

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertanian merupakan bidang yang paling menjadi sorotan oleh pemerintah dikarenakan adanya peran penting dalam pertumbuhan ekonomi masyarakat. Di Indonesia sendiri tanaman pangan merupakan suatu kebutuhan yang mendasar bagi masyarakat. Tanaman pangan di Indonesia menjadi produk yang dirasakan paling banyak manfaatnya karena memenuhi kebutuhan pasar domestik. Tanaman padi merupakan jenis kelompok tanaman pangan yang banyak terdapat di Indonesia. Tanaman padi produksinya dari tahun ke tahun mengalami pasang surut.

Saat ini perkembangan Teknologi Informasi semakin pesat terutama pada pengolahan data, terdapat banyaknya data dalam suatu instansi atau lembaga pertanian, sehingga memberi kesulitan terhadap pengelompokan data. Namun semakin pesatnya Teknologi Informasi dapat diatasi permasalahan tersebut dengan menggunakan teknik *Data Mining*.

Berdasarkan data hasil pertanian di Badan Pusat Statistik, menampilkan wilayah dengan hasil pertanian yang bervariasi dari jumlah data luas panen padi dan hasil produksi padi. Pada penelitian ini dilakukan pengelompokan setiap provinsi di Indonesia berdasarkan data hasil potensi pertanian yang dimiliki oleh masing-masing provinsi. Pada pembagian kelompok wilayah dapat dilakukan berdasarkan luas panen padi dan hasil produksi padi pada tahun 2018. Setiap provinsi akan dibagi/diklusterkan ke dalam 3 kluster yang mengindikasikan

bahwa kluster 1 merupakan kelompok provinsi dengan hasil potensi yang baik, kluster 2 merupakan kelompok provinsi dengan hasil potensi yang sedang, dan sedangkan kluster 3 merupakan kelompok provinsi dengan hasil potensi yang kurang baik dengan melihat centroid awal yang sudah ditentukan.

Beberapa algoritma yang digunakan dalam teknik *data mining* dengan metode *Algoritma K-Means* dan *Algoritma Silhouette Coefficient*. Diharapkan dalam proses *clustering* yang diterapkan menggunakan *Algoritma K-Means* dapat mengelompokkan dan menentukan *cluster* yang akurat. Dan juga dengan adanya *Algoritma Silhouette Coefficient* diharapkan bisa melihat kualitas dan kekuatan klaster, seberapa baik suatu objek ditempatkan dalam suatu klaster.

Menyikapi permasalahan tersebut, pada penelitian ini penulis mencoba untuk menerapkan teknik *Data Mining* dengan menggunakan *Algoritma K-Means* dan *Silhouette Coefficient* sebagai metode penyelesaian masalah tersebut, sehingga penulis menjadikannya sebuah judul "**Analisis Potensi Produksi Padi di Indonesia Menggunakan Algoritma K-Means dan Algoritma Silhouette Coefficient**".

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka perlu dirumuskan adalah

1. Bagaimana menerapkan *Algoritma K-Means* untuk mengelompokkan potensi produksi padi di Indonesia?
2. Bagaimana menguji kualitas dari cluster yang dihasilkan menggunakan *Algoritma Silhouette Coefficient*?

1.3 Batasan Masalah

Batasan masalah penelitian meliputi :

1. Data yang digunakan adalah data tanaman padi.
2. Data tersebut meliputi data luas panen padi dan data produksi padi.
3. Sistem yang dirancang menggunakan Bahasa Pemrograman *PHP* dan menggunakan *database PHP MySQL*.

1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan pada penelitian ini adalah untuk mengelompokkan produktivitas provinsi-provinsi di Indonesia dalam memproduksi tanaman padi dengan menerapkan *Algoritma K-Means*.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah :

1. Memberikan gambaran langkah-langkah penerapan *Algoritma K-Means* dan *Silhouette Coefficient* pada data penelitian ini.
2. Memberikan referensi, rekomendasi maupun contoh bagi penelitian yang berkaitan dengan *clustering* dalam bidang pertanian.
3. Dapat memberikan informasi atau masukan bagi dinas pertanian dalam upaya peningkatan produktivitas tanaman padi.

1.6 Metodologi Penelitian

Metodologi penelitian merupakan cara utama yang digunakan peneliti untuk mencapai tujuan dan menentukan jawaban atas masalah yang diajukan. Beberapa metode-metode penelitian yang akan digunakan antara lain:

1.6.1 Metode Pengumpulan Data

Metode pengumpulan data diperlukan untuk mendapatkan berbagai informasi yang dibutuhkan dalam mencapai tujuan penelitian. Dalam melakukan penelitian ini penulis menerapkan beberapa metode antara lain:

a. Studi Pustaka

Penulis mempelajari dan memperoleh data dari jurnal, artikel, dan buku-buku serta mencari informasi yang diperlukan terkait penyusunan laporan.

b. Sumber Data

Data yang digunakan dalam penelitian ini diperoleh secara online melalui <https://www.bps.go.id/> sebagai sumber data yang untuk penelitian.

1.6.2 Metode Analisis

Dalam tahap analisis yang dilakukan penulis yaitu :

- a. Menganalisis tahap demi tahap dari proses clustering sehingga mencapai hasil akhir.
- b. Menganalisis mengenai cara kerja dari *Algoritma K-Means* dalam mengclustering data tanaman padi.
- c. Menganalisis mengenai cara kerja dari *Algoritma Silhouette Coefficient* dalam menguji kualitas dari cluster yang dihasilkan.

1.6.3 Metode Perancangan

Metode perancangan yang digunakan untuk membuat sistem terkait penelitian yaitu :

- a. Merancang sebuah aplikasi analisis data produktivitas tanaman padi berbasis web.
- b. Merancang database sebagai penyimpanan data dan hasil analisis.

1.6.4 Implementasi

Tahap ini, penulis melakukan proses pembuatan sistem dari pemecahan masalah yang telah di rancang ke dalam Bahasa pemrograman *PHP* dan menyimpan data ke dalam *MySQL* sebagai databasenya, kemudian di implementasi secara interface.

1.6.5 Metode Testing

Pada tahap ini, penulis melakukan pengujian sistem menggunakan Black Box testing dan Matrix Confusion agar sistem dapat berjalan dengan baik. Apabila sistem telah di uji akan di lanjutkan dengan pemasangan pada domain dan hosting.

1.7 Sistematika Penulisan

Penulis memberikan sistematika berdasarkan bab-bab yang berurutan berdasarkan pokok-pokok permasalahan untuk mempermudah penyusunan dalam penulisan skripsi, sebagai berikut :

BAB I. Pendahuluan

Dalam bab ini tentang latar belakang penelitian, rumusan penelitian, batasan-batasan penelitian, maksud dan tujuan penelitian, metode penelitian, dan sistematika penulisan.

BAB II. Landasan Teori

Dalam bab ini berisi tentang teori yang dapat menunjang penelitian, yaitu berupa pengertian *data mining*, proses *data mining*, algoritma *K-Means*, algoritma *Silhouette Coefficient*.

BAB III. Analisis & Perancangan

Dalam bab ini berisi mengenai langkah-langkah metodologi pada penelitian ini.

BAB IV. Hasil dan Pembahasan

Dalam bab ini berisi tentang implementasi program computer berdasarkan hasil perancangan yang di buat pada BAB III, dan pengujian sistem.

BAB V. Penutup

Dalam bab ini berisi tentang kesimpulan dan saran dari seluruh penulisan penelitian ini.

DAFTAR PUSTAKA

