

CLUSTERING AKTIVITAS KAPAL YANG MASUK KE PELABUHAN KEPULAUAN SERIBU DENGAN FUZZY C- MEANS

Michelle Angela Thena

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara
Jl. Letjen S. Parman No.1, Jakarta Barat, DKI Jakarta, Indonesia 11410

Email: michelle.535200020@stu.untar.ac.id

ABSTRAK

Sebagai salah satu prasarana transportasi, pelabuhan memiliki peran strategis untuk mendukung sistem transportasi karena menjadi titik simpul hubungan antar daerah/negara. Pelabuhan memiliki fungsi sosial dan ekonomi. Secara ekonomi, pelabuhan berfungsi sebagai salah satu penggerak roda perekonomian karena menjadi fasilitas yang memudahkan distribusi hasil-hasil produksi. Secara sosial, pelabuhan menjadi fasilitas publik tempat berlangsungnya interaksi antarpengguna (masyarakat), termasuk interaksi yang terjadi karena adanya aktivitas perekonomian. Analisis ini dilakukan dengan menggunakan metode *clustering fuzzy c-means* untuk mengamati banyaknya aktivitas kapal dan penumpang yang tiba di beberapa pelabuhan Kepulauan Seribu pada bulan Januari 2021. Dari percobaan *clustering* tersebut, didapatkan hasil rata-rata *silhouette* terbaik dari 2 *cluster* dengan nilai 0.9228549970193373.

Kata kunci: Klasterisasi, Klaster, *Fuzzy C-Means*

ABSTRACT

As one of the transportation infrastructures, ports have a strategic role to support the transportation system because they are the nodes of relations between regions/countries. The port has a function both from an economic and social perspective. Economically, the port functions as one of the driving wheels of the economy because it is a facility that facilitates the distribution of production results. Socially, the port is a public facility where interactions between users (community) take place, including interactions that occur due to economic activity. This analysis was carried out using the fuzzy c-means clustering method to observe the many activities of ships and passengers arriving at several Kepulauan Seribu ports in January 2021. From the clustering experiment, the best average silhouette results were obtained from 2 clusters with a value of 0.9228549970193373.

Keywords: *Clustering, Cluster, Fuzzy C-Means*

1. PENDAHULUAN

Pelabuhan adalah tempat yang terdiri dari daratan dan perairan dan sekitarnya dengan batas-batas tertentu sebagai tempat kegiatan Pemerintahan dan kegiatan layanan jasa. Utamanya pelabuhan adalah tempat kapal bersandar, berlabuh, naik turun penumpang dan/atau bongkar muat barang yang dilengkapi dengan fasilitas keselamatan pelayaran dan kegiatan penunjang pelabuhan serta sebagai tempat perpindahan intra dan antar moda transportasi (Gurning & Budiyanto, 2007).

Pelabuhan adalah tempat yang terdiri dari daratan dan perairan disekitarnya dengan batas-batas tertentu sebagai tempat kegiatan pemerintahan dan kegiatan ekonomi yang dipergunakan sebagai tempat kapal bersandar, berlabuh, naik-turun penumpang, dan/atau bongkar muat barang yang dilengkapi dengan fasilitas keselamatan pelayaran dan kegiatan penunjang pelabuhan dan serta sebagai tempat perpindahan intra dan antarmoda transportasi. Pelabuhan umum adalah pelabuhan yang diselenggarakan untuk kepentingan pelayanan masyarakat umum (Suranto, 2004).

Pengembangan transportasi diarahkan untuk menjembatani kesenjangan antar wilayah dan mendorong pemerataan hasilhasil pembangunan. Transportasi laut memegang peranan penting dalam kelancaran perdagangan karena memiliki nilai ekonomis yang tinggi antara lain daya angkut banyak, dan biaya relatif murah. Guna menunjang perdagangan dan lalulintas muatan, pelabuhan diciptakan sebagai titik simpul perpindahan muatan barang dimana kapal dapat berlabuh, bersandar, melakukan bongkar muat barang dan penerusan ke daerah lainnya (Kramadibrata, 1985).

