

BAB I

PENDAHULUAN

Bab ini menjelaskan tentang latar belakang penelitian yang dilakukan dan menjelaskan permasalahan serta solusi yang dilakukan.

1.1 Latar Belakang

Kedelai adalah komoditas pangan utama di Indonesia selain padi dan jagung. Kedelai tanaman sumber protein yang murah, sehingga dapat digunakan dalam memenuhi kebutuhan gizi masyarakat. Kebutuhan terhadap kedelai semakin meningkat dari tahun ketahun sejalan dengan bertambahnya penduduk dan meningkatkan kesadaran masyarakat terhadap makanan berprotein nabati[1]. Namun bila dilihat dari hasil produksinya masih belum memuaskan karena kendala yang harus dihadapi para petani dalam budidaya tanaman kedelai seperti gagal panen karena terserang hama dan penyakit. Penyakit pada tanaman kedelai dapat dibedakan menjadi dua kelompok besar yaitu penyakit hama dan penyakit dan usaha terbaik untuk menanggulangi penyakit tanaman kedelai dengan cara pemantauan, varietas tahan penyakit, pengendalian penyakit dengan budidaya, serta pengendalian secara mekanis, hayati, nabati dan secara kimia[2].

Gejala dapat dijadikan dasar untuk diagnosis maupun identifikasi awal suatu penyakit. Pengetahuan mengenai gejala maupun cara identifikasi diperlukan petani untuk mengetahui jenis penyakit dan dasar pengambilan keputusan pengendaliannya, namun kemampuan tersebut seringkali hanya dimiliki kalangan tertentu saja yaitu pakar, namun pakar tidak selalu ada disetiap saat atau seringkali sulit ditemui terutama di daerah pedesaan di Jawa Timur[3]. Berdasarkan data gejala penyakit hama dan tanaman terdapat 70 gejala yang menyebabkan adanya kompleksitas gejala yang tinggi sehingga dibutuhkan metode untuk menentukan gejala yang paling berpengaruh, yaitu menggunakan metode Information Gain, sedangkan untuk menentukan klasifikasi supaya lebih tepat menggunakan KNN serta efektif dalam memproses data yang besar.

Dengan Seiring dengan perkembangan teknologi yang semakin pesat, maka dibutuhkan teknologi informasi yang dapat mengklasifikasi penyakit pada tanaman kedelai sehingga dapat membantu para petani dalam mengidentifikasi suatu hama dan penyakit pada tanaman kedelai. Klasifikasi adalah proses penemuan model yang menggambarkan dan membedakan kelas data atau konsep yang berfungsi untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui[4]. Dalam penelitian ini akan menggunakan metode KNN yang digunakan untuk mengklasifikasikan hama dan penyakit tanaman kedelai.

Dalam proses melakukan klasifikasi Hama dan Penyakit pada tanaman kedelai pada penelitian ini dibutuhkan beberapa metode yaitu KNN dan seleksi fitur Information Gain. Algoritma KNN adalah sebuah metode untuk melakukan klasifikasi terhadap objek dari data uji yang memiliki jarak terdekat dengan objek tersebut[5]. Sedangkan Seleksi fitur dilakukan untuk mendapatkan fitur yang relevan dan tidak berlebihan sehingga dapat mempercepat proses klasifikasi dan dapat meningkatkan akurasi dari algoritma klasifikasi[6]. Metode seleksi fitur yang digunakan dalam penelitian ini adalah *Information gain*.

Kelebihan algoritma KNN yaitu metode klasifikasi yang simpel, mudah diimplementasikan, dan efektif dalam memproses data yang besar. Namun kekurangan algoritma ini adalah memakan waktu yang lama dalam melakukan klasifikasi, sehingga dibutuhkan seleksi fitur dalam meningkatkan performa untuk proses klasifikasi. Metode seleksi fitur yang digunakan dalam penelitian ini adalah *Information gain* yang dapat mengurangi noise yang disebabkan oleh fitur yang tidak relevan untuk mendapatkan hasil akurasi yang lebih baik [6].

Kelebihan *Information gain* adalah penerapannya yang sederhana dengan melakukan perangkingan atribut dan banyak digunakan dalam pembuatan aplikasi[7]. *Information gain* adalah metode seleksi fitur yang dapat mengoptimalkan kinerja model dengan mengurutkan fitur dan mendeteksi fitur yang mempunyai bobot gain paling besar berdasarkan kelas tertentu,

Sehingga ditemukan fitur yang paling dominan[5]. Fitur yang akan dipilih dilihat berdasarkan nilai dari *Information gain* terbesar.

