Bab I

Pendahuluan

1.1 Latar Belakang

Komoditas adalah benda atau barang yang diperdagangkan, dapat disimpan dalam waktu yang tertentu (Ester Kojongian 2017). Pasar komoditas di Jawa Timur memiliki peran penting dalam mendukung perekonomian daerah dan nasional, terutama melalui kontribusinya sebagai penghasil utama berbagai komoditas agraris seperti padi, jagung, bawang merah, cabai, dan berbagai produk hortikultura lainnya. Menurut (Fackler. & Goodwin 2001) integrasi pasar adalah tingkat pergerakan harga pada wilayah yang berbeda, dimana produk yang sama akan memiliki harga yang sama, meskipun dijual pada tempat yang berbeda dan sinyal harga serta informasi pasar ditransmisikan secara merata (M.Ghosh 2000). Dinamika pasar di Jawa Timur tidak hanya mencerminkan potensi besar dalam pemenuhan kebutuhan pangan nasional tetapi juga menunjukkan tantangan serius berupa fluktuasi harga yang berulang. Fluktuasi ini dapat dipengaruhi oleh berbagai faktor, seperti perubahan musim panen, permintaan pasar, kualitas infrastruktur distribusi, kebijakan pemerintah, hingga gangguan eksternal seperti bencana alam. Harga komoditas di Jawa Timur memainkan peran penting dalam perekonomian daerah dan nasional, mengingat wilayah ini merupakan salah satu pusat produksi dan distribusi utama di Indonesia. Sebagai provinsi dengan kontribusi signifikan terhadap PDB nasional, dinamika harga komoditas di Jawa Timur dipengaruhi oleh berbagai faktor, termasuk produksi lokal, distribusi, dan perubahan harga global. Misalnya, studi oleh (Cahyaningrum 2021) menunjukkan bahwa dinamika harga global sangat memengaruhi inflasi inti di Jawa Timur melalui komposit harga global yang mencakup komoditas makanan seperti gandum dan minyak kelapa sawit.

Fenomena fluktuasi harga yang tinggi di Jawa Timur juga dapat diperburuk oleh ketimpangan distribusi komoditas antara daerah produsen dan konsumen. Daerah yang menjadi sentra produksi sering kali menghadapi tantangan distribusi yang menyebabkan *oversupply* lokal sehingga harga jatuh, sementara wilayah konsumen harus menghadapi harga yang lebih tinggi akibat biaya distribusi yang mahal. Ketimpangan ini memperlihatkan perlunya analisis pola

harga berbasis data untuk mengidentifikasi wilayah yang mengalami ketidakseimbangan dan pola distribusi yang perlu dioptimalkan. Selain itu, pengaruh musim dan cuaca juga menjadi faktor dominan yang memengaruhi harga komoditas agraris di Jawa Timur, Studi oleh (Hartono et al. 2023) mengungkapkan bahwa harga komoditas hortikultura seperti cabai dan bawang merah sering kali berfluktuasi tinggi, terutama karena sifatnya yang mudah rusak dan ketergantungan pada musim produksi. Pola musiman ini menciptakan fluktuasi harga yang dapat diprediksi tetapi sering kali sulit untuk dikelola tanpa dukungan data analitis yang tepat.

Dalam beberapa tahun terakhir, pandemi COVID-19 memperburuk volatilitas harga dengan mengganggu distribusi dan menciptakan ketidakpastian pasar. Sebuah studi oleh (Roy et al. 2020) mencatat bahwa pembatasan sosial selama pandemi meningkatkan kesenjangan harga dan fluktuasi komoditas strategis seperti bawang merah di Jawa Timur. Dalam upaya memahami pola harga pasar di Jawa Timur, analisis berbasis *clustering* atau pengelompokan data telah menjadi metode yang efektif. Metode ini memungkinkan data dikelompokkan berdasarkan kesamaan karakteristik tertentu, seperti pola harga, volume distribusi, atau lokasi geografis. Dengan demikian, pola-pola tersembunyi dalam data dapat diungkap untuk memberikan wawasan strategis bagi para pemangku kepentingan.

K-Means merupakan salah satu algoritma pengelompokan yang paling populer karena kesederhanaan dan efisiensinya dalam mengelompokkan data. Algoritma ini bekerja dengan membagi data menjadi beberapa *cluster* berdasarkan kedekatan jarak antara titik data dan *centroid* yang mewakili *cluster*. Keunggulan utama K-Means adalah kemampuannya dalam mengidentifikasi pola umum pada dataset yang terstruktur dan memiliki distribusi yang relatif beragam. Namun, keterbatasan utama algoritma ini terletak pada sensitivitasnya terhadap outlier, di mana data yang menyimpang dapat menggeser posisi *centroid* secara signifikan dan mengurangi akurasi hasil pengelompokan (Sinaga & Yang 2020).

