

BAB I

PENDAHULUAN

1.1 Latar Belakang

Klasifikasi adalah suatu proses pengelompokan atau pemetaan data ke dalam suatu kategori atau kelas tertentu berdasarkan dari karakteristik, fitur, atau atribut yang dimiliki. Pada konteks *data mining* atau *machine learning*, klasifikasi termasuk dalam salah satu contoh jenis dari *supervised learning*, yaitu suatu metode pembelajaran yang menggunakan data latih (*training*) atau data yang berlabel untuk digunakan ketika memprediksi label dari data baru yang belum diketahui. Sebagai contoh pada penelitian penentuan kelayakan tanah, klasifikasi dapat digunakan untuk menentukan apakah tanah dari suatu lahan termasuk dalam kategori layak atau tidak layak untuk ditanami suatu tanaman berdasarkan parameter kualitas tanah seperti pH, kelembapan, dan suhu.

Tanaman jagung merupakan salah satu komoditas pangan penting yang pertumbuhan dan produktivitasnya sangat dipengaruhi oleh kondisi lahan dan lingkungan. Kesesuaian lahan perlu diperhatikan karena setiap parameter lingkungan memiliki pengaruh terhadap proses pertumbuhan, penyerapan unsur hara, serta hasil produktivitas jagung. Oleh karena itu, identifikasi karakteristik lahan menjadi satu langkah penting dalam menentukan tingkat kelayakan suatu lahan untuk budidaya jagung. Beberapa parameter yang mempengaruhi pertumbuhan jagung adalah pH tanah, kelembapan tanah, suhu udara, dan kelembapan udara.

Nilai pH tanah berperan dalam menentukan ketersediaan unsur hara yang dapat diserap oleh tanaman. Jagung umumnya dapat tumbuh dengan baik pada tanah dengan pH berkisar antara 5,5 hingga 7,0. Pada kondisi yang terlalu rendah maupun tinggi, penyerapan unsur hara dapat terganggu sehingga berpengaruh terhadap pertumbuhan tanaman. Kelembapan tanah memiliki peran penting dalam menyediakan ketersediaan air bagi tanaman. Jagung membutuhkan kondisi tanah dengan kandungan air yang cukup, kelembapan yang terlalu rendah dapat menyebabkan tanaman kekurangan air dan kelembapan terlalu tinggi dapat menghambat pertukaran udara dalam tanah dan meningkatkan resiko gangguan

sistem perakaran. Suhu udara berpengaruh terhadap beberapa proses fisiologis tanaman seperti fotosintesis, respirasi, dan pertumbuhan jaringan. Tanaman jagung termasuk tanaman yang dapat tumbuh optimal pada kondisi suhu hangat. Perubahan suhu terlalu rendah atau terlalu tinggi dapat menghambat pertumbuhan dan menurunkan produktivitas tanaman (Waqas et al., 2021). Kelembapan udara berpengaruh terhadap pertumbuhan jagung karena berhubungan dengan proses transpirasi dan keseimbangan air dalam tanaman. Kelembapan udara yang sesuai dapat membantu menjaga laju penguapan air sehingga tanaman dapat mempertahankan kondisi fisiologis yang optimal selama masa pertumbuhan. Berdasarkan karakteristik tersebut, parameter pH tanah, kelembapan tanah, suhu udara, dan kelembapan udara dipilih dalam penelitian ini sebagai indikator untuk menentukan kelayakan lahan bagi tanaman jagung (FAO, 2022).

Kelayakan tanah menjadi salah satu faktor yang penting dalam keberhasilan budidaya pada tanaman jagung. Hal tersebut karena tingkat kondisi tanah dapat berpengaruh pada proses pertumbuhan dan hasil panen tanaman. Penentuan kelayakan tanah yang dilakukan secara manual memerlukan waktu yang lebih lama dan masih bergantung pada pengamatan secara langsung sehingga hasil penilaian dapat bersifat subjektif. Berdasarkan hal tersebut maka diperlukan suatu sistem yang dapat membantu dalam melakukan klasifikasi kelayakan tanah secara lebih cepat, objektif, dan berbasis data. Dengan melalui klasifikasi, sistem dapat mengenali pola tertentu dari data historis dan memanfaatkannya untuk melakukan prediksi maupun penentuan kategori pada data baru. Dalam penerapan pada sistem berbasis sensor, klasifikasi dapat digunakan untuk membantu menentukan kelayakan tanah berdasarkan parameter seperti pH, kelembapan, dan suhu sehingga dapat membantu dalam proses menentukan kelayakan tanah untuk tanaman jagung.

