

BAB I

PENDAHULUAN

Latar Belakang

Pengelompokkan desa menjadi sebuah kebutuhan yang penting dalam rangka meningkatkan pengelolaan dan pengembangan sektor pertanian di Kabupaten Sumenep. Kabupaten Sumenep merupakan salah satu kabupaten di Jawa Timur yang memiliki potensi pertanian yang signifikan. Sebagai daerah agraris, Kabupaten ini memiliki banyak desa yang memiliki peran penting dalam sektor pertanian dan produksi hasil pertanian. Dalam melakukan pengelompokkan desa berdasarkan produktivitas hasil pertanian dibutuhkan metode yang efektif dan efisien. Dengan adanya pengelompokan desa berdasarkan produktivitasnya, Dinas Pertanian memiliki kesempatan untuk memberikan bantuan dan arahan yang lebih tepat kepada desa-desa tersebut. Melalui pemahaman yang lebih baik tentang tingkat produktivitas setiap desa, dinas pertanian dapat merancang program bantuan yang lebih efektif dan strategis.

Salah satu metode yang sering digunakan dalam pengelompokkan adalah algoritma k-means dan k-medoids. Algoritma-algoritma tersebut telah terbukti efektif dalam mengelompokkan data berdasarkan karakteristik tertentu.

Metode k-means dan k-medoids memiliki tujuan untuk melakukan pengelompokkan desa berdasarkan produktivitas hasil pertanian yang mereka hasilkan. Pada penelitian ini, terdapat sejumlah indikator yang digunakan dalam perhitungan pengelompokkan desa tersebut. Indikator tersebut meliputi luas tanaman, varietas/jenis tanaman, stadia/umur tanaman, jenis organisme pengganggu tanaman (OPT), pupuk bersubsidi organik dan anorganik yang digunakan, luas tanaman yang terserang, intensitas serangan, luas tanaman yang dalam kondisi waspada, serta hasil panen yang diperoleh.

K-Means adalah metode pengelompokan data non-hierarki yang bertujuan membagi data ke dalam satu atau lebih kelompok atau *cluster*, di mana data dengan karakteristik serupa dikelompokkan ke dalam satu *cluster*, sementara data dengan karakteristik yang berbeda dikelompokkan ke dalam kelompok yang berbeda pula[1]. Metode k-medoids adalah salah satu teknik partisi dalam analisis klaster, yang menggunakan objek medoids sebagai pusat klaster[2]. Objek medoids dipilih

sebagai titik tengah dari kelompok data dalam klaster[2]. Data yang telah di dapat akan dilakukan analisis menggunakan algoritma k-medoids *clustering* guna melakukan pengelompokkan desa berdasarkan kebutuhan arahan yang diperlukan. Selanjutnya, hasil pengelompokkan tersebut akan dievaluasi untuk menentukan apakah desa-desa tersebut layak untuk mendapatkan arahan dari pihak pertanian, dengan tujuan untuk meningkatkan hasil panen yang optimal. Penelitian ini mengadopsi pendekatan pengukuran dua metode, yakni metode k-means dan k-medoids, dikarenakan kelebihan yang dimiliki oleh metode k-medoids. Metode k-medoids menggunakan objek aktual dari kumpulan data sebagai pusat *cluster*, sementara metode k-means menggunakan rata-rata. Hal ini memberikan keuntungan dalam menjaga keakuratan representasi *cluster* yang dihasilkan. Di samping itu, metode k-means sering menghadapi kendala konvergensi ketika jumlah data sangat besar.

Untuk menentukan metode yang paling akurat akan di gunakan yaitu dengan menggunakan metode Sum of Squared Errors sebagai penentu metode paling akurat. *Sum of Squared Errors (SSE)* adalah salah satu metode statistik yang digunakan untuk mengukur selisih total antara nilai sebenarnya dan nilai yang diperoleh. Metode ini berguna dalam menganalisis keakuratan atau kesalahan prediksi dalam suatu model atau algoritma. Dengan mengkuadratkan selisih antara nilai sebenarnya dan nilai yang diprediksi, *SSE* memberikan gambaran tentang sejauh mana model atau algoritma tersebut berhasil menggambarkan atau memprediksi data dengan akurat. Semakin kecil nilai *SSE*, semakin baik model atau algoritma dalam mencocokkan data dengan nilai yang sebenarnya[3].

