

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pentingnya pengelolaan sumber daya ikan di Provinsi Jawa Timur, Indonesia. Provinsi Jawa Timur dikenal sebagai salah satu wilayah dengan potensi perikanan yang tinggi, dengan berbagai jenis ikan dan sumber daya laut yang melimpah [1]. Namun, pengelolaan sumber daya ikan yang efektif memerlukan pemahaman yang mendalam tentang karakteristik daerah-daerah potensi ikan di wilayah tersebut. Peran sektor perikanan di Jawa Timur sangat vital, karena sebagian besar penduduknya, khususnya para nelayan, mengandalkan sektor ini sebagai sumber penghasilan utama. Selain itu, sektor perikanan juga memiliki dampak positif dalam mendorong pertumbuhan ekonomi provinsi Jawa Timur. Wilayah ini memiliki potensi yang besar dalam industri perikanan, baik itu dalam penangkapan ikan maupun budidaya. Selama beberapa tahun terakhir, produktivitas sektor perikanan di provinsi ini terus meningkat. Berdasarkan data dari Dinas Kelautan dan Perikanan (DKP) Jawa Timur, produksi perikanan tangkap pada tahun 2020 mencapai 134.194 ton. Sementara itu, produksi perikanan budidaya mencapai jumlah yang lebih impresif, yaitu 539.568 ton. Dengan menggabungkan kedua jenis produksi tersebut, total produksi perikanan mencapai angka mendekati 674.000 ton. Kontribusi sektor perikanan terhadap Produk Domestik Bruto (PDRB) Jawa Timur pada tahun 2020 mencapai Rp 50,990 miliar atau sekitar 2,53% dari total PDRB provinsi tersebut. Dari banyaknya wilayah di Jawa Timur sehingga membutuhkan metode pengelompokan data yang sesuai dengan indikator tertentu untuk mempermudah pendistribusian dan penjualan hasil dari industri perikanan tangkap dan juga budidaya. Bukan hanya mempermudah pengelompokan saja dengan adanya metode ini membantu dalam pemahaman data, segmentasi pasar, pengenalan pola, dan berbagai aplikasi lain dalam analisis data. Metode pengelompokan data atau *clustering* yang umum digunakan adalah *K-Means Clustering*, *Self-Organizing Maps (SOM)*, dan *Agglomerative Hierarchical Clustering*.

K-Means adalah salah satu algoritma *clustering* yang paling umum digunakan. Tujuan utama dari algoritma ini adalah untuk mengelompokkan data

menjadi beberapa kelompok, atau kluster, di mana setiap kluster memiliki pusatnya sendiri. Algoritma *K-Means* bekerja dengan cara menginisialisasi pusat-pusat kluster secara acak, kemudian mengalokasikan setiap data ke kluster terdekat berdasarkan jarak (sering menggunakan jarak *Euclidean*), lalu memperbarui pusat kluster dengan rata-rata data dalam kluster tersebut. Proses ini diulang hingga pusat kluster tidak berubah secara signifikan atau telah mencapai jumlah iterasi tertentu [2]. Meskipun algoritma *K-Means* relatif sederhana dan cepat, hasilnya dapat sangat tergantung pada inisialisasi pusat kluster, dan algoritma ini dapat terjebak dalam solusi lokal yang tidak optimal [3]. Oleh karena itu, seringkali perlu menjalankan algoritma beberapa kali dengan inisialisasi yang berbeda untuk memastikan hasil yang lebih baik.

Self-Organizing Maps (SOM), juga dikenal sebagai Kohonen Maps, adalah algoritma klustering yang digunakan untuk analisis data dan pemetaan. Ini berbeda dari algoritma klustering tradisional seperti *K-Means* karena menghasilkan pemetaan berdimensi rendah dari data ke dalam ruang dua dimensi atau tiga dimensi yang biasanya disebut "peta" atau "grid" [4]. SOM bekerja dengan memetakan data input ke dalam grid yang biasanya terdiri dari neuron. Setiap neuron mewakili suatu kelompok atau kluster. Pada awalnya, bobot dari setiap neuron diinisialisasi secara acak. Kemudian, data input dipetakan ke neuron yang memiliki bobot terdekat berdasarkan jarak *Euclidean*. Selanjutnya, bobot neuron yang terpilih dan neuron sekitarnya diperbarui untuk menyesuaikan diri dengan data input [5]. Proses ini diulang dengan data input berulang kali hingga grid konvergen.