Kinerja pelabuhan dapat digunakan untuk mengetahui tingkat pelayanan pelabuhan kepada pengguna pelabuhan (kapal dan barang), yang tergantung pada waktu pelayanan kapal selama berada di pelabuhan. Kinerja pelabuhan yang tinggi menunjukkan bahwa pelabuhan tersebut dapat memberikan pelayanan yang baik. Kinerja suatu pelabuhan dapat di evaluasi dari sudut pandang efisiensi teknis, efisiensi biaya dan efektifitas. (Triatmodjo, 2009) Efektifitas terkait dengan seberapa baik pelabuhan menyediakan pelayanan arus barang kepada pengguna perusahaan pelayaran atau pun ekspedisi (Talley, 2009)

Tujuan

Tujuan dari penelitian ini adalah untuk menganalisis banyaknya aktivitas kapal yang tiba di beberapa pelabuhan Kepulauan Seribu beserta dengan jumlah penumpang kapal yang turun. Hasil dari analisis tersebut kemudian akan dipakai sebagai referensi untuk mengatur jadwal datangnya kapal ke pelabuhan guna menghindari terjadinya penumpukan kapal yang masuk. Sehingga aktivitas keluar masuk kapal di pelabuhan pun dapat berjalan dengan kondusif,

2 METODE

Metode yang akan digunakan adalah metode *clustering fuzzy c-means*. *Clustering* merupakan proses pengelompokan data dalam kelas-kelas atau *cluster-cluster* sehingga data dalam suatu *cluster* memiliki tingkat persamaan yang tinggi satu dengan lainnya tetapi sangat berbeda dengan data dalam *cluster* lain. (Luthfi, 2007) Algoritma *fuzzy c-means* merupakan salah satu algoritma pengklasteran fuzzy yang menghasilkan soft-partisi terkendala dan merupakan suatu metode partisi iteratif yang bertujuan menemukan pusat *cluster* yang meminimumkan fungsi ketidakmiripan sehingga menghasilkan cpartisi optimal. Metode tersebut menghitung pusat klaster (centroid) dan membangkitkan matriks kelas keanggotaan. Metode ini dikembangkan oleh Dunn (1973) dan diperbaiki oleh Bezdek (1981). Metode ini juga sering digunakan dalam pengenalan pola. (Abdy, 2015)

Secara detail, algoritma *fuzzy c-means* mempunyai tahapan-tahapan sebagai berikut (Abdy, 2015):

1. Pilih suatu nilai untuk parameter fuzziness klaster m , dengan $m > 1$
2. Pilih suatu nilai untuk kriteria penghentian iterasi ε (yaitu $\varepsilon = 0.0001$ memberikan suatu konvergensi yang layak)
3. Pilih suatu ukuran jarak dalam variabel-space (yaitu jarak Euclidean)
4. Pilih banyaknya kelas atau grup c , dengan $c = 2, 3, \dots, n - 1$
5. Inisialisasi secara acak matriks keanggotaan (U) dengan kendala

$$\sum_{i=1}^c \mu_{ij} = 1$$

untuk setiap $j = 1, 2, \dots, n$

6. Hitung pusat klaster (v_i) dengan menggunakan persamaan (4)
7. Hitung ketidakmiripan diantara pusat klaster dan titik data dengan menggunakan persamaan (3). Hentikan iterasi jika $\|U^{[k+1]} - U^{[k]}\| < \varepsilon$
8. Hitung suatu U baru dengan persamaan (5). Lanjutkan ke tahap 6

Konsep dasar FCM (*fuzzy c-means*) pertama kali adalah menentukan pusat *cluster* yang akan menandai lokasi rata-rata untuk tiap *cluster*. Pada kondisi awal, pusat *cluster* ini masih belum

akurat. Tiap-tiap titik data memiliki derajat keanggotaan untuk tiap-tiap *cluster*. Dengan cara memperbaiki pusat *cluster* dan derajat keanggotaan tiap-tiap titik data secara berulang, maka akan dapat dilihat bahwa pusat *cluster* akan bergerak menuju lokasi yang tepat. Perulangan ini didasarkan pada derajat keanggotaan yang menggambarkan jarak dari titik data yang diberikan ke pusat *cluster* yang terbobot oleh derajat keanggotaan titik data tersebut (Kusumadewi & Purnomo, 2010).

