

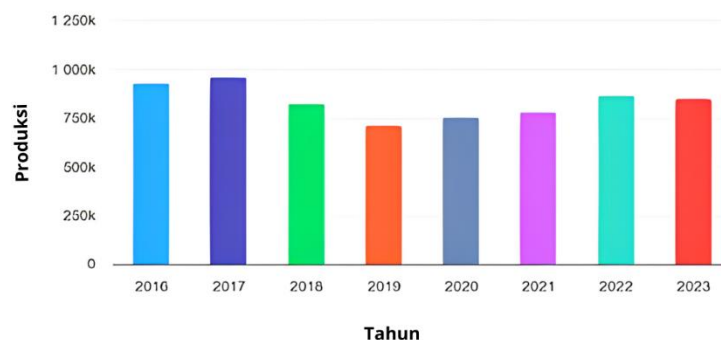
BAB I

PENDAHULUAN

1.1 Latar Belakang

Padi merupakan jenis tanaman yang berperan penting dalam memenuhi kebutuhan makanan bagi sebagian besar populasi dunia, terutama di Bojonegoro. Kabupaten Bojonegoro dijuluki sebagai lumbung pangan andalan Jawa Timur[1]. Padi (*Oryza sativa*) adalah tanaman pangan utama dengan sejarah panjang dalam budidaya, memiliki peran penting dalam aspek sosial dan ekonomi, serta berkontribusi signifikan terhadap ketahanan pangan global. Kabupaten Bojonegoro memiliki lahan yang luas dan memproduksi berbagai jenis tanaman, khususnya pada tanaman padi. Selain itu, sebagian besar masyarakat di Kabupaten Bojonegoro berprofesi sebagai petani, terutama petani padi sawah.

Berdasarkan data Badan Pusat Statistik (BPS) Kabupaten Bojonegoro, produktivitas padi mengalami perubahan sebagai berikut meningkat dari tahun 2016 ke 2017, menurun dari tahun 2017 ke 2018, menurun lagi dari tahun 2018 ke 2019, kemudian meningkat dari tahun 2019 ke 2020, naik lagi dari tahun 2020 ke 2021, terus naik dari tahun 2021 ke 2022, dan akhirnya menurun dari tahun 2022 ke 2023. Dari penjelasan tersebut, data produksi padi dapat dilihat pada Gambar 1.1 sebagai berikut:



Gambar 1. 1 Data Produksi Tanaman Padi Kab. Bojonegoro Tahun 2016-2023

Berdasarkan Gambar 1.1 hasil produksi padi tahun ke tahun menunjukkan

ketidakstabilan. Produksi yang tidak konsisten ini mempengaruhi masyarakat karena padi merupakan komoditas utama dan sumber pangan pokok bagi masyarakat. Ada beberapa faktor yang menyebabkan penurunan pada produktivitas padi yaitu seperti kondisi cuaca ekstrem, serangan penyakit dan hama, kesehatan tanah yang buruk, dan strategi pertanian yang tidak efektif. Kekeringan atau banjir yang berkepanjangan dapat menghambat pertumbuhan tanaman, sedangkan serangan penyakit seperti blas atau hawar daun dan ditambah beberapa tahun ini serangan hama tikus dapat menyebabkan hasil panen menurun. Kualitas tanah yang tidak subur atau kurangnya nutrisi, bersama dengan pemilihan varietas padi yang tidak sesuai, juga dapat menyebabkan penurunan produktivitas. Selain itu, penggunaan pupuk yang tidak tepat, praktek irigasi yang tidak efisien, serta kurangnya strategi yang digunakan oleh petani, dan kurangnya strategi yang dimiliki pemerintah untuk tetap mempertahankan produksi dari tahun ke tahun[2]. Dengan merencanakan pengoptimalan pemanfaatan lahan di wilayah yang masih belum dimanfaatkan untuk produksi komoditas, pemerintah bertanggung jawab untuk meningkatkan produktivitas pertanian padi secara berkelanjutan setiap tahunnya.