Agglomerative Hierarchical Clustering adalah salah satu metode dalam analisis klustering yang mengikuti pendekatan berhirarki. Pendekatan ini dimulai dengan menganggap setiap titik data sebagai satu kluster dan kemudian secara iteratif menggabungkan kluster-kluster yang paling mirip satu sama lain hingga semua data tergabung menjadi satu kelompok besar [6]. Proses ini menciptakan struktur berhirarki yang dapat divisualisasikan dalam bentuk pohon atau dendrogram [7]. *Agglomerative hierarchical clustering* memiliki beberapa kelebihan yang membuatnya menjadi pilihan yang baik dalam analisis *clustering*, Hierarki yang dapat dianalisis, interpretabilitas yang baik dengan dendrogram [8].

Cocok untuk data multivariat. Algoritma ini dapat digunakan untuk data dengan banyak fitur (data multivariat) dan berbagai jenis jarak atau metrik kesamaan. Kemampuan menerapkan berbagai metrik jarak yang sesuai dengan jenis data, seperti euclidean, manhattan, cosine, dan lain-lain. Hasil *agglomerative hierarchical clustering* relatif stabil, yang berarti tidak terlalu dipengaruhi oleh inisialisasi awal. Ini berbeda dengan beberapa algoritma *clustering* yang mungkin menghasilkan solusi yang berbeda setiap kali dijalankan [9].

1.2 Perumusan Masalah

Rumusan Masalah pada penelitian ini adalah: “Bagaimana penerapan metode *Agglomerative Hierarchy Clustering* dalam mengelompokkan daerah potensi ikan di Provinsi Jawa Timur berdasarkan fitur-fitur yang relevan dan bagaimana hasil *clustering* tersebut dapat memberikan informasi berharga untuk pengelolaan sumber daya ikan di wilayah tersebut”.

1.2.1 Permasalahan

Permasalahan dalam penelitian ini adalah adanya kebutuhan untuk mengidentifikasi dan mengelompokkan daerah-daerah potensi ikan di Provinsi Jawa Timur agar dapat lebih efektif dalam perencanaan pengembangan sektor perikanan.

1.2.2 Usulan Solusi

Usulan solusinya adalah menggunakan metode *Agglomerative Hierarchy Clustering* untuk mengelompokkan daerah-daerah tersebut berdasarkan karakteristik potensi perikanan yang dimilikinya. Dengan metode ini, diharapkan dapat ditemukan pola-pola atau kelompok-kelompok yang dapat memberikan panduan kepada pemerintah dan *stakeholder* terkait dalam mengambil keputusan terkait pengelolaan sumber daya perikanan di wilayah ini. dalam penelitian ini akan dicari *cluster* terbaik menggunakan berbagai metode *linkage* yang tersedia dalam *Agglomerative Hierarchy Clustering*, seperti *single linkage* dan *double linkage*, Tujuan dari pencarian ini adalah untuk mengevaluasi berbagai kemungkinan pengelompokan dan menentukan metode *linkage* yang paling sesuai untuk menggambarkan pola-pola dalam data potensi ikan di Provinsi Jawa Timur. Setiap metode *linkage* memiliki karakteristik yang berbeda dalam membentuk *cluster*, dan

pemilihan metode *linkage* yang tepat akan mempengaruhi hasil akhir dari analisis *clustering*. Setelah terbentuknya *cluster-cluster* tersebut, analisis hasil *cluster* akan dilakukan dengan mengukur *mean* (rata-rata) dan standar deviasi (deviasi standar) dari berbagai fitur yang digunakan dalam pengelompokan, seperti produksi ikan, jenis ikan yang dominan, atau variabel lain yang relevan. Analisis ini akan memberikan pemahaman yang lebih dalam tentang perbedaan antara *cluster-cluster* yang terbentuk, termasuk sejauh mana setiap *cluster* memiliki karakteristik yang serupa atau berbeda dalam hal potensi perikanan. *Mean* akan memberikan gambaran tentang nilai tengah dari setiap fitur dalam suatu *cluster*, sementara standar deviasi akan menunjukkan sejauh mana data dalam *cluster* tersebut tersebar dari nilai *mean*. Informasi ini akan membantu dalam membedakan *cluster-cluster* yang ada dan memberikan wawasan tentang keragaman dalam potensi perikanan di daerah-daerah tersebut.