Sebaliknya, DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) menawarkan kemampuan untuk mendeteksi cluster dengan bentuk yang tidak beraturan dan lebih toleran terhadap *outlier*. Algoritma ini sangat cocok untuk data yang tidak teratur atau penuh dengan *noise*, seperti pola harga musiman yang tidak konsisten akibat permintaan dan distribusi yang dinamis. DBSCAN juga tidak memerlukan jumlah *cluster* yang telah ditentukan sebelumnya, membuatnya lebih fleksibel dalam menangani data yang kompleks seperti pola harga pasar di Jawa Timur. Hal ini menjadikannya sangat efektif untuk digunakan pada dataset yang kompleks, seperti pola harga komoditas di pasar Jawa Timur, di mana fluktuasi harga dapat sangat beragam dan tidak terstruktur dengan baik (Gholizadeh et al. 2020)

Selain K-Means dan DBSCAN, SOM (*Self-Organizing Maps*) merupakan

algoritma yang menawarkan keunggulan unik dalam memvisualisasikan data berdimensi tinggi ke dalam peta dua dimensi. SOM memungkinkan analisis visual yang lebih intuitif terhadap data yang kompleks, seperti pola harga berdasarkan musim, lokasi geografis, dan karakteristik pasar lainnya. Kemampuan ini membuat SOM ideal untuk menganalisis data pasar komoditas yang melibatkan banyak variabel, seperti jenis komoditas, lokasi pasar, dan periode waktu (Kuo et al. 2005).

Penerapan kedua algoritma ini diharapkan dapat memberikan wawasan baru tentang dinamika pasar di Jawa Timur. Dengan menggunakan DBSCAN, pola harga pasar yang tidak teratur dapat dikelompokkan secara akurat, sedangkan dengan SOM, pola multidimensi dalam data dapat divisualisasikan dengan jelas. Kombinasi ini memungkinkan identifikasi pasar dengan karakteristik harga serupa, pendeteksian anomali harga yang mencurigakan, serta penyusunan rekomendasi kebijakan yang lebih berbasis data. Dalam penelitian (Fidayani et al. 2017) menunjukkan bahwa integrasi informasi pasar jagung di Jawa Timur dapat meningkatkan efisiensi distribusi komoditas dan mengurangi volatilitas harga.

Dengan memahami kelebihan dan kekurangan algoritma K-Means, DBSCAN, dan SOM, penelitian ini bertujuan untuk mengevaluasi efektivitas ketiga algoritma tersebut dalam mengelompokkan pola harga komoditas di Jawa Timur. Analisis ini diharapkan dapat memberikan kontribusi signifikan dalam mendukung kebijakan stabilisasi harga, perbaikan distribusi, serta pengelolaan pasar yang lebih adaptif dan berbasis data. Penelitian ini juga memberikan wawasan yang relevan bagi para pemangku kebijakan, pelaku pasar, dan masyarakat umum untuk mengelola tantangan fluktuasi harga yang menjadi masalah utama dalam pasar komoditas di Jawa Timur.

1.2 Perumusan Masalah

Berikut rumusan masalah yang ingin saya angkat adalah

1. Bagaimana kinerja algoritma K-Means, DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*), dan SOM (*Self-Organizing Maps*) dalam mengelompokkan pasar berdasarkan pola harga komoditas di Jawa Timur?
2. Metode mana yang memberikan hasil terbaik dalam hal akurasi, interpretasi, dan relevansi hasil pengelompokan?

1.3 Tujuan Penelitian

Berikut adalah tujuan yang ingin dicapai pada penulisan proposal/TA.

1. Menganalisis dan membandingkan kinerja algoritma K-Means, DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*), dan SOM

(*Self-Organizing Maps*) untuk pengelompokan pasar berdasarkan pola harga komoditas di Jawa Timur.

