



Perbandingan K-Means dan DBSCAN dalam Analisis Pola Pergerakan Kapal Menggunakan Data Automatic Identification System (AIS)

Darmansah^{1,*}, Koko Handoko², Novri Adhiatma¹, Pastima Simanjuntak²

¹Fakultas Teknik dan Komputer, Sistem Informasi, Universitas Putera Batam, Batam

Jl. Raden Patah No.12A, Lubuk Baja Kota, Kec. Lubuk Baja, Kota Batam, Kepulauan Riau, Indonesia

²Fakultas Teknik dan Komputer, Teknik Informatika, Universitas Putera Batam, Batam

Jl. Raden Patah No.12A, Lubuk Baja Kota, Kec. Lubuk Baja, Kota Batam, Kepulauan Riau, Indonesia

Email: ^{1,*}darmansah071@gmail.com, ²kokohandoko01@gmail.com, ³ad.novri@gmail.com, ⁴pastima@puterabatam.ac.id⁴

Correspondence Author Email: darmansah071@gmail.com

Submitted: 07/03/2026; Accepted: 24/04/2026; Published: 26/04/2026

Abstrak—Wilayah perairan Batam merupakan salah satu jalur pelayaran tersibuk di Indonesia yang memiliki tingkat kepadatan lalu lintas kapal yang tinggi dan pola pergerakan yang kompleks. Kondisi ini membutuhkan teknik analisis data yang dapat mengidentifikasi pola pergerakan kapal secara akurat dan menyesuaikan. Tujuan dari penelitian ini adalah untuk mempelajari pola pergerakan kapal menggunakan data Sistem Identifikasi Otomatis (AIS), dan juga untuk melihat bagaimana algoritma K-Means dan DBSCAN bekerja dalam proses clustering data. Data AIS yang digunakan mencakup koordinat geografis, waktu pengamatan, kecepatan, dan arah pergerakan kapal di perairan Batam. Penelitian ini mencakup penerapan algoritma K-Means dan DBSCAN, ekstraksi dan normalisasi fitur, dan pra-pemrosesan data untuk meningkatkan kualitas data. metrik validasi internal yang digunakan untuk menilai kualitas kluster yaitu Silhouette Score dan Davies–Bouldin Index. Hasil penelitian menunjukkan bahwa algoritma DBSCAN memiliki tingkat kohesi kluster dan separasi antar kluster yang lebih baik daripada K-Means. Algoritma K-Means menghasilkan nilai Silhouette Score sebesar 0,48 dan nilai Davies–Bouldin Index sebesar 0,91, sedangkan algoritma DBSCAN menghasilkan nilai Silhouette Score sebesar 0,62 dan nilai Davies–Bouldin Index sebesar 0,67. Selain itu, DBSCAN dapat menemukan data suara sebesar 19,96% dari kumpulan data, yang menunjukkan pergerakan kapal yang tidak normal atau tidak membentuk pola kepadatan tertentu. Hasilnya menunjukkan bahwa algoritma DBSCAN menganalisis pola pergerakan kapal dengan data AIS di wilayah perairan Batam lebih baik daripada K-Means. Penelitian ini diharapkan dapat menjadi dasar untuk pengembangan sistem informasi maritim yang membantu memantau lalu lintas kapal, membuat keputusan tentang keselamatan, dan mengelola perairan.

Kata Kunci: Sistem Identifikasi Otomatis (AIS); Pengelompokan; K-Means; DBSCAN; Pola Pergerakan Kapal; Batam

Abstract—Batam waters are one of the busiest shipping lanes in Indonesia, with high ship traffic density and complex movement patterns. This condition requires data analysis techniques that can accurately identify and adapt ship movement patterns. The purpose of this study is to study ship movement patterns using Automatic Identification System (AIS) data, and also to see how the K-Means and DBSCAN algorithms work in the data clustering process. The AIS data used includes geographic coordinates, observation time, speed, and direction of ship movement in Batam waters. This study includes the application of the K-Means and DBSCAN algorithms, feature extraction and normalization, and data pre-processing to improve data quality. Internal validation metrics used to assess cluster quality are the Silhouette Score and the Davies–Bouldin Index. The results of the study show that the DBSCAN algorithm has a better level of cluster cohesion and separation between clusters than K-Means. The K-Means algorithm produces a Silhouette Score value of 0.48 and a Davies–Bouldin Index value of 0.91, while the DBSCAN algorithm produces a Silhouette Score value of 0.62 and a Davies–Bouldin Index value of 0.67. In addition, DBSCAN can find sound data of 19.96% of the data set, which indicates abnormal ship movements or does not form a certain density pattern. The results show that the DBSCAN algorithm analyzes ship movement patterns with AIS data in the Batam waters better than K-Means. This research is expected to be the basis for the development of maritime information systems that help monitor ship traffic, make decisions about safety, and manage waters.

Keywords: Automatic Identification System (AIS); Clustering; K-Means; DBSCAN; Vessel Movement Pattern; Batam

1. PENDAHULUAN

Batam adalah salah satu wilayah maritim paling strategis di Indonesia. Karena berada di dekat Selat Singapura, salah satu jalur pelayaran tersibuk di dunia [1],[2]. Kawasan ini menjadi hub utama untuk kapal internasional dan kapal penumpang, serta berbagai aktivitas maritim seperti pelabuhan, galangan kapal, dan zona industri [3],[4],[5]. Untuk memastikan keselamatan pelayaran, efisiensi operasional, dan pengawasan maritim, sistem pemantauan dan analisis pergerakan kapal yang berkualitas tinggi diperlukan karena banyaknya kapal yang melintasi perairan Batam. Automatic Identification System (AIS) adalah salah satu teknologi utama yang digunakan untuk memantau lalu lintas laut [6],[7],[8]. AIS memberikan informasi spasial dan temporal tentang kapal secara real-time, seperti posisi geografis, kecepatan, arah, dan identitas kapal [9],[10],[11]. Dalam konteks lalu lintas yang padat dan kompleks di perairan Batam, data AIS sangat berguna untuk mengidentifikasi pola jalur pelayaran utama, area dengan kepadatan tinggi, dan kemungkinan adanya pelanggaran.