Pada perhitungan menggunakan FCM dengan menggunakan data yang sama tetapi diolah dengan jumlah *cluster* yang berbeda, maka hasil pengelompokannya akan sedikit berbeda, karena data tidak diolah dengan satu variabel saja tetapi dengan semua variabel. Perbedaan hasil pengelompokan itu dikarenakan data pada kelompok tertentu kemungkinan akan berpindah pada kelompok lain jika diolah dengan jumlah *cluster* yang berbeda, ini menunjukkan bahwa sistem aplikasi sudah berjalan dengan benar (Kusumadewi & Purnomo, 2010). Setelahnya, akan dilakukan evaluasi dengan menggunakan metode *silhouette coefficient*. *Silhouette coefficient* merupakan metode yang digunakan untuk melihat kualitas dan kekuatan dari *cluster*. Metode *silhouette coefficient* merupakan gabungan dari dua metode yaitu metode cohesiion yang berfungsi untuk mengukur seberapa dekat relasi antara objek dalam sebuah *cluster*, dan metode separation yang berfungsi untuk mengukur seberapa jauh sebuah *cluster* terpisah dengan *cluster* lain (Pramesti, Furqon, & Dewi, 2017).

Tahapan perhitungan *silhouette coefficient* (Handoyo, Mangkudjaja, & Nasution, 2014):

- a. Hitung rata-rata jarak dari suatu dokumen misalkan *i* dengan semua dokumen lain yang berada dalam satu *cluster*

$$a(i) = \frac{1}{|A| - 1} \sum_j \in A, j \neq i d(i, j)$$

dengan *j* adalah dokumen lain dalam satu *cluster* *A* dan *d(i,j)* adalah jarak antara dokumen *i* dengan *j*.

- b. Hitung rata-rata jarak dari dokumen *i* tersebut dengan semua dokumen di *cluster* lain, dan diambil nilai terkecilnya.

$$d(i, C) = \frac{1}{|A|} \sum_{j \in C} d(i, j)$$

dengan *d(i,C)* adalah jarak rata-rata dokumen *i* dengan semua objek pada *cluster* lain *C* dimana *A* ≠ *C*.

$$b(i) = \min_{C \neq A} d(i, C)$$

- c. Nilai *Silhouette Coefficient* nya adalah :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

Nilai hasil *silhouette coefficient* terletak pada kisaran nilai -1 hingga 1. Semakin nilai *silhouette coefficient* mendekati nilai 1, maka semakin baik pengelompokan data dalam satu *cluster*. Sebaliknya jika nilai *silhouette coefficient* mendekati nilai -1, maka semakin buruk pengelompokan data di dalam satu *cluster* (Pramesti, Furqon, & Dewi, 2017).

3 HASIL DAN PEMBAHASAN

3.1 Data

Data yang dipakai untuk penelitian kali ini adalah data penumpang kapal dari dan ke Kepulauan Seribu bulan Januari 2021. Data tersebut diperoleh dari sumber data resmi yang dipublikasikan oleh pemerintah DKI Jakarta melalui *website* Jakarta Open Data.

Di dalam dataset tersebut, terdapat beberapa atribut yang dapat digunakan untuk penelitian seperti tanggal, pelabuhan yang menunjukkan nama-nama pelabuhan di Kepulauan Seribu, kapal_berangkat yang menunjukkan jumlah kapal yang berangkat atau meninggalkan pelabuhan tersebut, kapal_tiba menunjukkan jumlah kapal yang tiba atau berlabuh di pelabuhan tersebut, penumpang_naik menunjukkan jumlah penumpang yang pergi meninggalkan pelabuhan tersebut, dan penumpang_turun menunjukkan jumlah penumpang yang menginjakkan kaki di pelabuhan tersebut.

Dari seluruh atribut yang tersedia di atas, yang akan digunakan adalah atribut kapal_tiba dan penumpang_turun untuk memantau aktivitas kapal dan penumpang yang masuk ke pelabuhan Kepulauan Seribu.

3.2 Pra-Pemrosesan Data

Pada tahapan ini, akan dilakukan pra-pemrosesan data. Data cleaning atau pembersihan data merupakan proses untuk memastikan bahwa data yang akan kita analisa memiliki kualitas yang tinggi dengan cara memeriksa jika terdapat data yang missing dan tidak konsisten, menghapus duplikasi data, dan memperbaiki kesalahan pada data. Data cleaning dilakukan bertujuan agar proses data visualization yang dihasilkan valid dan tidak terjadi kesalahan.