2. Menentukan metode yang paling optimal dalam konteks penelitian ini berdasarkan hasil evaluasi.

1.4 Batasan dan Asumsi Penelitian

Penelitian ini memiliki beberapa batasan dan asumsi yang dirancang untuk memastikan fokus dan keberlanjutan analisis. Batasan pertama adalah data yang digunakan dalam penelitian ini terbatas pada pola harga komoditas tertentu di wilayah Jawa Timur. Penelitian ini dibatasi pada analisis pola harga komoditas bahan pokok sembako yang terdiri dari beras medium (kg), minyak goreng curah (liter), telur ayam ras (kg), susu kental manis merk Indomilk (370 gr/kr1), bawang merah (kg), daging ayam ras (kg), garam beryodium halus (kg), gas LPG 3 kg, serta gula kristal putih (kg). Data yang digunakan mencakup periode satu tahun dalam harian. Analisis dilakukan pada data yang dikumpulkan dari sumber situs Siskaperbapo Jawa Timur, yang relevan dengan penelitian ini. Pemilihan data ini didasarkan pada ketersediaan informasi yang relevan dan mencerminkan variasi harga di berbagai pasar tradisional di Jawa Timur. Penelitian ini tidak mempertimbangkan faktor-faktor eksternal, seperti perubahan kebijakan ekonomi, kondisi cuaca, atau tren global, yang dapat memengaruhi pola harga komoditas.

Penelitian ini mengasumsikan bahwa data mentah memerlukan proses pra-pemrosesan sebelum tahap modeling dilakukan. Proses ini meliputi pembersihan data untuk menghilangkan anomali, penanganan data yang hilang, dan normalisasi agar data berada dalam skala yang konsisten. Dengan demikian, kualitas data dapat terjamin, dan hasil analisis menjadi lebih valid. Selain itu, asumsi lainnya adalah bahwa setiap metode pengelompokan (K-Means, DB-SCAN, dan SOM) dapat diimplementasikan secara optimal dengan parameter yang disesuaikan untuk mendapatkan hasil terbaik dalam konteks data yang dianalisis.

Penelitian ini juga mengasumsikan bahwa metrik evaluasi, seperti *Silhouette Score*, *Davies-Bouldin Index*, dan visualisasi kluster, cukup untuk mengukur kinerja masing-masing metode. Oleh karena itu, hasil pengelompokan diinterpretasikan secara kuantitatif dan kualitatif berdasarkan metrik tersebut tanpa mempertimbangkan preferensi subjektif dari pihak tertentu. Dengan batasan dan asumsi ini, penelitian diharapkan dapat memberikan hasil yang valid dan relevan dalam konteks pengelompokan pola harga komoditas di Jawa Timur, meskipun tidak mencakup semua variabel dan faktor yang mungkin memengaruhi pola harga secara keseluruhan.

1.5 Manfaat Penelitian

Kontribusi dari tulisan ini adalah

1. Memberikan rekomendasi metode pengelompokan yang paling sesuai untuk menganalisis pola harga komoditas di Jawa Timur.
2. Membantu pengambil kebijakan dalam memahami segmentasi pasar dan menyusun strategi perdagangan yang lebih baik.
3. Menyediakan referensi bagi penelitian serupa di masa depan yang melibatkan pengelompokan data kompleks.
4. Mendukung pengembangan algoritma pengelompokan yang lebih adaptif dan efisien untuk aplikasi di bidang ekonomi dan perdagangan.

1.6 Sistematika Penulisan

Secara sistematis isi dari Tugas Akhir ini disusun sebagai berikut:

1. BAB I Pendahuluan

Bab ini menguraikan latar belakang penelitian, perumusan masalah, tujuan penelitian, manfaat penelitian, serta batasan dan asumsi yang digunakan. Penjelasan pada bab ini memberikan gambaran umum mengenai ruang lingkup penelitian yang berfokus pada pengelompokan pasar di Jawa Timur berdasarkan pola harga komoditas menggunakan algoritma K-Means, DBSCAN, dan *Self-Organizing Maps* (SOM).

2. BAB II Kajian Pustaka

Bab ini memaparkan teori-teori yang relevan dengan algoritma K-Means, DBSCAN, dan *Self-Organizing Maps* (SOM). Selain itu, bab ini juga mencakup hasil penelitian terdahulu yang berkaitan dengan topik ini untuk mendukung landasan teoritis penelitian.

3. BAB III Metode Penelitian

Bab ini menjelaskan secara rinci metode yang digunakan dalam penelitian, termasuk lokasi penelitian, data yang digunakan, alur penelitian, dan implementasi algoritma K-Means, DBSCAN, dan SOM. Selain itu, bab ini juga memaparkan metode evaluasi yang diterapkan untuk mengukur kinerja algoritma dalam pengelompokan data harga komoditas.

4. BAB IV Pengumpulan dan Pengolahan Data.

Bab ini menguraikan proses pengumpulan data dari situs Siskaperbapo dan sumber pendukung lainnya. Proses tersebut mencakup pembersihan data, normalisasi, serta eksplorasi data yang dilakukan untuk memastikan data siap digunakan dalam proses pengelompokan. Data meliputi informasi harga komoditas, lokasi pasar, dan waktu pengambilan data.

