

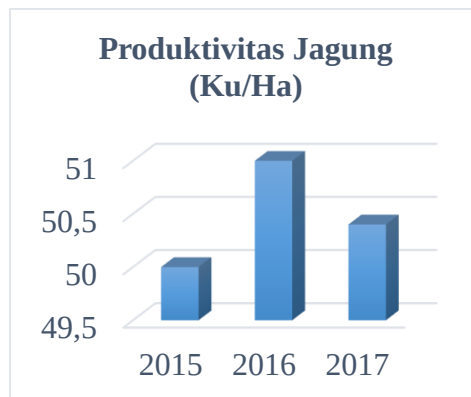
BAB I

PENDAHULUAN

1.1 Latar Belakang

Jagung merupakan tanaman jenis rerumputan tropis yang mampu beradaptasi terhadap perubahan iklim, dan memiliki umur 70-210 hari. Jagung dapat tumbuh setinggi 3 meter. Jagung yang bernama ilmiah *Zea mays*. ini berbeda dengan tanaman pangan lainnya karena satu satunya tanaman yang memiliki bunga jantan dan betina yang terpisah. Jagung (*Zea mays*) adalah salah satu sereal yang bernilai ekonomi dan strategis seluruh bagian dari tanaman jagung dapat diolah untuk segala macam kebutuhan. Jagung juga mempunyai peluang yang dapat dibesarkan sebab memiliki peran sebagai sumber utama protein dan karbohidrat setelah beras [1].

Berdasarkan Badan Pusat Statistika (BPS) provinsi Jawa Timur produktivitas jagung mengalami peningkatan dari tahun 2015 ke 2016 sebesar 1 (Ku/Ha) dan pada tahun 2016 menuju 2017 penurunan sebesar 0,6 (Ku/Ha). Dapat dilihat pada Gambar 1.1.



Gambar 1. 1 Data Produktivitas Jagung

Ada beberapa komoditas termasuk jagung, hasil produksinya tidak konsisten dari tahun ke tahun. Penurunan produktivitas jagung ini terjadi karena beberapa penyebab yaitu kurangnya pasokan air ke ladang, terjadinya bencana alam, terdapat serangan dari Organisme Pengganggu Tanaman (OPT), kurangnya kebutuhan pupuk, potensi hasil dari preferensi penggunaan varietas, serta cuaca yang tidak menentu sehingga menyebabkan petani jagung gagal panen. Selain itu kurangnya strategi yang dimiliki pemerintah untuk tetap mempertahankan produksi dari tahun ke tahun. Strategi yang dapat digunakan yaitu memaksimalkan daerah yang masih memiliki banyak lahan, namun belum menghasilkan komoditas tertentu. Hal ini menjadi tugas pemerintah agar terus meningkatkan produktivitas tanaman jagung.

Oleh sebab itu, salah satu solusi untuk permasalahan seperti pendekatan untuk analisis *data mining* dengan menggunakan teknik *clustering K-Means* dan *K-Medoids*. Metode *K-Means* ini merupakan teknik pengelompokan data kedalam beberapa kelompok. *K-Means* adalah salah satu cara pengelompokan data non-hirarki yang didasarkan penentuan jumlah *cluster*. Sedangkan *K-Medoids* merupakan algoritma yang masih bertautan dengan algoritma *K-Means* [2]. *K-Medoids* adalah algoritma *clustering* yang bertujuan untuk meminimalkan kriteria kesalahan absolut daripada *Sum of Squared Errors* (SSE) [3]. Pengelompokan metode *K-Medoids* menggunakan objek *representatif* yang disebut *medoids* sebagai titik pusatnya [4]. Pengelompokan didasari dengan kesamaan karakteristik di masing-masing wilayah berdasarkan jumlah *cluster* yang telah ditentukan sebelumnya. Metode *preprocessing* yang digunakan adalah *missing value* dan normalisasi data. Pada penanganan metode *missing values* menggunakan teknik imputasi. Metode imputasi yang digunakan adalah imputasi *mean* dan imputasi *median*. Hasil pengelompokan untuk mengevaluasi hasil cluster (mencari nilai K) terbaik menggunakan metode *Elbow*. Menganalisa kualitas cluster dari setiap *clusteringnya* serta memilih algoritma terbaik menggunakan *Davies-Bouldin Index* (DBI). Tujuan pengelompokan ini untuk melihat perbandingan kedua metode pengklateran dan mengetahui daerah mana yang memiliki hasil produksi jagung yang paling produktif pada setiap wilayah/kota di Jawa Timur. Dari hasil perbandingan tersebut memperoleh algoritma terbaik untuk pengclusteran tanaman jagung dan agar pemerintah memberi pupuk untuk wilayah-wilayah yang masih belum memaksimalkan hasil komoditasnya dengan rata.