Beberapa penelitian menggunakan analisis data dan pembelajaran mesin untuk menganalisis pola pergerakan kapal. Data AIS sangat potensial untuk digunakan untuk menunjukkan dinamika lalu lintas kapal secara spasial dan temporal [12]. Namun demikian, meskipun jumlah penelitian yang terkait dengan analisis data AIS terus meningkat, masih ada banyak kesenjangan penelitian (research gap) yang signifikan baik dari perspektif

metodologi, konteks wilayah studi, maupun kedalaman analisis [13],[14]. Masalah utama yang belum banyak dipelajari terletak pada penerapan algoritma terhadap karakteristik spasial-geometris wilayah perairan yang sangat kompleks, seperti Batam. Dengan banyak pulau kecil, jalur sempit (jalur sempit), area pelabuhan dengan aktivitas tinggi, dan area perairan terbuka yang saling terhubung, konfigurasi geografis perairan Batam unik. Kondisi ini menyebabkan distribusi data AIS yang tidak homogen, multidensitas, dan pola overlapping dan non-linear antar jalur pelayaran [6].

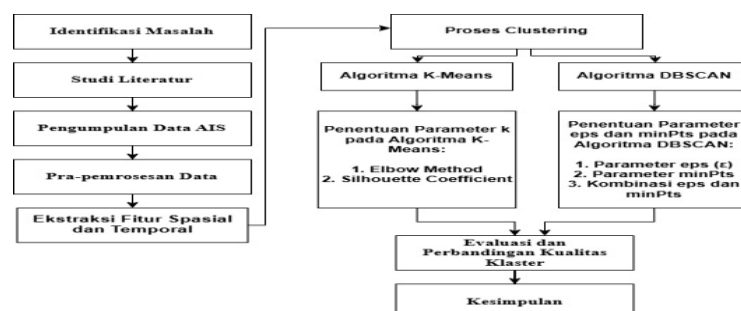
Penelitian sebelumnya biasanya memeriksa algoritma clustering pada dataset AIS yang lebih sederhana atau hanya berfokus pada skala makro tanpa mempertimbangkan kompleksitas geometris lokal. Akibatnya, belum ada penelitian khusus yang menilai kinerja algoritma seperti K-Means dan DBSCAN dalam menghadapi kondisi fragmentasi spasial, pola pergerakan yang saling beririsan, multi-skala kepadatan (multi-density), dan kondisi multi-skala kepadatan (multi-density). Selain itu, parameter DBSCAN seperti nilai ϵ dan minPts sangat sensitif terhadap struktur spasial data, sehingga studi kontekstual yang mempertimbangkan karakteristik lokal diperlukan. Kedua, dari perspektif geografis, sebagian besar penelitian yang melakukan analisis data AIS dilakukan pada wilayah perairan yang relatif homogen atau memiliki ciri lalu lintas yang lebih sederhana, seperti perairan terbuka atau jalur pelayaran utama [15]. Karena berada di persimpangan jalur pelayaran internasional, dekat dengan pelabuhan besar, dan berdekatan dengan wilayah perairan negara lain, wilayah perairan Batam memiliki kompleksitas yang jauh lebih tinggi [16],[17]. Karena kompleksitasnya, pola pergerakan kapal di Batam sangat berbeda dari segi kepadatan, arah pergerakan, dan jenis aktivitas kapal. Namun, penelitian akademik yang secara khusus menggunakan Batam sebagai lokasi studi untuk menganalisis pola pergerakan kapal dengan data AIS masih relatif terbatas. Ini menunjukkan adanya kesenjangan dalam penelitian tentang penggunaan data AIS untuk wilayah maritim strategis dengan karakteristik lalu lintas ekstrem seperti Batam. Ketiga, penelitian sebelumnya biasanya berfokus pada hasil visualisasi kluster tanpa melakukan evaluasi kuantitatif dan kontekstual kualitas kluster. Banyak studi hanya mengukur keberhasilan clustering berdasarkan kemampuan algoritma untuk memisahkan data ke dalam kelompok tertentu, tanpa memeriksa secara menyeluruh apakah kluster memiliki makna operasional dalam konteks maritim [18],[19].

Kesenjangan ini menunjukkan perlunya penelitian yang tidak hanya membandingkan algoritma secara teknis, tetapi juga mengaitkan hasil clustering dengan kondisi lalu lintas laut aktual di wilayah Batam. Meskipun DBSCAN sering disebut lebih baik untuk menangani data spasial dan suara, penelitian terdahulu mengenai seberapa baik DBSCAN menangani data AIS dengan variasi kepadatan tinggi tidak konsisten. Studi tertentu menunjukkan DBSCAN mampu menemukan pola jalur pelayaran dengan baik, tetapi studi lain menemukan bahwa parameter yang tepat sulit ditemukan [20],[21]. Sebaliknya, K-Means, yang sering dianggap tidak cocok untuk data spasial kompleks, masih digunakan karena mudah digunakan dan efisien dalam komputasi [22],[23]. Mencoba membandingkan kedua algoritma tersebut pada data AIS yang sangat padat dan dinamis seperti Batam masih jarang. Hal ini menimbulkan celah penelitian terkait konteks penggunaan algoritma, yakni pada kondisi seperti apa masing-masing algoritma memberikan performa terbaik. Dengan demikian, research gap penelitian ini terletak pada kebutuhan untuk mengevaluasi ulang dan memvalidasi kinerja algoritma clustering dalam konteks spasial-geometris yang kompleks dan tidak homogen, khususnya di wilayah perairan Batam.