Dalam data pertanian yang disediakan oleh Dinas Ketahanan Pangan dan Pertanian Kabupaten Bojonegoro, terdapat variasi signifikan dalam hasil produktivitas di berbagai kecamatan. Untuk memahami pola dan perbedaan dalam produktivitas, penting dilakukan pengelompokan data[3]. Pengelompokan ini akan didasarkan pada beberapa parameter, yaitu luas tanam, luas panen, produktivitas, dan total produksi per kecamatan. Melalui pendekatan ini, akan diperoleh gambaran yang lebih jelas mengenai distribusi produktivitas di tiap kecamatan, sehingga dapat diidentifikasi daerah dengan tingkat penyebaran dan pengelompokan produktivitas sesuai kemiripan antar wilayah.

Pengelompokan produktivitas padi diimplementasikan melalui teknik *clustering*, yang merupakan metode untuk mengelompokkan data berdasarkan karakteristik atau pola tertentu. Dengan menerapkan teknik ini, dapat lebih efektif menganalisis dan memahami pola produktivitas padi di berbagai wilayah, memungkinkan pengambilan keputusan yang lebih terarah dalam upaya meningkatkan hasil pertanian. Saat ini terdapat metode yang akan digunakan untuk

clustering seperti *K-Means Clustering* dan *Agglomerative Hierarchical Clustering*. *K-Means Clustering* memiliki beberapa keunggulan, antara lain mudah diterapkan dan dijalankan, kecepatan proses, serta fleksibilitas yang baik. Metode ini juga sering dipakai dalam praktik data mining[4]. Sedangkan keunggulan dari metode *Agglomerative Hierarchical Clustering* (AHC) adalah dapat memilih fungsi metrik atau fungsi jarak, sehingga memberikan keleluasaan untuk mengukur kemiripan antar data. Jenis metrik yang dipilih bisa berbeda-beda, seperti Euclidean, Manhattan, atau Mahalanobis, tergantung pada sifat data dan tujuan analisis. Kesesuaian pilihan metrik ini memungkinkan penyesuaian yang optimal dengan karakteristik data dan tujuan analisis yang diinginkan[5]. Dengan menggunakan metode dua tersebut diperlukannya evaluasi, penelitian ini menggunakan evaluasi *Silhouette Coefficient* untuk menentukan nilai cluster optimal dari perbandingan dua metode.

Berdasarkan penelitian terkait yaitu pengelompokan produktivitas tanaman padi di Jawa Tengah menggunakan metode *Clustering K-Means*. Penelitian ini bertujuan untuk melihat dan mengetahui kota atau kabupaten yang memiliki potensi dalam memproduksi padi sekaligus daerah yang memiliki produksi padi kurang maksimal. Setelah proses perhitungan sampai iterasi ke 5 mendapat hasil cluster C0 dengan produktivitas sedang memiliki 12 daerah, lalu C1 dengan produktivitas rendah memiliki 18 daerah, dan C2 dengan produktivitas tinggi memiliki 5 daerah[6].

Tinjauan penelitian terkait berikutnya yaitu Analisis Cluster Penyakit Malaria Provinsi Papua Menggunakan *Metode Single Linkage* Dan *K-Means* dengan data penyakit malaria berdasarkan rentang usia dan bulan terjadinya. Dimana cluster satu menunjukkan 50 pasien terdapat 47 pasien(tertinggi). Dan cluster kedua menunjukkan bahwa 50 pasien terhadap 11 pasien. Berdasarkan hasil evaluasi menggunakan *Silhouette* terdapat kebaikan maka $S(\text{metode single linkage}) = 76.2\%$ dan $S(\text{metode K-Means}) = 85.1\%$, yang artinya lebih akurat *single linkage* dibandingkan *K-Means*[7].