Pemanfaatan sensor dalam bidang pertanian memungkinkan proses pengumpulan data kondisi lahan dapat dilakukan secara lebih efektif dan otomatis (Garg & Alam, 2023). *Dataset* yang digunakan dalam penelitian ini berasal dari hasil pembacaan sensor yang terintegrasi dengan mikrokontroler Arduino. Sensor tersebut bekerja dengan cara mengukur beberapa parameter kualitas tanah dan lingkungan seperti pH tanah, kelembapan tanah, suhu, dan kelembapan udara yang

kemudian disimpan dalam bentuk data tabular. Setiap kombinasi dari nilai parameter tersebut merepresentasikan kondisi aktual lahan pada saat pengukuran dilakukan. Pada data tabular, setiap baris merepresentasikan satu observasi, sedangkan setiap kolom menunjukkan atribut atau fitur yang memiliki nilai terstruktur seperti numerik atau kategorikal. Data tabular banyak digunakan karena mudah dianalisis dan kompatibel dengan berbagai algoritma klasifikasi seperti *K-Nearest Neighbor* (KNN), *Decision Tree*, hingga *Naive Bayes*. Dalam konteks penelitian ini klasifikasi digunakan untuk menentukan tingkat kelayakan dari tanah berdasarkan dengan parameter-parameter terukur. Dengan memanfaatkan data hasil pembacaan sensor sebagai fitur *input*, proses klasifikasi dapat dilakukan untuk mengidentifikasi apakah kondisi tanah tergolong layak atau tidak untuk digunakan pada budidaya tanaman, sehingga dapat mendukung pengambilan keputusan berbasis data secara lebih objektif dan terukur.

Parameter seperti pH tanah, kelembapan tanah, suhu udara, dan kelembapan udara memiliki pengaruh terhadap tingkat kesuburan tanah dan pertumbuhan pada tanaman jagung. Nilai parameter tersebut dapat menunjukkan apakah kondisi lingkungan sesuai untuk mendukung proses budidaya tanaman. Dengan memanfaatkan sensor, proses pengambilan data dapat dilakukan secara lebih konsisten sehingga informasi yang diperoleh dapat digunakan sebagai dasar dalam proses klasifikasi kelayakan tanah.

Pada penelitian ini, metode *K-Nearest Neighbor* (KNN) dipilih karena memiliki konsep yang sederhana dan mampu melakukan klasifikasi berdasarkan dengan kedekatan jarak antar data. Selain itu, metode *K-Nearest Neighbor* (KNN) cukup sesuai untuk digunakan pada data numerik hasil dari pembacaan sensor. Hal tersebut karena proses klasifikasinya dilakukan berdasarkan dengan kemiripan nilai parameter antar data.

Pada penelitian (Jr, 2025) dilakukan analisis klasifikasi komunitas mikroba tanah berdasarkan berbagai sifat tanah dengan membandingkan kinerja antara algoritma *K-Nearest Neighbor* (KNN) dengan algoritma *Naive Bayes*. Penelitian tersebut dilakukan bertujuan untuk mencari model dengan akurasi dan efisiensi komputasi yang lebih tinggi. Dataset yang digunakan adalah dataset mikroba tanah

dengan fitur pH, suhu, dan tekstur tanah, beserta dengan data kelimpahan mikroba. Dari penelitian tersebut didapatkan akurasi dari algoritma *K-Nearest Neighbor* (KNN) sebesar 85% dan algoritma *Naive Bayes* sebesar 80%. Penulis menyimpulkan bahwa algoritma *K-Nearest Neighbor* (KNN) direkomendasikan untuk tugas yang mengutamakan akurasi karena kinerja klasifikasinya dinilai lebih baik, dapat menangkap hubungan kompleks dan linear dalam data, sedangkan algoritma *Naive Bayes* lebih cocok untuk dataset yang lebih besar yang membutuhkan komputasi cepat.