Penelitian ini tidak sekadar membandingkan algoritma secara keseluruhan, tetapi juga berfokus pada bagaimana karakteristik khusus wilayah Batam memengaruhi kualitas kluster yang dihasilkan oleh K-Means dan DBSCAN, dan bagaimana kualitas kluster yang dihasilkan. Oleh karena itu, penelitian ini membandingkan secara menyeluruh algoritma K-Means dan DBSCAN dalam analisis pola pergerakan kapal dengan data AIS di wilayah Batam. Penelitian ini juga memeriksa kinerja teknis kedua algoritma tersebut dan melihat apakah hasil clustering relevan dengan kondisi lalu lintas maritim yang sebenarnya dan apakah mereka dapat membantu dalam pengembangan sistem informasi maritim berbatasan. Oleh karena itu, diharapkan bahwa penelitian ini akan memperkuat posisi akademik penelitian sekaligus memberikan kontribusi teoritis dan praktis yang signifikan.

2. METODOLOGI PENELITIAN

Adapun tahapan penelitian dapat dilihat pada digaram yang terdapat pada Gambar 1.



Gambar 1. Metodologi Penelitian



Penelitian ini dirancang menggunakan pendekatan kuantitatif eksperimental dengan fokus pada perbandingan kinerja algoritma clustering K-Means dan DBSCAN dalam menganalisis pola pergerakan kapal berbasis data AIS di wilayah perairan Batam. Desain penelitian disusun secara bertahap untuk memastikan validitas, reproduktibilitas, dan relevansi hasil penelitian terhadap konteks sistem informasi maritim.

2.1 Identifikasi Masalah

Tahap awal penelitian difokuskan pada identifikasi permasalahan yang berkaitan dengan tingginya kompleksitas lalu lintas kapal di wilayah perairan Batam serta keterbatasan metode analisis data AIS dalam mengungkap pola pergerakan kapal secara optimal. Permasalahan dirumuskan berdasarkan kebutuhan akan metode clustering yang mampu menangani data spasial berdimensi besar, mengandung noise, serta memiliki variasi kepadatan tinggi. Tahap ini menjadi dasar dalam menentukan fokus penelitian, batasan masalah, serta pemilihan algoritma yang relevan untuk dibandingkan.

2.2 Studi Literatur

Studi ini menyelidiki teori dan penelitian sebelumnya tentang analisis data AIS, clustering data spasial, algoritma K-Means dan DBSCAN, dan aplikasinya dalam sistem informasi maritim. Tujuan dari langkah ini adalah untuk menemukan keadaan saat ini, memperluas ruang lingkup penelitian, dan menentukan variabel dan teknik evaluasi yang tepat. Literatur yang digunakan dalam penelitian ini yaitu Jurnal, prosiding konferensi, dan referensi akademik yang relevan dengan penelitian.

2.3 Pengumpulan Data AIS

Penelitian ini menggunakan data Automatic Identification System (AIS) yang mencakup wilayah perairan Batam dan sekitarnya. AIS yang dikumpulkan termasuk informasi tentang posisi geografis (latitude dan longitude), waktu pengiriman data, kecepatan kapal (SOG), dan arah pergerakan kapal (COG). Wilayah Batam dipilih karena karakteristik lalu lintasnya yang padat dan berbeda, yang menimbulkan masalah nyata dalam analisis data maritim.

2.4 Pra-pemrosesan Data

Sebelum data diproses lebih lanjut, tahap pra-pemrosesan dimaksudkan untuk meningkatkan kualitas data. Karena kualitas data input sangat bergantung pada kualitas data input, terutama data AIS real-time dan raw, proses ini meliputi penghapusan data duplikat, penanganan data yang tidak ada nilainya (missing values), penyaringan suara bising yang tidak relevan, dan normalisasi atau standarisasi data jika diperlukan.

2.5 Ekstraksi Fitur Spasial dan Temporal

Pada tahap ini dilakukan pemilihan dan pembentukan fitur-fitur yang relevan untuk proses clustering. Untuk memastikan bahwa pola pergerakan kapal dapat digambarkan secara menyeluruh, baik dari sisi lokasi maupun perilaku pergerakan, fitur yang digunakan terdiri dari kombinasi atribut spasial, seperti latitude dan longitude, serta atribut pendukung lainnya, seperti kecepatan dan arah kapal.

2.6 Proses Clustering

2.6.1 K-Means

Dua metode kuantitatif digunakan untuk menentukan jumlah kluster (k) yaitu Elbow Method dan Silhouette Analysis. Metode Elbow dilakukan dengan menghitung nilai Within-Cluster Sum of Squares (WCSS) untuk beberapa nilai k , misalnya $k = 2$ hingga $k = 10$. Nilai k optimal ditemukan pada titik "siku", atau elbow, ketika penurunan WCSS mulai melambat secara signifikan. Silhouette Analysis Untuk memvalidasi hasil Elbow Method, digunakan Silhouette Score yang mengukur tingkat kohesi dan separasi kluster. Nilai k dengan skor silhouette tertinggi dipilih sebagai kandidat optimal. Kriteria pemilihan akhir k yaitu Titik elbow pada grafik WCSS Didukung oleh nilai silhouette yang relatif tinggi dan stabil.

