

## **BAB 1**

### **PENDAHULUAN**

#### **1.1. Latar Belakang**

Kakao merupakan salah satu komoditas perkebunan strategis di Indonesia yang memberikan kontribusi besar terhadap perekonomian nasional dan daerah. Kabupaten Aceh Tenggara memiliki potensi besar dalam produksi kakao dengan luas lahan yang signifikan dan banyaknya petani yang bergantung pada komoditas ini. Namun, produksi kakao di wilayah ini menunjukkan ketimpangan yang cukup besar antar daerah. Beberapa kecamatan memiliki produktivitas tinggi, sementara yang lain mengalami hasil panen rendah akibat faktor seperti teknik budidaya, kualitas tanah, akses infrastruktur, dan fluktuasi harga jual kakao.

Ketimpangan ini diperparah dengan belum adanya pemetaan produksi kakao yang jelas, sehingga kebijakan pemerintah sering kali tidak tepat sasaran. Bantuan pertanian, subsidi, dan pelatihan petani sulit dialokasikan secara efektif tanpa pengelompokan yang akurat mengenai daerah dengan produktivitas tinggi, sedang, dan rendah. Hal ini menyebabkan ketidakseimbangan dalam pengembangan sektor kakao serta menurunkan kesejahteraan petani, terutama di wilayah dengan hasil panen rendah.

Dinas Pertanian dipilih sebagai lokasi penelitian karena perannya dalam pengelolaan dan pengembangan sektor kakao di Kabupaten Aceh Tenggara. Lembaga ini memiliki data lengkap terkait produksi, luas lahan, jumlah petani, serta faktor lain yang mempengaruhi produktivitas kakao. Selain menyediakan data, Dinas Pertanian juga berperan dalam kebijakan peningkatan produksi, seperti penyediaan bibit unggul, distribusi pupuk, dan pelatihan petani. Hasil *clustering* dengan *Gaussian Mixture Model (GMM)* dapat membantu dalam perencanaan strategi yang lebih tepat sasaran, termasuk pengalokasian bantuan dan pengembangan infrastruktur pertanian. Dengan demikian, penelitian di Dinas Pertanian tidak hanya didukung oleh ketersediaan data, tetapi juga berkontribusi langsung dalam penyusunan kebijakan berbasis data untuk mendukung pengembangan kakao di Kabupaten Aceh Tenggara.

Penelitian ini bertujuan untuk mengelompokkan daerah produksi kakao di Kabupaten Aceh Tenggara menggunakan metode *Gaussian Mixture Model* (GMM), sebuah algoritma *clustering* yang lebih fleksibel dalam menangkap pola distribusi data dibandingkan metode lain seperti *K-Means*. Dengan menggunakan variabel produksi kakao, luas lahan, jumlah petani, tingkat produksi, dan produktivitas, penelitian ini akan mengidentifikasi pola produksi di setiap daerah.

Hasil penelitian ini diharapkan dapat membantu pemerintah dalam menyusun kebijakan yang lebih tepat, seperti pengalokasian bantuan teknis bagi daerah dengan produktivitas rendah, perencanaan infrastruktur pertanian yang lebih efisien, serta pelatihan petani berbasis data. Selain itu, penelitian ini juga berkontribusi pada pengembangan ilmu data mining dan penerapan metode *machine learning* dalam sektor pertanian, sehingga pengelolaan produksi kakao menjadi lebih efektif, berkelanjutan, dan berbasis data.

Data mining adalah proses iteratif dan interaktif untuk mengidentifikasi pola atau model baru yang signifikan, berguna, dan mudah dipahami dalam *database* berukuran besar (*massive database*). Proses ini melibatkan pencarian tren atau pola tertentu dalam kumpulan data besar untuk mendukung pengambilan keputusan di masa depan. Pola-pola tersebut dikenali melalui alat khusus yang mampu memberikan analisis data yang informatif dan bernilai wawasan, yang kemudian dapat dikaji lebih mendalam, sering kali dengan memanfaatkan alat pendukung keputusan lainnya (Sikumbang, 2022). Dalam penelitian ini, data minin g digunakan untuk mengelompokkan daerah berdasarkan tingkat hasil produksi tertinggi hingga terendah menggunakan salah satu metode algoritma yang dipilih untuk proses ini adalah algoritma *Gaussian Mixture Model*.