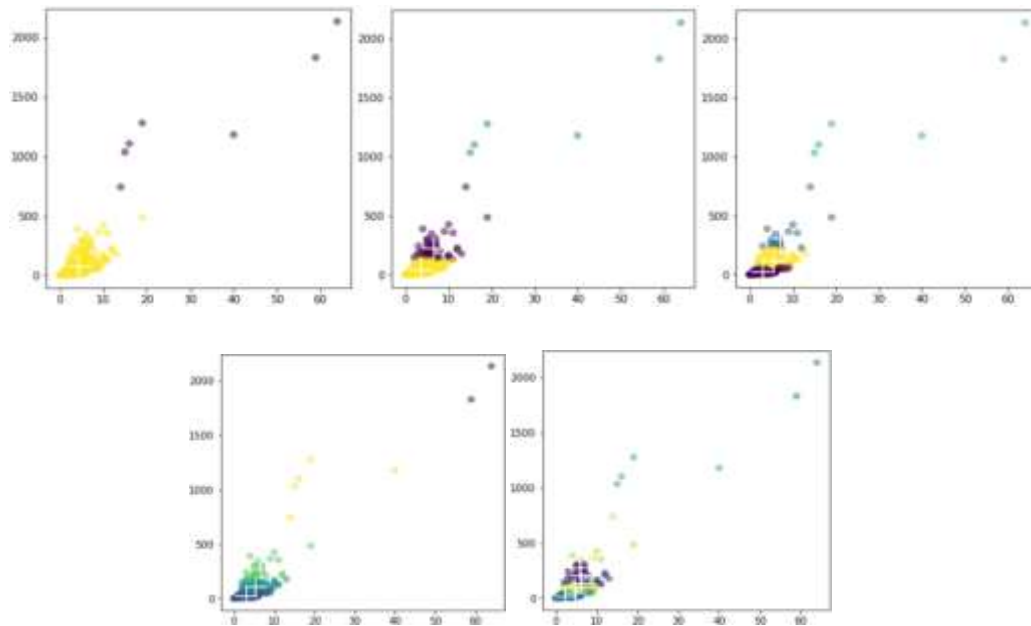
Dalam penelitian ini, yang akan dilakukan setelah berhasil meng-import data adalah melakukan check missing value untuk mengecek apakah terdapat kolom atau baris yang kosong atau tidak memiliki nilai. Setelah dilakukan pengecekan menggunakan metode `info()`, tidak ditemukan adanya missing value dalam data set tersebut. Sehingga data set ini sudah bisa digunakan untuk penelitian.

3.3 Eksperimen

Data set yang digunakan untuk mendukung penelitian ini adalah data set penumpang kapal dari dan ke Kepulauan Seribu Januari 2021 yang memiliki 6 variabel (tanggal, pelabuhan, kapal_berangkat, kapal_tiba, penumpang_naik, penumpang_turun) dan 310 jumlah sampel. Dengan data ini, dapat diperoleh informasi jumlah aktivitas kapal dan penumpang yang keluar masuk dari 10 pelabuhan di Kepulauan Seribu selama Januari 2021.

Pada penelitian ini, variabel yang akan dipakai sebagai fokus utama adalah variabel kapal_tiba dan penumpang_turun. Kemudian akan dilakukan uji coba *clustering* sebanyak 5 kali. Pada percobaan pertama data tersebut dikelompokkan menjadi 2 *cluster*, kedua menjadi 3 *cluster*, ketiga menjadi 4 *cluster*, keempat menjadi 5 *cluster*, dan terakhir menjadi 10 *cluster*.

Gambar (1) yang dilampirkan di bawah ini menunjukkan hasil dari percobaan *clustering* tersebut.



Gambar 1 Hasil Plot Clustering Fuzzy C-Means (2, 3, 4, 5, 10 Cluster)

- Evaluasi

Evaluasi dilakukan dengan menghitung nilai rata-rata *silhouette* dari tiap-tiap jumlah *cluster* yang dibuat untuk mencari nilai terbaik. Nilai rata-rata *silhouette* dapat terlihat dari yang dilampirkan pada Tabel (1).

Tabel 1 Nilai rata-rata *silhouette* dari tiap percobaan *clustering*.