Pada penelitian (Nurdiawan et al., 2020) dilakukan perbandingan antara algoritma *K-Nearest Neighbor* (KNN) dengan algoritma *Decision Tree* dalam klasifikasi kelembapan. Penelitian tersebut bertujuan untuk membandingkan antara kedua algoritma tersebut untuk dapat menentukan klasifikasi tanah dengan parameter hasil lembab dan kering. Dataset yang digunakan pada penelitian ini berasal dari pembacaan sensor. Berdasarkan hasil pembacaan sensor kelembapan pada algoritma *Decision Tree* didapatkan akurasi sebesar 55,77%, *precision* (kelas lembab) sebesar 47,37%, *recall* (kelas kering) sebesar 62,69% dan tingkat kesalahan (RMSE) lebih tinggi sebesar 0,0514. Sedangkan untuk algoritma *K-Nearest Neighbor* (KNN) didapatkan akurasi sebesar 72,69%, *precision* (kelas lembab) sebesar 73,47%, *recall* (kelas kering) sebesar 80,60% dan tingkat kesalahan (RMSE) lebih rendah sebesar 0,0294. Berdasarkan penelitian yang telah dilakukan, penulis menyimpulkan bahwa algoritma *K-Nearest Neighbor* (KNN) merupakan algoritma yang lebih handal dalam menangani variasi data sensor kelembapan karena pendekatannya yang berbasis jarak antar data poin, sementara algoritma *Decision Tree* cenderung mengalami penurunan performa pada dataset yang digunakan pada penelitian tersebut.

Berdasarkan dari penelitian terdahulu tersebut, metode *K-Nearest Neighbor* (KNN) memperoleh tingkat akurasi tinggi dan menunjukkan performa yang lebih baik daripada algoritma pembandingan. Namun penelitian sebelumnya umumnya masih berfokus pada parameter tertentu seperti kelembapan tanah, belum menggabungkan beberapa parameter lingkungan yang lain secara bersamaan. Selain itu, proses pelabelan data pada beberapa penelitian masih dilakukan secara

manual sehingga memerlukan waktu lebih lama dan berpotensi timbulnya ketidakkonsistenan data. Oleh karena itu, berdasarkan latar belakang tersebut maka dilakukan penelitian terkait klasifikasi kelayakan tanah pada tanaman jagung dengan menggunakan kombinasi parameter pH tanah, kelembapan tanah, suhu udara dan kelembapan udara dengan menggunakan metode *K-Nearest Neighbor* (KNN). Penelitian ini juga memanfaatkan proses pelabelan awal (*ground truth*) secara otomatis menggunakan logika pengkondisian *if else* pada sistem mikrokontroler. Klasifikasi kelayakan tanah ini bertujuan untuk menentukan apakah suatu lahan termasuk dalam kategori Layak dan Tidak Layak untuk ditanami tanaman jagung. Parameter yang digunakan pada penelitian ini adalah pH tanah, kelembapan tanah, suhu udara, dan kelembapan udara.

Pada penelitian ini proses klasifikasi digunakan untuk menentukan kelayakan tanah pada tanaman jagung berdasarkan parameter hasil pembacaan sensor tanah. Data penelitian diperoleh melalui proses akuisisi data menggunakan sensor yang terintegrasi dengan mikrokontroler Arduino UNO. Sistem tersebut digunakan untuk mengukur parameter tanah dan lingkungan yang kemudian menghasilkan data untuk digunakan sebagai *dataset* penelitian. Penggunaan sistem mikrokontroler dalam proses pengambilan data membantu proses pelabelan awal (*ground truth*) secara otomatis berdasarkan parameter yang telah ditentukan, sehingga data dapat langsung dikategorikan ke dalam kelas Layak dan Tidak Layak. Dengan adanya proses tersebut maka proses pelabelan tidak dilakukan secara manual melainkan otomatis oleh sistem sehingga menjadi lebih efisien dibandingkan dengan pelabelan manual satu per satu. Label yang dihasilkan kemudian digunakan dalam proses klasifikasi menggunakan metode *K-Nearest Neighbor* (KNN).

1.2 Rumusan Masalah

Berdasarkan pada uraian latar belakang maka dapat dipaparkan rumusan masalah sebagai berikut.

1.2.1 Permasalahan

Salah satu permasalahan dalam penerapan metode *K-Nearest Neighbor* (KNN) adalah penentuan nilai K yang optimal. Nilai K berperan penting dalam

proses klasifikasi karena menentukan jumlah tetangga terdekat yang digunakan dalam pengambilan keputusan. Pemilihan nilai K yang tidak tepat dapat mempengaruhi tingkat akurasi hasil klasifikasi. Oleh karena itu, penentuan nilai K menjadi salah satu permasalahan dalam penelitian ini.

1.2.2 Solusi Permasalahan

Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan pendekatan pengujian beberapa variasi nilai K dalam proses klasifikasi menggunakan metode *K-Nearest Neighbor* (KNN). Dengan melakukan pengujian beberapa variasi nilai K dan evaluasi performa menggunakan metrik akurasi, *precision*, *recall*, dan *f1-score*, diharapkan dapat diperoleh nilai k yang paling optimal dalam proses klasifikasi kelayakan tanah pada tanaman jagung.