1.3 Pertanyaan penelitian

Pertanyaan penelitian utama yang akan dijawab dalam penelitian ini adalah:

1. Berapa nilai *error DBI* karena pengaruh perubahan parameter pada metode *Agglomerative Hierarchy Clustering*?
2. Berapa nilai perbandingan kinerja metode *Agglomerative Hierarchy Clustering* dengan menggunakan perbedaan *single linkage* dan *double linkage* untuk mengelompokkan daerah-daerah potensi ikan di Provinsi Jawa Timur?

1.4 Batasan Masalah

Berikut ini adalah batasan masalah pada penelitian ini:

1. Data yang digunakan didalam penelitian ini adalah data yang diperoleh peneliti dari Dinas Kelautan dan Perikanan Provinsi Jawa timur tahun 2019.
2. Data yang digunakan terdiri dari 254 kecamatan dengan 11 fitur, yakni jenis ikan (ton) yang meliputi (Cakalang, Tongkol, Gabus, Lele, Mujair, Kakap Merah/Bambangan, Kerapu, Nila, Lemuru, Rajungan, Kepiting dan Tawes).
3. Metode *Clustering* yang digunakan adalah *Agglomerative Hierarchy Clustering* dengan jenis *single linkage* dan *double linkage*.
4. Evaluasi hasil *clustering* menggunakan *Davies-Bouldin Index (DBI)*.
5. Analisa klaster menggunakan *mean* dan standar deviasi.

1.5 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah:

1. Menerapkan metode *Agglomerative Hierarchy Clustering* untuk melakukan *clustering* terhadap daerah potensi ikan di Provinsi Jawa Timur berdasarkan fitur-fitur yang relevan.
2. Membentuk *cluster-cluster* daerah potensi ikan yang memiliki karakteristik serupa berdasarkan hasil dari proses *clustering* menggunakan metode *Agglomerative Hierarchy Clustering*.
3. Menganalisis hasil *clustering* untuk mendapatkan *insight* tentang pola distribusi daerah potensi ikan di Provinsi Jawa Timur dan memberikan informasi yang berharga bagi pengambilan keputusan dalam pengelolaan sumber daya ikan di wilayah tersebut.

Manfaat yang saling berkaitan dari penelitian ini adalah:

1. Penelitian ini akan memberikan cara efisien dan objektif dalam mengidentifikasi daerah-daerah potensi ikan di Provinsi Jawa Timur dengan menggunakan metode *clustering*, menggantikan pendekatan manual yang rumit dan memakan waktu.
2. Pembentukan *cluster-cluster* daerah potensi ikan akan membantu para pengambil keputusan, seperti pihak pemerintah dan lembaga terkait, untuk memahami pola distribusi dan karakteristik wilayah potensi ikan, sehingga dapat mengembangkan strategi dan kebijakan yang lebih tepat untuk pengelolaan sumber daya ikan secara berkelanjutan.
3. Dengan adanya informasi yang lebih jelas tentang daerah potensi ikan, penelitian ini akan mendukung upaya dalam pengelolaan sumber daya ikan di Provinsi Jawa Timur, termasuk dalam aspek konservasi, pengawasan, dan pengembangan sektor perikanan, guna memastikan pemanfaatan yang berkelanjutan dan melindungi keberlanjutan ekosistem laut.

1.6 Sistematika Penulisan

Proposal tugas akhir ini terdiri dari lima bab yang masing-masing bab memiliki rangkaian sistematis yang mengikuti perumusan masalah yang telah dicantumkan pada BAB I. Berikut adalah sistematika proposal tersebut:

BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, rumusan masalah, solusi permasalahan, batasan masalah, tujuan dan manfaat dilakukannya penelitian, dan sistematika proposal

BAB II TINJAUAN PUSTAKA

Bab ini membahas konsep dasar *clustering* dan pentingnya penerapannya dalam pengelolaan sumber daya ikan. Metode *Agglomerative Hierarchy Clustering* dijelaskan lebih lanjut, termasuk kelebihan dan keterbatasannya dalam konteks penelitian ini. Bab ini juga menguraikan beberapa contoh penggunaan *clustering* dalam bidang pengelolaan sumber daya ikan.

BAB III METODE PENELITIAN

Bab ini menjelaskan tentang arsitektur sistem, data preparation, langkah-langkah pada setiap metode usulan, metode evaluasi, skenario pengujian, dan jadwal penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan tentang hasil, pembahasan, dan analisis dari implementasi penelitian yang dilakukan.

BAB V PENUTUP

Bab ini menjelaskan tentang kesimpulan dari hasil penelitian yang dilakukan dan saran pengembangan untuk penelitian ini kedepannya.