Penelitian selanjutnya perbandingan analisis cluster algoritma *K- Means* dan AHC dalam perancangan persediaan barang pada perusahaan manufaktur.

Penelitian ini membandingkan algoritma AHC dan *K-Means* dalam pengelompokan perencanaan persediaan barang pada perusahaan manufaktur. Menggunakan dataset penjualan Juni 2021. Pengelompokan penjualan barang yang dihasilkan berdasarkan kesamaan rata-rata jumlah penjualan bulanan serta rekomendasi persediaan barang untuk proses penjualan di masa mendatang, berdasarkan rata-rata nilai penjualan setiap pusat klaster yang dibentuk[8].

Berdasarkan penelitian sebelumnya yang telah dilakukan dengan teknik *Clustering*, didapatkan kekurangan dalam mendapatkan hasil klaster awal yang unik dan belum adanya pengkajian atau validasi lebih lanjut, maka dari itu peneliti mengusulkan pendekatan analisa data mining dengan menggunakan metode *K-Means Clustering* dan *Agglomerative Hierarchical Clustering* (AHC).

Penelitian ini menganalisis 4 (empat) parameter tanaman padi yaitu luas panen, luas tanam, produktivitas, dan produksi. Mengetahui pengelompokan produksi padi pada kabupaten Bojonegoro dengan membandingkan dua metode tersebut.

Oleh karena itu, penelitian ini mengusulkan pengelompokan data tanaman padi dengan menggunakan metode *K-Means* dan *Agglomerative Hierarchical Clustering* (AHC) serta uji evaluasi menggunakan metode *Silhouette Coefficient*. Dengan adanya pengelompokan ini diharapkan dapat membantu dan mempermudah dalam melakukan pemantauan, penyebaran pupuk, serta perkembangan produksi padi pada Dinas Ketahanan Pangan dan Pertanian Kabupaten Bojonegoro.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

1.2.1 Permasalahan

Pengelompokan produksi padi di kabupaten Bojonegoro saat ini belum stabil oleh karena itu pemerintah kabupaten Bojonegoro perlu memperhatikan penyebaran dan pengelompokan wilayah dalam suatu kelompok ke kelompok lainnya yang memiliki kemiripan. Dengan pemantauan wilayah yang kurang pemerintah kabupaten Bojonegoro memerlukan pemetaan produktivitas padi untuk

pengambilan keputusan. Dengan ini pemerintah dapat mengoptimalkan dalam memberikan bantuan benih, pupuk, obat pembasmi hama serta pembangunan jaringan irigasi tingkat usaha tani serta upaya memaksimalkan potensi produksi padi di setiap wilayah dan juga mendorong laju ekonomi di sektor pertanian di kabupaten Bojonegoro. Dengan data yang digunakan terdiri dari 4 parameter yaitu luas panen, luas tanam, produktivitas, dan produksi.

1.2.2 Metode Usulan

Penelitian ini mengangkat topik data mining pada produktivitas padi di kabupaten Bojonegoro. Dalam memetakan produktivitas padi pada kabupaten Bojonegoro diperlukannya clustering menggunakan *K-Means* dan *Agglomerative Hierarchical Clustering* (AHC).

1.2.3 Pertanyaan Penelitian

Berdasarkan hasil uraian permasalahan, maka pertanyaan yang menjadi fokus penelitian ini yaitu

1. Berapa nilai hasil evaluasi menggunakan *Silhouette Coefficient* dari perbandingan metode *K-Means* dan *Agglomerative Hierarchical Clustering* (AHC) dalam pengelompokan produktivitas padi pada kabupaten Bojonegoro?
2. Bagaimana hasil penyebaran produktivitas padi di Kabupaten Bojonegoro berdasarkan *clustering K-Means* dan AHC dengan visualisasi dalam peta SIG?