2.6.2 DBSCAN

Untuk menghindari hasil clustering yang bias, parameter DBSCAN ditetapkan dengan hati-hati. pertama, penentuan ϵ (epsilon) dengan K-Distance Graph dan minPts determination. Langkah pertama dalam penentuan ϵ (epsilon) menggunakan Grafik Jarak K adalah menghitung jarak ke tetangga ke- k ($k = \text{minPts}$), mengurutkan jarak tersebut ke atas, dan kemudian memplot jarak k pada grafik. Nilai ϵ ditentukan pada titik "knee point" (titik perubahan gradien signifikan), yang merepresentasikan batas antara area padat dan jarang. Nilai minPts dimulai dengan menghitung dimensi data (d), eksperimen empiris (misalnya 4–10), dan stabilitas jumlah kluster dan proporsi suara. Dalam penelitian ini, dimensi data (d) adalah 4 (lat, lon, SOG, COG), dan minPts diuji dalam rentang 4–10. Nilai terbaik dipilih berdasarkan keseimbangan antara jumlah kluster dan suara.

2.7 Evaluasi dan Perbandingan Kualitas Kluster

Hasil clustering dari masing-masing algoritma dievaluasi menggunakan metrik evaluasi kluster yang sesuai, seperti indeks validitas internal. Tujuan evaluasi ini adalah untuk mengukur kualitas pemisahan kluster, konsistensi pola



yang terbentuk, dan kemampuan algoritma untuk menangani noise. Selain evaluasi kuantitatif, evaluasi kualitatif juga dilakukan dengan membandingkan hasil kluster dengan kondisi lalu lintas kapal nyata di sekitar Batam.

2.8 Kesimpulan dan Rekomendasi

Setelah analisis dan diskusi selesai, tahap akhir penelitian adalah membuat kesimpulan. Disusun juga saran untuk pengembangan penelitian lanjutan, penggunaan algoritma clustering yang paling sesuai, dan penerapan temuan penelitian dalam sistem informasi maritim.

3. HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah untuk membandingkan kinerja algoritma K-Means dan DBSCAN saat menganalisis pola pergerakan kapal menggunakan data Sistem Pengenalan Automatik (AIS) di wilayah perairan Batam. Proses analisis ini dimulai dengan pra-pemrosesan data, ekstraksi fitur, penerapan algoritma clustering, dan kemudian evaluasi kualitas kluster yang dihasilkan. Pada bagian ini disajikan hasil penelitian dalam dua bagian utama yaitu hasil dari penggunaan algoritma clustering dan perbandingan kinerja masing-masing algoritma dalam konteks karakteristik data AIS dan kondisi lalu lintas maritim di wilayah Batam.

3.1 Gambaran Umum Dataset Penelitian

Penelitian ini menggunakan data AIS yang mencakup wilayah perairan Batam dan sekitarnya. Mengingat lokasinya yang strategis di jalur pelayaran internasional Selat Singapura, wilayah ini dipilih karena memiliki karakteristik lalu lintas kapal yang sangat padat dan kompleks. Data AIS yang digunakan merekam posisi dan pergerakan kapal secara berkala, memungkinkan visualisasi dinamika lalu lintas laut secara spasial dan temporal. Atribut utama dalam penelitian ini adalah Identitas kapal (MMSI), koordinat geografis (latitude dan longitude), waktu pengamatan (timestamp), kecepatan kapal (Speed Over Ground/SOG), dan arah pergerakan kapal (Course Over Ground/COG). Kemudian berdasarkan hasil optimasi menggunakan Elbow Method dan Silhouette Analysis, diperoleh nilai optimal $k = 3$, yang menunjukkan bahwa data AIS di wilayah Batam secara alami dapat dikelompokkan menjadi tiga pola utama pergerakan kapal.

3.2 Hasil Pra-Pemrosesan Data

Pra-pemrosesan data dilakukan untuk memastikan bahwa data yang digunakan dalam proses clustering memiliki kualitas yang memadai. Proses ini menyaring data dengan nilai yang tidak valid, duplikat, dan data yang tidak termasuk dalam wilayah studi Batam. Hasil pra-pemrosesan data menunjukkan bahwa beberapa data AIS yang belum diproses tidak layak digunakan untuk analisis lanjutan. Hasil pra-pemrosesan disajikan dalam Tabel 1.

Tabel 1 Hasil pra-pemrosesan

Tahapan	Jumlah Data
Data Ais Awal	120.000
Penghapusan Data Duplikat	18.500
Penghapusan data tidak valid	12.300
Data siap analisis	89.200

3.3 Perhitungan K-Means

Proses pengelompokan data dilakukan dengan membentuk tiga kluster ($k = 3$). Pengukuran tingkat kemiripan antar data menggunakan metode Euclidean Distance, dengan mempertimbangkan fitur latitude, longitude, dan Speed Over Ground (SOG). Berikut ini adalah inisialisasi Centroid Awal yang di pilih secara acak yang terdapat pada Tabel 2.

Tabel 2. Centroid Awal

Centroid	Latitude	Longitude	SOG
C1	1.120	103.950	12
C2	1.300	104.100	2
C3	1.500	104.300	15

Kemudian perhitungan jarak antar data dalam penelitian ini menggunakan rumus Euclidean Distance, yang berfungsi untuk mengukur tingkat kedekatan antara dua titik berdasarkan beberapa variable, berikut Adalah rumus yang di pakai:

$$d(r, y) = \sqrt{(lat_r - lat_y)^2 + (lon_r - lon_y)^2 + (sog_r - sog_y)^2} \tag{1}$$

Hasil perhitungan, Data AIS dibagi secara relatif merata ke dalam tiga kluster utama, menurut hasil clustering K-Means dengan $k = 3$. Kluster pertama (C1) menunjukkan jalur pelayaran dengan kepadatan tinggi dan arah yang relatif homogen, yang dapat dikaitkan dengan koridor utama Selat Singapura. Kluster kedua (C2)

menunjukkan kecepatan rendah dan variasi arah yang lebih tinggi, yang menunjukkan aktivitas kapal di area pelabuhan atau zona labuh. Kluster ketiga (C3) menunjukkan jalur pelayaran dengan kecepatan tinggi dan pola yang lebih tersebar, yang dapat dikaitkan dengan koridor utama Selat Singapura. Namun, distribusi kluster yang seimbang ini menunjukkan bahwa K-Means cenderung membagi data secara global tanpa memperhitungkan variasi kepadatan lokal. Ini meningkatkan kemungkinan overlap antar kluster, terutama di daerah dengan pola pergerakan yang rumit seperti pelabuhan. Selain itu, ketika tidak ada identifikasi noise, seluruh data dipaksakan masuk ke dalam kluster tertentu. Akibatnya, kemampuan model untuk mengidentifikasi pergerakan kapal yang bersifat anomali berkurang. Adapun hasil pengelompokan kluster dapat di lihat secara lengkap pada Tabel 3.

