means* maritim global [7] [17]. Kluster 1 dan 2 memperlihatkan pola mendekati pelabuhan, sedangkan Kluster 3 menegaskan adanya jalur pergerakan utama yang konsisten.

3.2.2. Hubungan Volume, Kecepatan, dan Kepadatan

Hasil dari pengolahan dataset menunjukkan:

- a. Volume tinggi pada pukul 12.00.
- b. Kepadatan tertinggi juga berada pada jam dan area yang sama.
- c. Namun kecepatan rata-rata tidak turun ekstrem.

Hal ini membuktikan bahwa Selat Bali memiliki lalu lintas yang lancar, meskipun lalu lintasnya padat. Karakteristik yang juga ditemukan pada perairan dengan ruang gerak terbuka seperti di penelitian [14].

Jika diterapkan konsep *traffic flow*:

$$Flow = Volume \times Kecepatan$$

Penerapan konsep *traffic flow* dalam penelitian ini menunjukkan bahwa pendekatan yang umum digunakan pada transportasi darat dapat diadaptasi secara efektif ke dalam konteks transportasi laut. Meskipun lingkungan perairan memiliki karakteristik fisik yang berbeda, hubungan dasar antara *volume*, kecepatan, dan kepadatan tetap relevan untuk menggambarkan dinamika lalu lintas kapal secara kuantitatif.

Maka jam dengan *flow* tertinggi berada pada 12.00 dan 18.00, menegaskan bahwa aktivitas kapal memang mencapai titik puncak pada periode tersebut.

3.2.3. Identifikasi *Peak Hours* dan Implikasinya

Peak hours yang konsisten tiap hari menunjukkan bahwa pergerakan lalu lintas kapal di Selat Bali mengikuti ritme operasional yang stabil, jauh dari pola acak. Temuan ini sangat penting untuk:

- a. Penjadwalan kapal penyeberangan antar pelabuhan.
- b. Pemetaan risiko dan keselamatan.
- c. Optimasi jalur masuk dan keluar pelabuhan.
- d. Pengaturan petugas navigasi dan VTS.

Informasi mengenai peak hours ini juga memberikan nilai praktis bagi perencanaan operasional pelabuhan, khususnya dalam pengaturan slot sandar dan keberangkatan kapal. Dengan mengetahui periode aktivitas tertinggi, pengelola pelabuhan dapat mengalokasikan sumber daya secara lebih efisien dan meminimalkan potensi keterlambatan maupun antrean kapal.

Pembahasan ini sejalan dengan rekomendasi analisis temporal pada penelitian maritim berbasis AIS [20].

3.2.4. Keterkaitan Hasil dengan Studi Terdahulu

Hasil penelitian ini memiliki kedekatan dengan beberapa studi:

- a. Pola *cluster* mirip studi Park & Choi (2022) dan Han et al. (2021).
- b. Pola densitas mirip penelitian densitas spasial-temporal Kim et al. (2022).
- c. Pola peak hours sesuai dinamika trafik pelabuhan pada studi Liu (2024).

Konsistensi ini memperkuat validitas metodologi yang digunakan. Hasil dari penelitian ini juga selaras dengan tren literatur baru yang menunjukkan bahwa efisiensi pengolahan data AIS dapat ditingkatkan melalui pendekatan komputasi terdistribusi, seperti *MapReduce* pada *K-Means* [9] [12], serta metode incremental K-Means yang memungkinkan pembaruan kluster secara dinamis ketika data baru masuk [13]. Kedua

pendekatan tersebut mendukung temuan bahwa pola lalu lintas kapal dapat ditangkap secara lebih responsif dan efisien.

Keselarasan hasil penelitian ini dengan berbagai studi terdahulu memperkuat validitas pendekatan yang digunakan. Hal ini menunjukkan bahwa kombinasi metode *K-Means* dan *traffic flow* model merupakan pendekatan yang relevan dan memiliki potensi untuk diterapkan pada wilayah perairan lain dengan karakteristik lalu lintas yang serupa.

4. Kesimpulan

Penelitian ini menunjukkan bahwa pergerakan lalu lintas kapal di Selat Bali memiliki pola yang teratur dan tidak bersifat acak. Melalui penerapan metode *clustering K-Means*, berhasil mengidentifikasi tiga kelompok utama. Untuk menentukan jumlah *cluster* digunakan metode *Elbow* untuk atribut numerik (*speed, vessel length, latitude, longitude*). Hasil plot ini menunjukkan bahwa titik siku pada nilai $k = 3$, yang berarti tiga *cluster* dianggap paling representatif tanpa membuat kluster terlalu terfragmentasi. Dalam ketiga *cluster* terdapat segmentasi perilaku kapal, yaitu kapal dengan kecepatan rendah yang umumnya berada di sekitar area pelabuhan, kapal dengan kecepatan stabil yang beroperasi pada jalur pelayaran utama, serta kapal berkecepatan tinggi yang mendominasi lintasan penyeberangan. Temuan ini menunjukkan bahwa analisis berbasis AIS mampu memberikan gambaran yang terukur mengenai dinamika lalu lintas kapal di Selat Bali, serta berpotensi mendukung optimasi operasional, peningkatan keselamatan navigasi, dan pertimbangan penerapan *Traffic Separation Scheme* (TSS).

