BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Indonesia, sebagai negara agraris dengan populasi lebih dari 270 juta jiwa, sangat bergantung pada sektor pertanian sebagai sumber utama mata pencaharian. Salah satu komoditas utama yang dihasilkan adalah padi, yang berperan penting dalam menjaga ketahanan pangan nasional. Berdasarkan data BPS (2023), produksi padi nasional mencapai 54,6 juta ton GKG, menjadikan Indonesia sebagai salah satu produsen beras terbesar di kawasan Asia Tenggara. Meskipun demikian, sektor pertanian, khususnya di Provinsi Sumatera Utara, masih menghadapi berbagai tantangan dalam hal produktivitas, efisiensi pengelolaan lahan, dan adaptasi terhadap perubahan iklim. Tantangan tersebut disebabkan oleh keterbatasan adopsi teknologi pertanian modern, distribusi pupuk yang belum merata, serta minimnya infrastruktur irigasi di beberapa wilayah pedesaan.

Perbedaan kondisi geografis, iklim, dan sistem pengelolaan pertanian di setiap kabupaten menyebabkan ketimpangan hasil produksi yang signifikan. Sebagai contoh, produktivitas padi di Kabupaten Deli Serdang mencapai rata-rata 5,7 ton/ha, sedangkan di Mandailing Natal hanya sekitar 4,2 ton/ha (BPS, 2023). Kondisi ini menunjukkan adanya ketidakseimbangan dalam pengelolaan pertanian akibat distribusi sumber daya yang tidak merata, seperti pupuk, air irigasi, dan alat pertanian yang kurang sesuai dengan karakteristik lahan. Oleh karena itu,

dibutuhkan pendekatan analitik berbasis data untuk memahami karakteristik wilayah pertanian secara lebih mendalam dan terukur.

Pendekatan berbasis data (data-driven decision making) memungkinkan pemerintah dan pelaku pertanian merancang kebijakan yang lebih efisien dan tepat sasaran. Salah satu teknologi yang relevan dalam konteks ini adalah machine learning, yang mampu mengolah data besar dan menemukan pola tersembunyi yang sulit diidentifikasi secara manual. Machine learning memiliki cabang unsupervised learning, di mana data tanpa label dapat dikelompokkan berdasarkan kesamaan karakteristik. Dalam konteks ini, metode K-Means clustering digunakan karena mampu mengidentifikasi kelompok wilayah pertanian dengan kondisi homogen berdasarkan luas panen, produktivitas, dan total hasil produksi.

Proses segmentasi wilayah menggunakan K-Means clustering menghasilkan klaster dengan kondisi pertanian yang serupa. Hasil segmentasi ini dapat digunakan sebagai dasar dalam perumusan kebijakan pertanian yang lebih spesifik, seperti distribusi bantuan, penentuan teknologi tepat guna, atau alokasi sumber daya secara proporsional. Studi Li et al. (2022) menunjukkan bahwa penerapan metode K-Means clustering dalam agricultural zoning mampu meningkatkan efisiensi pengelolaan lahan hingga 15% dibanding pendekatan tradisional. Dengan demikian, hasil segmentasi ini menjadi landasan bagi kebijakan berbasis bukti (evidence-based policy) yang mendukung pemerataan produktivitas serta kesejahteraan petani di Sumatera Utara.

Keterbaruan penelitian ini terletak pada penerapan metode K-Means clustering untuk segmentasi wilayah pertanian di Indonesia, khususnya di Provinsi

Sumatera Utara. Sebagian besar penelitian terdahulu berfokus pada prediksi hasil panen tanpa mempertimbangkan segmentasi wilayah secara spasial dan terperinci. Misalnya, penelitian Prasetyo dan Nurhalimah (2021) hanya menitikberatkan pada prediksi hasil pertanian menggunakan Random Forest tanpa menganalisis pembagian wilayah berbasis klaster. Oleh karena itu, penelitian ini memberikan kontribusi baru dengan pendekatan yang lebih lokal dan aplikatif untuk mendukung kebijakan pertanian berbasis data serta peningkatan ketahanan pangan daerah.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, penelitian ini bertujuan untuk memberikan solusi terhadap permasalahan pengelolaan pertanian di Provinsi Sumatera Utara dengan menggunakan metode K-Means clustering. Untuk itu, masalah utama yang dihadapi dalam pengelolaan pertanian di wilayah ini perlu dirumuskan dalam bentuk pertanyaan penelitian yang logis dan terstruktur. Pertanyaan-pertanyaan ini akan menjadi landasan bagi analisis lebih lanjut dalam penelitian ini.