Jumlah <i>Cluster</i>	Rata-Rata <i>Silhouette</i>
2	0.9228549970193373
3	0.6822397522307128
4	0.6110398244761042
5	0.6169357546017968
10	0.5809119253699172

Dari hasil tersebut, diketahui bahwa nilai rata-rata *silhouette* terbaik ada pada *cluster* 2 dengan nilai yang paling mendekati 1. Oleh karena itu, *cluster* 2 akan dipakai sebagai jumlah *cluster* yang akan dilanjutkan ke tahap analisis pada penelitian ini. Analisis

Hasil analisis yang didapatkan dari penelitian dengan menggunakan jumlah *cluster* 2 adalah sebagai berikut:

- *Cluster* 0 merupakan data yang memiliki jumlah kapal tiba dalam rentang 14 sampai 64, dengan jumlah penumpang turun mulai dari 743 sampai 2134.
- *Cluster* 1 merupakan data yang memiliki jumlah kapal tiba dalam rentang 0 sampai 19, dengan jumlah penumpang turun mulai dari 0 sampai 484.
- Kondisi tersibuk pada tanggal 1 Januari 2021 dengan 64 kapal tiba dan 2134 penumpang turun di Pelabuhan Untung Jawa, 19 kapal tiba dan 1276 penumpang turun di Pelabuhan Muara Angke. Serta 2 Januari 2021 dengan 59 kapal tiba dan 1828 penumpang turun di Pelabuhan Untung Jawa, 16 kapal tiba dan 1101 penumpang turun di Pelabuhan Muara Angke.

- Kondisi tersepi tercatat ada pada tanggal 1, 3, 6, 15, 17, 18, 19, 20, 23, 26, 28, 31 Januari 2021 di Pelabuhan Sabira, tanggal 18 dan 19 di Pelabuhan Tidung/ Payung, tanggal 19 dan 24 Januari 2021 di Pelabuhan Untung Jawa. Tidak ada kapal dan penumpang yang tiba.

4. KESIMPULAN

Kesimpulan dari penelitian ini adalah:

1. Metode *fuzzy c-means* berhasil diimplementasikan pada *clustering* data penumpang kapal dari dan ke Kepulauan Seribu.
2. Pelabuhan yang menerima kedatangan kapal terbanyak adalah Pelabuhan Untung Jawa, lebih tepatnya di tanggal 1 Januari 2021 yang menerima 64 kapal tiba dengan 2134 penumpang turun dan tanggal 2 Januari 2021 yang menerima 59 kapal masuk dengan 1828 penumpang turun. Disusul oleh Pelabuhan Muara Angke yang menerima 1276 penumpang turun dari 19 kapal tiba di tanggal 1 Januari 2021 dan di 2 Januari 2021, pelabuhan tersebut menerima 1101 penumpang turun dari 16 kapal.
3. Pelabuhan yang memiliki aktivitas paling rendah adalah Pelabuhan Sabira yang sama sekali tidak terlihat adanya aktivitas kapal selama 12 hari, tepatnya pada tanggal 1, 3, 6, 15, 17, 18, 19, 20, 23, 26, 28, 31 Januari 2021.

DAFTAR PUSTAKA

- [1] Abdy, M. (2015). Pengklasteran dengan Algoritma Fuzzy C-Means. *Jurnal Matematika, Statistika dan Komputasi*, 12(1), 30-35.
- [2] Gurning, R., & Budiyanto, H. E. (2007). *Manajemen Bisnis Pelabuhan*. Surabaya: PT. Andika Prasetya Ekawahana.
- [3] Handoyo, R., Mangkudjaja, R. R., & Nasution, S. M. (2014). Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen. *Jurnal Sifo Mikroskil*, 15(2), 73-82.
- [4] Kramadibrata, S. (1985). *Perencanaan Pelabuhan*. Bandung: Ganeca Exact.
- [5] Kusumadewi, S., & Purnomo, H. (2010). *Aplikasi Logika Fuzzy Untuk Pendukung Keputusan*. Yogyakarta: Graha Ilmu.
- [6] Luthfi, E. T. (2007). Fuzzy C-Means untuk Clustering Data (Studi Kasus: Data Performance Mengajar Dosen). *Seminar Nasional Teknologi*, (pp. 1-7).
- [7] Pramesti, D. F., Furqon, M. T., & Dewi, C. (2017). Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas (Hotspot). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e- ISSN, 2548, 964X.*, 723-732.
- [8] Suranto, S. (2004). *Manajemen Operasional Angkutan Laut dan Kepelabuhanan Serta Prosedur Impor Barang*. Jakarta: PT Gramedia Pustaka Utama.
- [9] Talley, W. K. (2009). *Port Economics*. New York: Routledge. Triatmodjo, B. (2009). *Perencanaan Pelabuhan*. Yogyakarta: Beta Offset.