Tabel 3. Hasil Pengelompokan K-Means

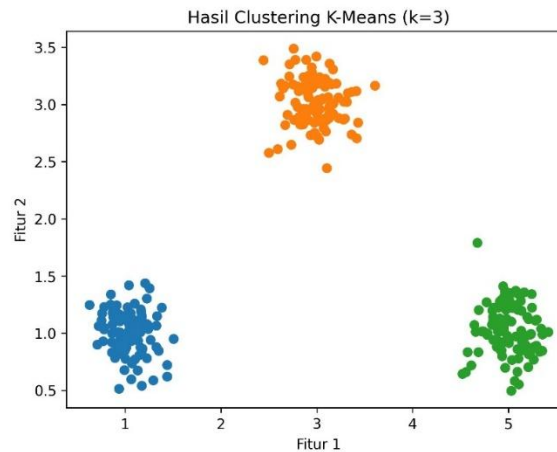
Kluster	Jumlah Data	Persentase
C1	30.500	34,19%
C2	28.700	32,18%
C3	30.000	33,63%
Total	89.200	100%

Adapun hasil dari centroid akhir dalam perhitungan k-means ini bisa dilihat pada Tabel 4.

Tabel 4. Centroid Akhir

Kluster	Latitude	Longitude	SOG	COG
C1	0,25	0,4	0,58	0,7
C2	0,5	0,65	0,25	0,45
C3	0,78	0,22	0,72	0,15

Dari Iterasi berikutnya tidak mengubah keanggotaan kluster sehingga algoritma berhenti. Berikut Adalah grafik dari perhitungan algoritma K-Means yang dapat di lihat pada gambar 2 dibawah ini:



Gambar 2. Hasil Clustering K-Means

3.4 Perhitungan Detail DBSCAN

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) adalah algoritma clustering tidak terlihat yang mengelompokkan data berdasarkan kepadatan (density). Ini memiliki kemampuan untuk membentuk kluster dengan cara yang diinginkan dan dapat mengidentifikasi noise atau outlier dalam data AIS yang memiliki distribusi tidak merata. Pada penelitian ini, data AIS digunakan untuk menemukan pola pergerakan kapal di wilayah perairan Batam dengan DBSCAN. DBSCAN menggunakan dua fungsi utama yaitu epsilon yaitu Jarak maksimum antar titik agar dianggap bertetangga (neighborhood radius) dan minPts Jumlah minimum titik dalam radius ϵ yang memungkinkan suatu titik dianggap sebagai titik inti (core point). Pada penelitian ini digunakan parameter optimal ϵ (epsilon) = 0,015 dan minPts = 6. Fitur yang digunakan yaitu Latitude, Longitude, SOG dan COG serta Berikutkah rumus dalam perhitungan Euclidean Distance:

$$d(r_i, r_j) = \sqrt{\sum_{k=1}^m (r_{rk} - r_{jk})^2} \tag{8}$$

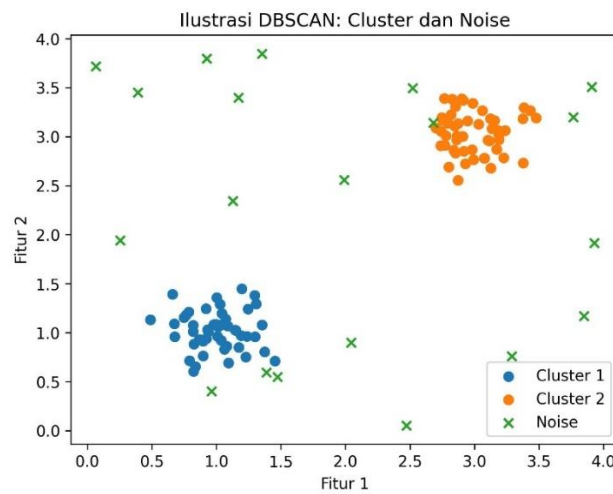
Adapun hasil dari perhitungan Clustering DBSCAN dapat dilihat pada Tabel 5.

Tabel 5. Hasil Clustering DBSCAN

Kategori	Jumlah
Kluster utama	3
Total data	89.200

Kategori	Jumlah
Data dalam klaster	71.400
Noise	17.800 (±19,96%)

Dari 89.200 data AIS, 71.400 (±80,4%) berhasil dikelompokkan ke dalam tiga klaster utama, dan 17.800 (±19,96%) diidentifikasi sebagai noise. Hasilnya menunjukkan bahwa pola kepadatan sebagian besar pergerakan kapal di perairan Batam cukup jelas dan dapat digambarkan dalam struktur klaster yang signifikan. Ada tiga klaster utama, yang menunjukkan pembagian pola pergerakan kapal yang alami. Klaster-klaster ini dapat dikaitkan dengan koridor pelayaran utama, area aktivitas pelabuhan, dan jalur sekunder. Sebaliknya, proporsi suara yang besar menunjukkan bahwa data AIS tidak konsisten, dengan sebagian besar pergerakan kapal tidak mengikuti pola kepadatan tertentu. Ini sekaligus menegaskan keunggulan DBSCAN dalam mengidentifikasi dan memisahkan data anomali, yang menghasilkan representasi pola pergerakan kapal yang lebih realistis dan sesuai dengan kompleksitas spasial wilayah perairan Batam. Berikut merupakan grafik dari hasil Algoritma DBSCAN yang bisa dilihat pada Gambar 3.