Lebih lanjut, analisis terhadap *traffic volume* dan *traffic flow* menunjukkan adanya pola jam puncak yang konsisten pada pukul 12.00, dengan tingkat kepadatan tertinggi terpusat di sekitar Pelabuhan Ketapang dan Gilimanuk. Hal ini membuktikan bahwa dinamika lalu lintas kapal dapat dipetakan secara jelas dan memiliki potensi besar untuk dimanfaatkan dalam pengaturan jadwal operasional, pengawasan risiko, serta peningkatan efisiensi operasional layanan penyeberangan.

Selain itu, penelitian ini membuka peluang pengembangan lebih lanjut, seperti analisis musiman, prediksi *Estimated Time of Arrival* (ETA), deteksi anomali, maupun penerapan algoritma *clustering* alternatif guna memperoleh pemahaman yang lebih komprehensif terhadap perilaku kapal. Dengan demikian, penelitian ini tidak hanya menjawab tujuan awal penelitian, tetapi juga menyediakan dasar yang kuat bagi pengambilan keputusan yang lebih adaptif dan inovatif dalam pengelolaan lalu lintas laut di Selat Bali.

Daftar Rujukan

- [1] Badan Pusat Statistik, Statistik Transportasi Laut Indonesia. Jakarta: Badan Pusat Statistik, 2022.
- [2] Kementerian Perhubungan Republik Indonesia, "Data Lalu Lintas Penyeberangan Ketapang-Gilimanuk," Jakarta, 2021.
- [3] M. Vespe, L. Alessandrini, F. Tarchi, and P. Braca, "Evaluation of AIS Data Integrity by Exploiting Vessel Movement Statistics," IEEE Trans Aerosp Electron Syst, vol. 51, no. 1, pp. 1–14, 2015.
- [4] Kementerian Perhubungan Republik Indonesia, Keputusan Menteri Perhubungan Nomor KM 129 Tahun 2020 tentang Penetapan Sistem Rute di Selat Lombok; dan Keputusan Menteri Perhubungan Nomor KM 130 Tahun 2020 tentang Penetapan Sistem Rute di Selat Sunda, Jakarta, 2020.
- [5] D. Laut dan Pesisir Dalam Peningkatan Daya Saing Indonesia, A. Risiko Tubrukan Kapal Dan Pemodelan Persebaran, D. Aji Purnomo, A. Dinariyana, and K. Buda Artana, "Seminar Nasional Kelautan XIV " Implementasi Hasil Riset Sumber C5-1," 2019.
- [6] J. Park and M. Choi, "A K-means clustering algorithm to determine representative operational profiles of a ship using AIS data," J Mar Sci Eng, vol. 10, no. 9, p. 1245, 2022.
- [7] H. Feng, M. Grifoll, Z. Yang, and P. Zheng, "Collision risk assessment for ships' routing waters: An information entropy approach with Automatic Identification System (AIS) data," Ocean Coast Manag, vol. 224, p. 106184, 2022.
- [8] X. Han, C. Armenakis, and M. Jadidi, "Modeling vessel behaviours by clustering AIS data using optimized DBSCAN," Sustainability, vol. 13, no. 15, p. 8162, 2021.
- [9] T.-A. Pham, X.-K. Dang, and N.-S. Vo, "Optimising maritime big data by K-means Clustering with Mapreduce Model," in International Conference on Industrial Networks and Intelligent Systems, Springer, 2022, pp. 136–151.
- [10] D. A. Napitupulu, "Prediksi Anomali Lintasan Kapal Berdasarkan Data AIS Menggunakan LSTM dan DBSCAN," Institut Teknologi Bandung, 2023.
- [11] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhajja, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," Inf Sci (N Y), vol. 622, pp. 178–210, 2023.
- [12] A. Abernathy and M. E. Celebi, "The incremental online k-means clustering algorithm and its application to color quantization," Expert Syst Appl, vol. 207, p. 117927, 2022.
- [13] L. Wang, P. Chen, L. Chen, and J. Mou, "Ship AIS trajectory clustering: An HDBSCAN-based approach," J Mar Sci Eng, vol. 9, no. 6, p. 566, 2021.
- [14] D. Liu, Z. Liu, H.-S. Kang, C.-L. Siow, and C. G. Soares, "Traffic complexity assessment on the Malacca Strait with traffic zone matrix based on AIS data," Ocean Engineering, vol. 314, p. 119687, 2024.
- [15] Y.-J. Kim et al., "Maritime traffic evaluation using spatial-temporal density analysis based on big AIS data," Applied Sciences, vol. 12, no. 21, p. 11246, 2022.
- [16] J. R. K. Bokau, R. Samad, Y. Park, and D. Kim, "From managing risk to reality: A case of maritime safety in Makassar Port, Indonesia using FRAM and AIS data analysis," Ocean Engineering, vol. 339, p. 122012, 2025.
- [17] W. M. Wijaya and Y. Nakamura, "Port performance indicators construction based on the AIS-generated trajectory segmentation and classification," Int J Data Sci Anal, vol. 20, no. 3, pp. 2473–2492, 2025.
- [18] J.-S. Lee, H.-T. Lee, and I.-S. Cho, "Maritime traffic route detection framework based on statistical density analysis from AIS data using a clustering algorithm," IEEE Access, vol. 10, pp. 23355–23366, 2022.
- [19] M. Mieczysńska and I. Czarnowski, "K-means clustering for SAT-AIS data analysis," WMU Journal of Maritime Affairs, vol. 20, no. 3, pp. 377–400, 2021.
- [20] D. Liu, H. Rong, and C. G. Soares, "Shipping route modelling of AIS maritime traffic data at the approach to ports," Ocean Engineering, vol. 289, p. 115868, 2023.