Berdasarkan penelitian sebelumnya oleh Insanul Kamila, Ulya Khairunnisa, dan Mustakim (2019) terkait perbandingan metode *K-Means* dan *K-Medoids* untuk pengelompokan data transaksi bongkar muat di Provinsi Riau, dengan melakukan 10 pengujian. Pada algoritma *K-Means cluster* terbaik terletak pada C3 dengan nilai DBI = 0,112. Sedangkan pada *K-Medoids cluster* terbaik juga terletak pada C3 dengan nilai DBI = 0,119. Dari hasil nilai DBI terlihat bahwa algoritma *K-Means* lebih rendah dibandingkan algoritma *K-Medoids* [5].

Kemudian pada penelitian oleh Fitriana Harahap (2021) mengenai pengelompokan siswa tunagrahita dengan membandingkan algoritma *K-Means* dan *K-Medoids*. Kedua algoritma ini menghasilkan tiga *cluster* optimal. Algoritma *K-Means* terdapat delapan siswa tunagrahita ringan, 14 siswa tunagrahita sedang, dan 14 siswa tunagrahita berat, nilai DBI = 0,161. Sedangkan pada algoritma *K-Medoids* terdapat terdapat 7 siswa tunagrahita ringan, 19 siswa tunagrahita sedang,

dan 10 siswa tunagrahita berat, dengan nilai DBI = 0,281. Sehingga algoritma *K-Means* memiliki hasil yang lebih baik daripada algoritma *K-Medoids* [6].

Pada penelitian ini menganalisa 5 (lima) parameter tanaman jagung yaitu *luas panen, produktivitas, produksi, varietas, dan luas area tanam*. Mengetahui pengelompokan produksi jagung di provinsi Jawa Timur dengan membandingkan algoritma *K-Means* dan *K-Medoids*.

Oleh karena itu, penelitian dilakukan untuk *clustering* data pertanian tanaman jagung dengan menggunakan algoritma *K-Means* dan algoritma *K-Medoids* serta uji validasi metode *Davies-Bouldin Index*. Adanya *clustering* ini diharapkan dapat membantu dan mempermudah Dinas Pertanian Dan Ketahanan Pangan dalam melakukan pemantauan, perkembangan produksi jagung serta pengetahuan dalam mengambil kebijakan terkait penanganan terkait tanaman jagung yang ada di setiap kota/kabupaten Provinsi Jawa Timur. *Davies-Bouldin Index* berguna untuk memperoleh jumlah dan nilai *cluster* yang optimal serta memilih algoritma terbaik.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Indonesia memiliki banyak potensi sumber daya alam yang sangat besar untuk mencukupi kebutuhan pangan penduduknya. Salah satu provinsi yang berperan aktif dalam laju pertumbuhan ekonomi di sektor pertanian tanaman jagung adalah Jawa Timur. Di Jawa Timur, pemerintah masih sulit untuk memantau pertumbuhan jagung setiap wilayah sehingga produksi tanaman jagung di Jawa Timur kurang dapat dioptimalkan, baik dari segi kuantitas maupun kualitasnya. Pemantauan yang kurang optimal dapat berimplikasi dalam tindakan evaluasi atau penanganan yang buruk atas kendala yang terjadi dalam produksi atau jagung. Oleh karena itu, pemetaan luas dan produksi jagung di setiap wilayah Jawa Timur sangat diperlukan untuk menunjang pengambilan keputusan pemerintah dalam mengoptimalkan produksi jagung di setiap Kabupaten/Kota di Provinsi Jawa Timur. Contohnya dengan pemberian pupuk dan mengoptimalkan aliran air pada daerah yang produksi jagungnya kurang. Terlebih lagi, dengan adanya sistem pemantauan yang memadai, pemerintah lewat Dinas Pertanian Dan Ketahanan Pangan dapat lebih memaksimalkan potensi produksi jagung sehingga dapat semakin mendorong laju ekonomi di sektor pertanian di Provinsi Jawa Timur.

1.2.2 Metode Usulan

Dalam memetakan tanaman jagung wilayah Jawa Timur diperlukannya *clustering* menggunakan algoritma *K-Means* dan algoritma *K-Medoids*. Data yang digunakan terdiri dari 5 (lima) parameter yaitu *luas panen, produktivitas, produksi, varietas, dan luas area tanam*.

1.2.3 Pertanyaan Penelitian

Berdasarkan permasalahan dan metode yang telah dipaparkan, dirumuskan :

1. Bagaimana hasil pengelompokan berdasarkan Produksi Jagung di Provinsi Jawa Timur menggunakan data imputasi *mean*, imputasi *median*, dan normalisasi *min max* dengan metode *K-Means* dan *K-Medoids*?
2. Berapa nilai *Davies-Bouldin Index* pada *K-Means* dan *K-Medoids* untuk mengelompokkan tanaman jagung pada daerah Provinsi Jawa Timur?