Gambar 3. Grafik hasil Algoritma DBSCAN

Dari grafik tersebut terlihat secara konseptual bahwa area dengan konsentrasi titik yang tinggi memenuhi kriteria sebagai core region, di mana setiap titik memiliki jumlah tetangga dalam radius ϵ yang lebih besar atau sama dengan $minPts$. Titik-titik di sekitar core region yang masih dapat diakses dalam radius ϵ disebut border points, dan titik-titik yang tidak memenuhi kriteria kepadatan disebut noise. Ini menunjukkan kemampuan DBSCAN untuk mengidentifikasi jalur pelayaran utama sebagai klaster padat, memisahkan aktivitas pelabuhan lokal menjadi klaster terpisah, dan menemukan pergerakan kapal yang tidak beraturan sebagai suara.

3.5 Perbandingan Algoritma K-Means dan DBSCAN

Adapun hasil perbandingan kedua algoritma ini dapat dilihat pada tabel 6 dibawah ini. Tabel ini menunjukkan bahwa algoritma DBSCAN dan K-Means berbeda dalam kinerja dan fitur untuk mengelompokkan data AIS di wilayah perairan Batam.

Tabel 6. Perbandingan Algoritma K-Means dan DBSCAN

Aspek Perbandingan	K-Means	DBSCAN
Pendekatan Clustering	Berbasis partisi (partition-based clustering)	Berbasis kepadatan (density-based clustering)
Penentuan Jumlah Klaster	Harus ditentukan di awal (nilai k)	Tidak perlu ditentukan di awal
Sensitivitas terhadap Skala Data	Sangat sensitif, memerlukan normalisasi	Relatif tidak sensitif terhadap skala
Kemampuan Menangani Noise	Tidak mampu mendeteksi noise	Mampu mengidentifikasi noise/outlier
Bentuk Klaster	Cenderung berbentuk bulat dan homogen	Fleksibel, mengikuti bentuk alami data
Ketahanan terhadap Variasi Kepadatan	Rendah	Tinggi
Interpretasi Klaster	Global dan umum	Spesifik dan berbasis kepadatan
Kinerja pada Data AIS Batam	Cukup baik untuk gambaran umum	Sangat baik dan lebih representatif

Aspek Perbandingan	K-Means	DBSCAN
Silhouette Score	0,48	0,62
Davies–Bouldin Index	0,91	0,67
Kompleksitas Komputasi	Rendah	Relatif lebih tinggi
Kesesuaian untuk Sistem Informasi Maritim	Analisis awal dan eksploratif	Pemantauan kepadatan dan deteksi anomali

Data AIS Batam menunjukkan variasi kepadatan yang tinggi antara jalur pelayaran utama, area pelabuhan, dan perairan terbuka. K-Means juga cenderung membentuk kluster global dengan variasi internal yang cukup besar, dan seluruh data dimasukkan ke dalam kluster tertentu, sehingga pergerakan kapal yang tidak biasa tidak dapat diidentifikasi. Sebaliknya, DBSCAN menggunakan metode berbasis kepadatan, yang memungkinkan pembentukan kluster sesuai dengan struktur data alami. Algoritma ini mampu mengidentifikasi data noise dengan mudah dan tidak memerlukan penentuan jumlah kluster. Kemampuan ini sangat bermanfaat dalam hal data AIS karena pergerakan kapal seringkali sporadis dan tidak selalu mengikuti pola yang seragam. Hasil evaluasi kualitas kluster menunjukkan bahwa DBSCAN memiliki kohesi internal yang lebih baik serta separasi antar kluster yang lebih jelas daripada K-Means. Dengan mempertimbangkan aspek kuantitatif dan kualitatif, DBSCAN dianggap lebih sesuai untuk analisis pola pergerakan kapal menggunakan data AIS di wilayah Batam. Ini terutama berlaku untuk kebutuhan sistem informasi maritim yang berfokus pada pemantauan kepadatan lalu lintas, identifikasi jalur pelayaran, dan deteksi pergerakan kapal yang bersifat anomali. Namun, K-Means masih relevan sebagai alat analisis awal dan metode pembandingan yang memberikan gambaran umum distribusi pergerakan kapal.

3.6 Hasil Evaluasi Kualitas Kluster Menggunakan Silhouette Score dan Davies–Bouldin Index

Berikut ini adalah hasil dari Evaluasi Kualitas Kluster Menggunakan Silhouette Score dan Davies–Bouldin Index yang bisa di lihat pada tabel 7 dibawah ini:

Tabel 7. Hasil Evaluasi Kualitas Kluster Menggunakan Silhouette Score dan Davies–Bouldin Index

Algoritma	Parameter	Jumlah Kluster	Silhouette Score	Davies–Bouldin Index	Interpretasi
K-Means	$k = 4$	4	0,48	0,91	Kohesi kluster sedang, separasi antar kluster relatif lemah
DBSCAN	$\epsilon = 0,015$; $\text{minPts} = 6$	3 kluster + noise	0,62	0,67	Kohesi kluster baik, separasi jelas, struktur lebih optimal