Penelitian terhadap proses *clustering* telah banyak dilakukan seperti penelitian yang dilakukan oleh Riva Arsyad Farissa, Rini Mayasari, Yuyun Umaidah dalam melakukan perbandingan algoritma k-means dan k-medoids untuk pengelompokkan data obat dengan silhouette coefficient. Berdasarkan hasil eksperimen, ditemukan bahwa kedua algoritma tersebut menunjukkan perbedaan dalam Silhouette Coefficient yang diperoleh. Algoritma K-Means menghasilkan nilai yang lebih tinggi, yakni sebesar 0,627, sedangkan K-Medoids menghasilkan nilai 0,536. Hal ini menunjukkan bahwa hasil pengelompokkan yang diperoleh dari algoritma K-Means memiliki kualitas yang lebih baik dibandingkan dengan K-

Medoids[4]. Penelitian lain juga dilakukan oleh Yunita Hilda Susanti, Edy Widodo dalam penelitian Perbandingan K-Means dan K-Medoids *Clustering* terhadap Kelayakan Puskesmas di DIY Tahun 2015. Hasil penelitian ini, penggunaan metode K-Means *Clustering* menghasilkan performa yang lebih baik dibandingkan dengan metode K-Medoids *Clustering*. Hal ini dapat dilihat dari nilai rasio simpangan baku yang lebih kecil yang diperoleh menggunakan metode K-Means *Clustering*[5]. Penelitian lain juga dilakukan oleh Novianti Puspitasari, Gidion Lempas, Hamdani, Haviluddin, dan Anindita Septiarini dalam penelitian Perbandingan Algoritma K-Means dan Algoritma K-Medoids Pada Kasus Covid-19 di Indonesia periode 2023 K-Medoids memiliki keunggulan dalam menghasilkan *cluster* dengan nilai kesalahan acak terkecil berdasarkan pengujian metode Sum Squared Error (SSE)[6].

Berdasarkan penelitian yang telah dilakukan sebelumnya dapat di simpulkan bahwa proses *clustering* dengan menggunakan metode k-means dan k-medoids sebagai perbandingan.

Rumusan Masalah

Permasalahan

Berdasarkan penjelasan latar belakang di atas, banyaknya desa dan banyaknya indikator pengelompokan sehingga dinas pertanian kesulitan dalam melakukan *clustering* atau pengelompokan.

Metode Usulan

Berdasarkan dengan permasalahan di atas maka di Kabupaten Sumenep membutuhkan sebuah sistem yang dapat mengelompokkan menggunakan metode Algoritma K-Means dan Algoritma K-Medoids dengan optimasi Sum of Squared Errors untuk mencari solusi metode paling optimal dengan melakukan perhitungan dengan tiga *cluster*.

Pertanyaan Penelitian

Berdasarkan latar belakang di atas, sehingga rumusan masalah yang akan di kaji yakni:

1. Bagaimana mengimplementasikan metode K-Means dan K-Medoids untuk pengelompokan desa berdasarkan hasil pertanian?
2. Bagaimana menentukan *cluster* yang paling optimal menggunakan metode *SSE* dan permintaan 3 *cluster* yang di inginkan oleh dinas pertanian Kabupaten Sumenep?

Tujuan dan Manfaat

Tujuan

1. Untuk mengimplementasikan metode K-Means dan K-Medoids untuk pengelompokan desa berdasarkan hasil pertanian.
2. Untuk menentukan *cluster* yang paling optimal menggunakan metode *SSE* dan permintaan 3 *cluster* yang di inginkan oleh dinas pertanian Kabupaten Sumenep

Manfaat

Manfaat dari penelitian bagi penulis yaitu menambah wawasan tentang pengelompokan menggunakan metode K-Means dan K-Medoids untuk membangun sebuah sistem untuk mengelompokkan daftar desa yang membutuhkan bimbingan dari pihak pertanian dalam bercocok tanam yang efektif di Kabupaten Sumenep.

Batasan Masalah

Dalam batasan masalah ini, penelitian ini di batasi permasalahan yang perlu:

1. Data yang digunakan adalah data produktivitas hasil panen Kabupaten Sumenep.
2. Data yang digunakan memiliki 9 indikator yaitu luas tanaman, varietas, umur tanaman, jenis OPT, pupuk, luas terserang, intensitas, luas waspada, dan hasil panen.
3. Dataset berjumlah 690 pada tahun 2022.
4. Jumlah cluster 3 yang di inginkan dari pihak dinas pertanian Kabupaten Sumenep.

Sistematika Penelitian

Dalam pembuatan sistematika penulisan skripsi di buat agar memudahkan penulis supaya menulis dengan runtut serta penulis fokus menganalisis dan terarah, berikut sistematika penulisan laporan skripsi sebagai berikut:

BAB I : PENDAHULUAN

Dalam bab ini berisi tentang penjelasan ringkas latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, dan sistematika penelitian.

BAB II : TINJAUAN PUSTAKA

Dalam bab ini berisi tentang penjelasan teori yang dapat menunjang penelitian, yaitu berupa Ketahanan Pangan, Data Mining, *Clustering*, K-Means, K-Medoids, dan *Sum of Squared Errors (SSE)*

BAB III : METODOLOGI PENELITIAN

Dalam bab ini berisi tentang cara penerapan konsep dasar yang telah di uraikan pada Bab II untuk menganalisis dan merancang metode penelitian, dan dataset yang di gunakan.

BAB IV : HASIL DAN PEMBAHASAN

Dalam bab ini berisis tentang implementasi program computer berdasarkan hasil perancangan yang dibuat menggunakan algoritma K-Means dan K-Medoids *Clustering*.

BAB V : PENUTUP

Dalam bab ini berisi tentang kesimpulan dan saran dari deluruh penulisan tugas akhir.