5. BAB V Analisis dan Pembahasan.

Bab ini menyajikan hasil dari pengelompokan yang dilakukan menggunakan algoritma K-Means, DBSCAN, dan SOM. Penilaian kinerja algoritma dilakukan dengan menggunakan metrik evaluasi seperti *Silhouette Score* dan *Davies-Bouldin Index*. Hasil yang diperoleh kemudian dianalisis untuk mengidentifikasi pola harga di berbagai pasar serta membahas implikasinya.

6. BAB VI Kesimpulan dan Saran

Bab ini merangkum temuan utama dari penelitian, termasuk algoritma yang memberikan hasil terbaik dalam pengelompokan data harga komoditas. Selain itu, diberikan rekomendasi untuk pengembangan penelitian lebih lanjut, seperti penggunaan data tambahan atau eksplorasi metode lain untuk meningkatkan akurasi dan efisiensi pengelompokan.

1.7 Rencana Kegiatan

Rencana kegiatan yang akan saya lakukan adalah sebagai berikut:

1. Studi literatur

Tahap awal penelitian ini dimulai dengan mengumpulkan informasi mendalam dari berbagai sumber seperti jurnal ilmiah, buku, dan laporan terkait teori serta penelitian yang relevan. Fokus utama pada tahap ini adalah memahami konsep dasar dan teknis algoritma K-Means, DBSCAN, dan SOM, serta mengeksplorasi studi kasus serupa yang telah dilakukan sebelumnya. Tujuannya adalah untuk mendapatkan wawasan menyeluruh mengenai keunggulan dan kelemahan masing-masing metode yang akan diterapkan dalam konteks pengelompokan pasar.

2. Pengumpulan Data Harga Pasar

Data pola harga komoditas yang akan digunakan dalam penelitian ini dikumpulkan dari berbagai sumber terpercaya, seperti data resmi instansi pemerintah, laporan pasar tradisional, atau basis data daring. Selain itu, diperlukan, metode seperti teknik *web scraping* juga akan digunakan untuk melengkapi data. Data yang diperoleh akan mencakup informasi harga komoditas tertentu di berbagai pasar di Jawa Timur dalam kurun waktu yang telah ditentukan.

3. Pra-pemrosesan Data

Setelah data terkumpul, langkah berikutnya adalah memastikan bahwa data siap untuk digunakan dalam analisis. Tahap ini melibatkan berbagai proses, seperti pembersihan data untuk menghilangkan duplikasi dan inkonsistensi, penanganan nilai yang hilang menggunakan metode

interpolasi atau estimasi lainnya, normalisasi data agar berada dalam skala yang sama, serta transformasi data untuk memenuhi format yang sesuai dengan algoritma yang akan diterapkan.

4. Implementasi Algoritma Pengelompokan

Pada tahap ini, algoritma K-Means, DBSCAN, dan SOM akan diterapkan pada data yang telah diproses. Implementasi ini melibatkan konfigurasi parameter untuk masing-masing metode, seperti menentukan jumlah kluster untuk K-Means, parameter epsilon dan minimum samples untuk DBSCAN, serta konfigurasi jaringan SOM seperti ukuran grid dan parameter pembelajaran. Proses ini bertujuan untuk menghasilkan kluster yang mencerminkan pola harga komoditas secara optimal.

5. Evaluasi Kinerja Metode

Hasil pengelompokan dari ketiga metode akan dievaluasi secara menyeluruh menggunakan metrik evaluasi yang relevan. Metrik ini mencakup *siluet score* untuk mengukur kohesi dan pemisahan antar kluster, *Davies-Bouldin Index* untuk menilai validitas kluster, serta analisis visual untuk mempermudah interpretasi hasil. Evaluasi ini bertujuan untuk menentukan sejauh mana setiap metode mampu menghasilkan pengelompokan yang akurat dan bermakna.

6. Analisis dan Perbandingan Hasil

Tahap akhir dari penelitian ini adalah melakukan analisis mendalam terhadap hasil pengelompokan. Hasil dari masing-masing metode akan dibandingkan untuk mengidentifikasi metode yang paling optimal dalam konteks penelitian ini. Selain itu, analisis ini juga mencakup perbandingan dengan penelitian sebelumnya untuk menilai konsistensi hasil serta mengidentifikasi area yang masih memerlukan perbaikan untuk pengembangan penelitian di masa mendatang.