1.3 Tujuan dan Manfaat

Tujuan dan manfaat dari penelitian ini sebagai berikut:

1.3.1 Tujuan Penelitian

1. Mengetahui nilai evaluasi *Silhouette Coefficient* dari perbandingan metode *K-Means* dan *Agglomerative Hierarchical Clustering* (AHC) dalam pengelompokan produktivitas padi di Kabupaten Bojonegoro.
2. Mengetahui pola penyebaran produktivitas padi berdasarkan hasil *clustering* dari kedua metode tersebut, serta menampilkan hasilnya dalam peta SIG.

1.3.2 Manfaat Penelitian

Manfaat dari penelitian ini bagi pihak pemerintah dan pembaca sebagai berikut:

1. Manfaat Bagi Pihak Pemerintah

Manfaat penelitian bagi pihak pemerintah yaitu sebagai berikut:

- a. Sebagai dasar untuk mengetahui penyebaran dan pengelompokan daerah, dengan itu dapat digunakan sebagai salah satu acuan untuk distribusi pupuk, obat pembasmi hama dan pengambilan keputusan dalam memaksimalkan dan mengoptimalkan produksi tanaman padi di kabupaten Bojonegoro
- b. Sebagai informasi yang berguna untuk dinas ketahanan pangan dan pertanian Bojonegoro.

2. Manfaat Bagi Pembaca

Manfaat penelitian bagi pembaca yaitu sebagai berikut:

- a. Dapat meningkatkan pengetahuan dan wawasan pembaca mengenai metode K-Means dan *Agglomerative Hierarchical Clustering* (AHC).
- b. Sebagai referensi bagi pembaca untuk melakukan penelitian mengenai Clustering.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian ini dilakukan pada Dinas Pertanian Kabupaten Bojonegoro, dengan total 28 kecamatan yang akan digunakan untuk penelitian.
2. Dataset yang digunakan yaitu data produktivitas padi di kabupaten Bojonegoro dengan 4 parameter yaitu luas tanam (Ha), luas panen (Ha), produktivitas (Ku/Ha) dan produksi (Ton).
3. Dataset yang digunakan dari tahun 2016-2023, diambil dari website “Satu Data Bojonegoro” <https://data.bojonegorokab.go.id/>.
4. Data diolah menggunakan metode perbandingan algoritma *K-Means Clustering* dan *Agglomerative Hierarchical Clustering* (AHC) dengan *Single Linkage*.
5. Metode uji validasi menggunakan *Silhouette Coefficient*.
6. Sistem Informasi Geografis (SIG) berupa peta statis untuk pemetaan.

1.5 Sistematika Proposal

Sistematika penulisan proposal ini terdiri dari beberapa bab, yaitu:

BAB I PENDAHULUAN

Pada bab ini berisi tentang latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah dan sistematika penulisan

BAB II TINJAUAN PUSTAKA

Bab ini berisi mengenai pemaparan teori-teori penunjang yang berkaitan dengan padi, data mining, *Clustering*, *K-Means*, *Agglomerative Hierarchical Clustering* (AHC), *Elbow*, *Silhouette Coefficient* serta penelitian-penelitian terdahulu yang dijadikan peneliti sebagai acuan pada penelitian ini.

BAB III METODE USULAN

Bab ini berisi mengenai metodologi penelitian seperti alur penelitian, dataset, algoritma sistem pengelompokkan dari pengumpulan data, proses preprocessing data, proses kedua metode dan pengujian penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan mengenai hasil dan pembahasan dari proses penelitian yang telah dilakukan berupa source code, penjelasan source code, uji scenario, hasil analisa, dan impelentasi antar muka sistem.

BAB V PENUTUP

Bab ini merumuskan kesimpulan dari hasil penelitian yang dilakukan dengan memberi saran untuk saran penelitian selanjutnya.

DAFTAR PUSTAKA

Bab berisi tentang keterangan mengenai sumber rujukan yang digunakan dalam penyusunan Tugas Akhir. Pada penelitian ini daftar pustaka disusun secara otomatis dengan menggunakan aplikasi *mendeley*.