Penelitian yang dilakukan oleh Tyas Setiyorini, Rizky Tri Asmono dengan judul Penerapan Metode K-Nearest Neighbor dan Information Gain Pada Klasifikasi Kinerja Siswa. Penelitian ini bertujuan untuk mengklasifikasi dan mengevaluasi kinerja siswa. Hasil dari 10 percobaan dengan nilai k dengan metode KNN didapatkan akurasi terbesar yaitu 74,068 sedangkan dengan metode KNN dan *Information Gain* didapatkan akurasi terbesar yaitu 76,553. Pada penelitian ini disimpulkan bahwa *Information Gain* mampu meningkatkan kinerja dan mengurangi jumlah fitur yang tidak relevan, sehingga penerapan KNN dan *Information Gain* dapat meningkatkan akurasi klasifikasi kinerja siswa yang lebih baik dibanding dengan menggunakan metode KNN saja[8].

Berdasarkan uraian diatas maka penulis ingin membangun sistem “Penerapan Seleksi Fitur *Information Gain* dan K-NN Untuk Klasifikasi Hama dan Penyakit Tanaman Kedelai”. Metode ini dipilih karena pada penelitian yang dilakukan oleh Tyas Setiyorini, dan Rizky Tri Asmono dengan judul “Penerapan Metode K-Nearest Neighbor dan Information Gain Pada Klasifikasi Kinerja Siswa”. Kesimpulan dari penelitian ini adalah pemilihan seleksi fitur *Information Gain* diperlukan untuk mereduksi fitur yang tidak relevan untuk mendapatkan hasil akurasi yang lebih baik daripada menggunakan metode KNN saja. Diharapkan sistem dapat menghasilkan secara akurat dalam mengklasifikasi penyakit tanaman kedelai untuk mempercepat dalam mengklasifikasi suatu hama dan penyakit pada tanaman kedelai secara dini pada tanaman kedelai sehingga meningkatkan produktivitas petani dalam budidaya tanaman kedelai.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Adapun dasar rumusan masalah antara lain :

1. Kurangnya pengetahuan para petani tentang gejala serta hama dan penyakit pada tanaman kedelai.
2. Kompleksitas gejala yang tinggi sehingga dibutuhkan metode seleksi Information Gain untuk menentukan gejala yang paling berpengaruh dan KNN untuk metode klasifikasi.

1.2.2 Pertanyaan Penelitian

Berapa akurasi terbaik yang dihasilkan metode KNN dengan seleksi fitur Information Gain?

1.3 Tujuan & Manfaat

1.3.1 Tujuan Penelitian

Bertujuan untuk mencari akurasi terbaik dengan menerapkan seleksi fitur Information Gain dan KNN untuk menentukan klasifikasi hama dan penyakit tanaman kedelai.

1.3.2 Manfaat Penelitian

Hasil dari penelitian ini diharapkan sistem mampu mengklasifikasi hama dan penyakit pada tanaman kedelai secara dini sehingga membantu para petani dalam budidaya tanaman kedelai untuk meningkatkan produktivitas tanaman kedelai.

1.4 Batasan Masalah

1. Dataset yang digunakan dalam penelitian didapat dari dinas UPT Proteksi Tanaman Pangan dan Holtikultura Provinsi Jawa Timur yang terdiri dari 352 dataset, 70 gejala (fitur), tahun 2019.
2. Parameter yang digunakan untuk pengklasifikasian ini adalah hasil hitung jarak terkecil pada metode KNN dengan 70 fitur kedelai.
3. Penelitian ini menggunakan KNN dalam mengklasifikasi hama dan penyakit pada tanaman kedelai.
4. Seleksi fitur yang digunakan dalam metode ini adalah *information gain*

5. Dalam Mengklasifikasi pada penelitian ini adalah berdasarkan 2 kategori organisme pengganggu tanaman (OPT) yaitu hama dan penyakit.
6. Hasil akhir dari penelitian ini adalah untuk menentukan hama atau penyakit pada tanaman kedelai berdasarkan gejalanya menggunakan metode KNN dan seleksi fitur *Information Gain*.

1.5 Sistematika Penulisan

Sistematika penulisan penelitian ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini membahas mengenai tentang latar belakang, rumusan masalah, batasan masalah, metode penelitian, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II KAJIAN PUSTAKA

Bab ini mengkaji beberapa teori sebagai pendukung penelitian berupa teori klasifikasi, seleksi fitur *Information Gain* dan KNN.

BAB III METODE USULAN

Bab ini mengulas tentang algoritma metode, Analisa permasalahan, dan solusi dari suatu masalah dengan memanfaatkan metode KNN dan seleksi fitur *Information Gain* pada sistem.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan tentang hasil dan pembahasan dari sistem yang sedang dikembangkan menggunakan metode yang telah diusulkan sebelumnya.

BAB V PENUTUP

Bab ini berisi tentang kesimpulan yang diambil dari penelitian ini dan saran untuk penelitian selanjutnya.