Tabel 7 menunjukkan bahwa algoritma DBSCAN memiliki tingkat kohesi intra-kluster yang lebih tinggi, yang berarti bahwa anggota dalam satu kluster memiliki kemiripan yang lebih besar daripada anggota dalam kluster lain. Nilai Silhouette Score algoritma adalah 0,62, lebih tinggi dari K-Means. Selain itu, nilai Davies–Bouldin Index pada DBSCAN sebesar 0,67 lebih rendah daripada K-Means sebesar 0,91, menunjukkan bahwa jarak antar kluster lebih terpisah dan tingkat overlap yang lebih rendah. Nilai DB Index yang lebih rendah menunjukkan kualitas kluster yang lebih baik dalam hal separasi dan kompaksi. Hasil yang berbeda ini secara langsung menunjukkan bahwa, dibandingkan dengan pendekatan K-Means yang berbasis partisi dan cenderung menghasilkan kluster dengan bentuk isotropik, pendekatan berbasis kepadatan DBSCAN lebih mampu menangkap struktur data AIS multi-densitas dan non-linear.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa karakteristik spasial data AIS yang dianalisis sangat memengaruhi kinerja algoritma clustering. Algoritma DBSCAN lebih baik daripada K-Means untuk menggambarkan struktur alami pergerakan kapal di wilayah perairan Batam yang memiliki pola multi-density dan non-linear. Hasil ini menegaskan bahwa pendekatan clustering yang memperhatikan konteks sangat penting. Selain itu, metode analisis harus disesuaikan dengan kompleksitas data dan kondisi geografis wilayah studi. Secara praktis, hasil penelitian ini membantu pengembangan sistem informasi maritim, terutama dalam hal pemantauan lalu lintas kapal dan deteksi anomali dengan kemampuan DBSCAN untuk mengidentifikasi suara. Karena keunggulannya dalam efisiensi komputasi, sementara itu K-Means tetap relevan sebagai metode eksplorasi awal. Namun demikian, penelitian ini memiliki keterbatasan, terutama berkaitan dengan skalabilitas DBSCAN pada dataset AIS berskala besar, yang dapat menyebabkan beban komputasi yang lebih besar. Hasil clustering juga sensitif terhadap pemilihan parameter ϵ dan minPts . Selain itu, penelitian ini belum mempertimbangkan faktor-faktor eksternal seperti kondisi lingkungan maritim. Oleh karena itu, untuk meningkatkan akurasi dan efisiensi model, penelitian lebih lanjut harus menggabungkan metode pengolahan data besar dan analisis spatiotemporal.

REFERENCES

- [1] J. P. Simanjuntak, J. Sihombing, “Optimalisasi Pelabuhan Dan Logistik Maritim Dalam Meningkatkan Daya Saing