Algoritma *Gaussian Mixture Model* adalah metode *clustering* non-hierarkis yang bekerja dengan memodelkan data dalam distribusi. *Gaussian* menggunakan parameter *mean* dan *varians* tertentu. Dalam proses *clustering*, algoritma ini termasuk dalam kategori *soft clustering*, di mana satu data point dapat menjadi anggota dari dua atau lebih *cluster* dengan derajat keanggotaan tertentu. Sebagai model statistik yang populer, *Gaussian Mixture Model* sering digunakan karena efisiensinya dalam mengolah data dalam jumlah besar. Berdasarkan hasil

penelitian dan alasan bahwa data point tidak selalu terkласifikasi secara eksklusif ke dalam satu *cluster*, melainkan dapat berada dalam beberapa *cluster* dengan probabilitas tertentu.

Berdasarkan penelitian sebelumnya oleh (Prabiantissa & Yuliastuti, 2021) tentang Prediksi Pergerakan Ikan Di Pesisir Pulau Madura Menggunakan Metode *Gaussian Mixture Model* Dan *K-Means Clustering* menunjukkan hasil prediksi menggunakan algoritma *Gaussian Mixture Model* lebih akurat karena memiliki kemiripan antara hasil estimasi dan data hasil ikan. Penelitian selanjutnya dilakukan oleh (Rizal et al., 2024) tentang Pengelompokan Data Penjualan Produk Monitor Di Amazon Menggunakan Algoritma *Gaussian Mixture Model*. Penelitian ini berhasil melakukan *clustering* data penjualan monitor di Amazon menggunakan metode *Gaussian Mixture Model* (GMM). Hasil analisis menunjukkan bahwa GMM efektif dalam mengidentifikasi pola penjualan dan segmen pasar yang berbeda. Dengan menggunakan GMM, penelitian ini memberikan wawasan mendalam tentang karakteristik unik dari setiap segmen pasar monitor di Amazon.

Dari hasil penelitian sebelumnya yang menunjukkan bahwa algoritma *Gaussian Mixture Model* efektif pada berbagai studi kasus. Peneliti tertarik untuk meneliti tentang “*Clustering Daerah Penghasil Kakao Di Aceh Tenggara Menggunakan Algoritma Gaussian Mixture Model (GMM)*” dengan adanya penelitian ini diharapkan dapat menambah pengetahuan masyarakat mengenai faktor-faktor apa saja yang menyebabkan daerah tersebut termasuk daerah penghasil kakao tertinggi di Aceh Tenggara.

## 1.2. Rumusan Masalah

Rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan algoritma *Gaussian Mixture Model* (GMM) dalam clustering daerah produksi kakao di Kabupaten Aceh Tenggara?
2. Bagaimana menentukan jumlah cluster yang optimal dalam pengelompokan daerah produksi kakao menggunakan GMM?

### **1.3 Tujuan Penelitian**

Berdasarkan latar belakang dan rumusan masalah yang telah disusun diatas, maka tujuan penelitian ini adalah :

1. Penerapan algoritma *Gaussian Mixture Model* (GMM) untuk clustering daerah produksi kakao di Kabupaten Aceh Tenggara.
2. Menentukan jumlah cluster yang optimal berdasarkan pola produksi kakao menggunakan GMM.

### **1.4 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat baik dari sisi akademis, praktis, maupun pengembangan ilmu pengetahuan, sebagai berikut :

#### **1. Manfaat Akademis**

Penelitian ini menambah referensi dalam bidang data mining terkait penerapan algoritma Gaussian Mixture Model (GMM) untuk clustering, serta memberikan wawasan yang dapat menjadi acuan bagi penelitian lanjutan di sektor pertanian dan perkebunan.

#### **2. Manfaat Praktis**

Hasil penelitian ini dapat membantu Dinas Pertanian dan Perkebunan dalam memahami pola distribusi produksi kakao serta mendukung pengambilan keputusan terkait alokasi bantuan, pelatihan petani, dan pengembangan infrastruktur pertanian di Kabupaten Aceh Tenggara.

#### **3. Manfaat bagi Pengembangan Ilmu Pengetahuan**

Penelitian ini memperluas pemahaman tentang penerapan metode clustering dalam sektor pertanian dan mendorong pengambilan keputusan berbasis data untuk pengelolaan sumber daya secara lebih efektif dan efisien.

### **1.5 Ruang Lingkup dan Batasan Penelitian**

1. Penelitian ini dilakukan di Kabupaten Aceh Tenggara dengan fokus pada pengelompokan daerah penghasil kakao.
2. Menggunakan algoritma Gaussian Mixture Model (GMM) untuk proses clustering.

3. Data yang dianalisis berasal dari instansi resmi dan sumber sekunder yang relevan.
4. Tidak dilakukan perbandingan metode dengan algoritma clustering lain seperti K-Means atau DBSCAN.
5. Hasil penelitian hanya difokuskan pada pola pengelompokan, tanpa analisis terhadap kebijakan atau strategi pengembangan.