1.3 Tujuan dan Manfaat

Berdasarkan rumusan yang telah disusun, maka tujuan yang dicapai dalam penelitian yaitu :

1. Untuk mengetahui hasil pengelompokan dan perbandingan Kabupaten/Kota berdasarkan produksi tanaman jagung di Provinsi Jawa Timur menggunakan data imputasi *mean* dan imputasi *median* dengan metode *K-Means* dan *K-Medoids*.
2. Untuk mengetahui nilai dari *Davies-Bouldin Index* pada *K-Means* dan *K-Medoids* yang optimal.

Manfaat yang diharapkan yaitu:

1. Membantu Dinas Pertanian untuk menjalankan program terkait isu-isu mengenai tanaman jagung wilayah Jawa Timur.
2. Sebagai informasi yang berguna untuk Dinas Pertanian dan Ketahanan Pangan dan masyarakat Jawa Timur.
3. Sebagai dasar pengambilan keputusan dalam mengoptimalkan produksi tanaman jagung di Provinsi Jawa Timur
4. Sebagai referensi untuk melaksanakan penelitian selanjutnya tanaman jagung di Provinsi Jawa Timur.

1.4 Batasan Masalah

Beberapa batasan masalah dalam penelitian ini yaitu :

1. Pada Provinsi Jawa Timur mempunyai 9 (sembilan) kota dan 29 (dua puluh sembilan) kabupaten.

2. *Dataset* yang digunakan adalah data potensi tanaman jagung di kabupaten/kota Provinsi Jawa Timur, 5 (lima) parameter yang terkait *luas panen (Ha)*, *produktivitas (Ku/Ha)*, *produksi (Ton)*, *varietas (Bisi-16, Bisi-2, P-21, Bisi-18, Pertiwi-3, Bisi-816, Nk-33, Bisi-222, Nk-22, dan Lokal)*, dan *luas area tanam (Ha)*.
3. Data potensi pertanian wilayah Provinsi Jawa Timur pada tahun 2017-2022 diambil dari Dinas Pertanian dan Ketahanan Pangan Jawa Timur.
4. Keseluruhan jumlah data pada 5 parameter yaitu 3230 data.
5. Jumlah *missing values* pada parameter *varietas* terdapat 959 data.
6. Jumlah *cluster K* ditentukan menggunakan metode *Elbow*
7. Nilai dari $C1 > 80.000$, $C2 \geq 50.000$, dan $C3 \leq 49.000$ untuk *parameter luas panen (Ha)*, *produksi (Ton)* dan *luas area tanam (Ha)* berdasarkan rata-rata data.
8. Nilai dari $C1 > 60$, $C2 \geq 35$, dan $C3 \leq 34$ untuk parameter *produktivitas (Ku/Ha)* berdasarkan rata-rata data.
9. Nilai dari $C1 > 600$, $C2 \geq 400$, dan $C3 \leq 399$ untuk parameter *varietas (Bisi-16, Bisi-2, P-21, Bisi-18, Pertiwi-3, Bisi-816, Nk-33, Bisi-222, Nk-22, dan Lokal)* berdasarkan rata-rata data.

1.5 Metodologi Penulisan Skripsi

Sistematika penulisan skripsi ini dibagi menjadi lima bab, yang masing-masing bab telah dirancang untuk tujuan tertentu, berikut penjelasan masing-masing bab :

BAB I : PENDAHULUAN

Pada bab ini yang membahas tentang latar belakang, perumusan masalah, tujuan dan manfaat serta batasan masalah.

BAB II : KAJIAN PUSTAKA

Bab ini memaparkan teori teori penunjang yang berkaitan dengan jagung, *data mining*, *Clustering K-Means*, *K-Medoids*, *DBI (Davies-Bouldin Index)*, *Elbow*, *Normalisasi Min Max*, *Missing values* serta penelitian-penelitian terdahulu yang berhubungan dengan permasalahan yang diteliti dan menjadi acuan konseptual.

BAB III : METODE PENELITIAN

Pada bab ini mendeskripsikan mengenai metodologi penelitian seperti *dataset* yang digunakan, serta alur kerja sistem.

BAB IV : HASIL DAN PEMBAHASAN

Pada bab ini menjelaskan mengenai hasil dan pembahasan dari proses penelitian yang dilakukan setelah mendapatkan data pendukung untuk memperoleh hasil penelitian.

BAB V : PENUTUP

Pada bab ini merumuskan kesimpulan dari hasil penelitian yang dilakukan serta memberi saran terdapat penelitian selanjutnya.

DAFTAR PUSTAKA

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