- Batam Sebagai Pusat Perdagangan Internasional,” *Manag. Stud. Entrep. J.*, vol. 6, no. 6, pp. 1147–1153, 2025, doi: <https://doi.org/10.37385/msej.v6i6.9852>.
- [2] A. Asyikin and A. Setiawan, “Kerjasama Ekonomi Singapura-Malaysia di Johor dan Dampaknya terhadap Kawasan Ekonomi Khusus Batam,” *Rev. Int. Relations*, vol. 7, no. 1, pp. 47–73, 2025, doi: [10.24252/rir.v7i1.56799](https://doi.org/10.24252/rir.v7i1.56799).
- [3] M. B. Pandjaitan, K. Khusaini, M. F. Aminuddin, and P. Suwarno, “Peran Kemaritiman Dalam Memperkuat Perekonomian Indonesia,” *J. Sains Teknol. Transp. Marit.*, vol. 7, no. 1, pp. 17–27, 2025, doi: [10.51578/j.sitektransmar.v7i1.104](https://doi.org/10.51578/j.sitektransmar.v7i1.104).
- [4] D. A. Purnama, P. C. Marifa, R. C. Shinta, and R. C. Shinta, “Unlocking Indonesia’s Maritime Potential: Optimizing Hub Port Development using a Principal Component Analysis and K-Means Clustering,” *J. Sist. Tek. Ind.*, vol. 27, no. 1, pp. 35–46, 2025, doi: [10.51578/j.sitektransmar.v7i1.104](https://doi.org/10.51578/j.sitektransmar.v7i1.104).
- [5] F. N. T. Aritonang, J. C. Arkady, L. A. R. Torindatu, and V. A. J. M. Silalahi, “Indonesia’s Position as a Maritime Axis in the Global Logistics Chain,” *JIPOWER J. Intellect. Power*, vol. 2, no. 1, pp. 110–123, 2025, doi: [10.63786/jipower.v2i1.38](https://doi.org/10.63786/jipower.v2i1.38).
- [6] D. A. Setyawan, A. Purwatiningsih, and F. S. Budi, “Analisis Pola Lalu Lintas Kapal Selat Bali Berbasis AIS , K-Means , Traffic Flow,” *J. Fasilkom*, vol. 15, no. 3, pp. 621–629, 2026, doi: <https://doi.org/10.37859/jf.v15i3.10782>.
- [7] L. Barasa, A. L. Togatorop, and M. Szesze, “Optimalisasi Pengoperasian AIS (Automatic Identification System) Dalam Upaya Menjaga Keselamatan Pelayaran,” *J. Ilm. Nas. Sekol. Tinggi Ilmu Pelayaran Jakarta*, vol. 14, no. 2, pp. 100–129, 2021, doi: <https://doi.org/10.36101/msm.v16i1.268>.
- [8] I. M. D. P. Asana, I. M. O. Widyantara, L. Linawati, D. M. Wiharta, and I. G. N. S. Wikananda, “Real-Time Web-Based Ship Collision Risk Detection Using AIS Data and Collision Risk Index (CRI),” *Sinkron*, vol. 9, no. 4, pp. 1942–1952, 2025, doi: [10.33395/sinkron.v9i4.15106](https://doi.org/10.33395/sinkron.v9i4.15106).
- [9] S. Basu, B. W. Lee, R. E. Clark, and S. Bera, “Studi Optimasi Implementasi AIS (Automatic Identification System) Tipe B Pada Kapal “Kapal Di Pelabuhan Perikanan Pantai (PPP) Mayangan Probolinggo,” *Basic Appl. Ecol.*, vol. 64, no. 11, pp. 57–67, 2022, doi: <https://doi.org/10.36418/syntax-literat.v7i11.10019>.
- [10] D. A. Rehardendi, A. Arfianto, D. P. Riananda, R. Burjulius, and S. Lena, “Pengiriman Pesan Darurat Berbasis Ais Untuk Jacket Keselamatan Di Kapal Penumpang,” *J. 7 Samudra*, vol. 10, no. 1, pp. 29–36, 2025, doi: [10.54992/7samudra.v10i1.330](https://doi.org/10.54992/7samudra.v10i1.330).
- [11] D. D. Apriani et al., “Evaluasi Penerapan Pemasangan Automatic Identification System (AIS) Pada Kapal Penyeberangan Yang Beroperasi Di Pelabuhan Paciran,” *IWJ Inl. Waterw. J.*, vol. 6, no. 1, pp. 19–25, 2024, doi: [10.54249/iwj.v6i1.207](https://doi.org/10.54249/iwj.v6i1.207).
- [12] M. S. Education and T. Management, “Integrating AIS-Based Probabilistic Collision Risk Assessment into Maritime Safety Education and Traffic Management,” *J. Ilm. Nas. Sekol. Tinggi Ilmu Pelayaran Jakarta*, vol. 8, no. 2, pp. 41–51, 2025, doi: <https://doi.org/10.36101/msm.v18i2.447>.
- [13] J. Yang, X. Bian, Y. Qi, X. Wang, Z. Yang, and J. Liu, “A spatial-temporal data mining method for the extraction of vessel traffic patterns using AIS data,” *Ocean Eng.*, vol. 293, no. November 2023, p. 116454, 2024, doi: [10.1016/j.oceaneng.2023.116454](https://doi.org/10.1016/j.oceaneng.2023.116454).
- [14] N. Ani, S. Catra Pratama, F. Aziz, and D. Fardiansyah, “Aplikasi Pemantauan Lalu Lintas Kapal di Perairan Laut dengan Menggunakan Metode Haversine Formula Article Info ABSTRAK,” *JSAI J. Sci. Appl. Informatics*, vol. 6, no. 3, pp. 511–516, 2024, doi: [10.36085](https://doi.org/10.36085).
- [15] I. I. J. Rifka Alkhiyatul Ma’rifat, I Made Suraharta, “Preliminary Study Of The Influence Of Seasonal Ocean Currents On Vessel Travel Time Through The Waters Of The Lombok Strait,” *J. Hidrogr. Indones.*, vol. 3, no. 2, pp. 306–312, 2024, doi: <https://doi.org/10.62703/jhi.v5i2.25>.
- [16] Y. Samiyono and Q. Amarona, “Identifikasi Peta Laut Dalam Rangka Pengembangan Kawasan Pelabuhan Batam, Bintan, Karimun (Bkk) Dan Sekitarnya,” *J. Matemar Manaj. Dan Teknol. Marit.*, vol. 4, no. 1, pp. 9–16, 2023, doi: [10.59225/matemar.v4i1.132](https://doi.org/10.59225/matemar.v4i1.132).
- [17] Y. Lisa Sitorus, L. Erdawati, and T. Marno Butarbutar, “Pengaruh Penetapan Kota Batam Sebagai Kawasan Ekonomi Khusus (KEK) Terhadap Pertumbuhan Ekonomi,” *J. Ilmu Huk.*, vol. 1, no. 4, pp. 174–181, 2024, doi: <https://doi.org/10.62017/syariah>.
- [18] D. Aditama, A. Taryana, and H. Siswantoro, “Di Lingkungan Indoor Bertingkat Dengan Algoritma K-Means,” *J. Sist. Inf. dan Teknol. Komputasi*, vol. 2, no. 3, pp. 133–141, 2025, doi: <https://doi.org/10.61124/sinta.v2i3.84>.
- [19] H. Kolengsusu, A. K. Henaulu, and F. Mony, “Comparison of Clusterization of Higher Education Institutions in Regions XII Maluku and North Maluku based on the Science and Technology Index (Sinta),” *J. Sains dan Teknol. Ind.*, vol. 21, no. 1, p. 124, 2023, doi: [10.24014/sitekin.v21i1.24826](https://doi.org/10.24014/sitekin.v21i1.24826).
- [20] D. M. Wiharta et al., “Identifikasi Aktivitas Illegal Transshipment Berbasis,” *J. Resist.*, vol. 5, no. 1, pp. 38–46, 2022, doi: <https://doi.org/10.31598/jurnalresistor.v5i1.1048>.
- [21] L. Busaina, N. R. Utami, S. Pramana, and D. Krismawati, “Pemetaan Daerah Aktivitas Perikanan Berbasis Data Automatic Identification System (AIS),” *Semin. Nas. Off. Stat.*, pp. 885–894, 2024, doi: <https://doi.org/10.34123/semnasoffstat.v2024i1.2299>.
- [22] Fadil Danu Rahman, M. I. Z. Mulki, and A. Taryana, “Clustering Dan Klasifikasi Data Cuaca Cilacap Dengan Menggunakan Metode K-Means Dan Random Forest,” *J. SINTA Sist. Inf. dan Teknol. Komputasi*, vol. 1, no. 2, pp. 90–97, 2024, doi: [10.61124/sinta.v1i2.15](https://doi.org/10.61124/sinta.v1i2.15).
- [23] F. Ramadhani, “Spatial Clustering Analysis of Stunting in North Sumatra Based on Environmental Factors Using K-Means Algorithm,” *Data Sci. J. Comput. Appl. Informatics*, vol. 9, no. 2, pp. 18–25, 2025, doi: [10.32734/jocai.v9i2-17179](https://doi.org/10.32734/jocai.v9i2-17179).