1. Bagaimana penerapan metode K-Means Clustering untuk mengelompokkan wilayah pertanian di Provinsi Sumatera Utara berdasarkan variabel luas panen, produktivitas, dan produksi padi tahun 2023?
2. Bagaimana karakteristik kapasitas produksi padi pada setiap klaster yang terbentuk berdasarkan nilai centroid luas panen, produktivitas, dan

produksi padi hasil pengelompokan menggunakan metode K-Means Clustering?

3. Bagaimana hasil segmentasi wilayah pertanian menggunakan metode K-Means Clustering dapat dimanfaatkan sebagai informasi awal dalam menentukan prioritas pengembangan pertanian berbasis data di Provinsi Sumatera Utara?

1.3. Tujuan Penelitian

Tujuan utama dari penelitian ini adalah untuk:

1. Menerapkan algoritma K-Means Clustering untuk melakukan segmentasi wilayah pertanian di Provinsi Sumatera Utara berdasarkan variabel luas panen, produktivitas, dan produksi padi.
2. Mengidentifikasi karakteristik setiap klaster yang terbentuk berdasarkan hasil pengelompokan wilayah pertanian.
3. Mengetahui potensi pemanfaatan hasil segmentasi wilayah pertanian sebagai dasar pengambilan keputusan dan perencanaan kebijakan pertanian berbasis data.

1.4 Hipotesis Penelitian

Berdasarkan tujuan penelitian, hipotesis kerja yang digunakan adalah sebagai berikut:

1. Algoritma K-Means dapat mengelompokkan 32 kabupaten/kota di Provinsi Sumatera Utara ke dalam tiga cluster berdasarkan kemiripan nilai luas panen, produktivitas, dan produksi padi.

2. Setiap cluster yang terbentuk memiliki karakteristik centroid yang berbeda sehingga dapat diinterpretasikan sebagai kelompok wilayah dengan potensi pertanian relatif rendah, menengah, dan tinggi.
3. Hasil segmentasi dapat digunakan sebagai informasi awal dalam mendukung penyusunan prioritas pengembangan pertanian berbasis data.

Hipotesis tersebut bersifat deskriptif dan tidak digunakan untuk menyimpulkan hubungan sebab-akibat antara hasil cluster dan faktor geografis, iklim, irigasi, atau kesuburan tanah karena variabel tersebut tidak dianalisis dalam penelitian ini.

1.5. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini sangat luas, baik dari segi kebijakan, praktis, maupun akademis:

1. **Manfaat Akademik**: Penelitian ini memperluas literatur penerapan machine learning di bidang pertanian serta menjadi referensi bagi penelitian berbasis unsupervised learning di Indonesia. Penelitian ini juga berkontribusi terhadap pengembangan kurikulum data science pertanian di perguruan tinggi.
2. **Manfaat Praktis**: Memberikan panduan analitik bagi pemerintah daerah dan lembaga pertanian dalam perumusan kebijakan berbasis segmentasi wilayah. Hasil penelitian diharapkan membantu optimalisasi distribusi sumber daya pertanian dan pemerataan bantuan pemerintah.
3. **Manfaat Sosial**: Mendukung pemerataan hasil pertanian dan peningkatan kesejahteraan petani melalui kebijakan berbasis bukti dan efisiensi distribusi

sumber daya. Penelitian ini juga berkontribusi pada peningkatan ketahanan pangan lokal dan pemberdayaan petani kecil di Sumatera Utara.