1.2.3 Pertanyaan Penelitian

1. Bagaimana pengaruh penentuan nilai K terhadap hasil klasifikasi kelayakan tanah pada tanaman jagung menggunakan metode *K-Nearest Neighbor* (KNN)?
2. Bagaimana performa dari metode *K-Nearest Neighbor* (KNN) dalam mengklasifikasikan kelayakan tanah pada tanaman jagung berdasarkan nilai akurasi, *precision*, *recall*, dan *f1-score*?
3. Nilai K berapakah yang menghasilkan performa terbaik dalam proses klasifikasi kelayakan tanah pada tanaman jagung?

1.3 Tujuan & Manfaat

Adapun tujuan dan manfaat dari penelitian ini adalah sebagai berikut.

1.3.1 Tujuan

Berdasarkan pada uraian latar belakang maka dapat dipaparkan tujuan dari penelitian ini adalah sebagai berikut.

1. Untuk mengetahui pengaruh dari penentuan nilai K terhadap hasil klasifikasi kelayakan tanah pada tanaman jagung menggunakan metode *K-Nearest Neighbor* (KNN).
2. Untuk mengevaluasi performa dari metode *K-Nearest Neighbor* (KNN) dalam klasifikasi kelayakan tanah pada tanaman jagung menggunakan

metrik akurasi, *precision*, *recall*, dan *f1-score* berdasarkan *confusion matrix*.

3. Untuk mengetahui mana nilai K yang menghasilkan performa terbaik dalam proses klasifikasi kelayakan tanah pada tanaman jagung.

1.3.2 Manfaat

Adapun manfaat yang diharapkan dapat tercapai dari penelitian ini adalah sebagai berikut.

1. Mengetahui performa dari Algoritma *K-Nearest Neighbor* (KNN) untuk klasifikasi kelayakan tanah pada tanaman jagung.
2. Sebagai referensi untuk penelitian lebih lanjut mengenai klasifikasi kelayakan tanah untuk tanaman jagung.

1.4 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut.

1. Parameter fitur yang digunakan dalam penelitian ini dibatasi pada empat parameter, yaitu pH tanah, kelembapan tanah, suhu udara, dan kelembapan udara.
2. Klasifikasi kelayakan terdiri dari dua kelas, yaitu Layak dan Tidak Layak.
3. Data diperoleh menggunakan sistem berbasis Arduino yang terintegrasi dengan sensor untuk mengukur parameter lahan.
4. Klasifikasi dilakukan dengan menggunakan Algoritma *K-Nearest Neighbor* (KNN).
5. Analisis dilakukan dengan menggunakan bahasa pemrograman *Python*.
6. *Dataset* yang digunakan berjumlah 4.201 data hasil pengukuran lapangan.
7. Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik akurasi, *precision*, *recall*, dan *f1-score*.

1.5 Kontribusi Penelitian

Penelitian ini memberikan kontribusi dalam penerapan metode *K-Nearest Neighbor* (KNN) untuk klasifikasi kelayakan tanah pada tanaman jagung berdasarkan parameter hasil pembacaan sensor tanah dan lingkungan. Data yang digunakan diperoleh melalui sistem berbasis mikrokontroler Arduino UNO sehingga proses pengambilan data dapat dilakukan secara otomatis. Selain itu,

penelitian ini juga menerapkan proses pelabelan awal (*ground truth*) secara otomatis menggunakan aturan pengkondisian pada sistem mikrokontroler berdasarkan parameter kualitas tanah yang diperoleh dari literatur pertanian.

Pemanfaatan sensor tanah dalam bidang pertanian dapat membantu memperoleh data kondisi lingkungan dan lahan secara lebih efektif sehingga mendukung proses pengambilan keputusan. Pemanfaatan sensor tanah dapat memungkinkan pengumpulan data tanah dilakukan secara lebih efektif sehingga dapat digunakan sebagai dasar dalam proses analisis dan klasifikasi berbasis data secara lebih efisien dan *real-time* (Hendrawati et al., 2025). Oleh karena itu, hasil penelitian ini diharapkan dapat menjadi referensi dalam pengembangan sistem klasifikasi kelayakan tanah menggunakan metode *machine learning* pada bidang pertanian.

1.6 Sistematika Laporan

Adapun sistematika penulisan penelitian yang digunakan terdiri dari beberapa bagian, yaitu :

1. BAB I PENDAHULUAN

Bab I pendahuluan berisi tentang penjelasan latar belakang penelitian yang dilakukan, yang dilengkapi dengan beberapa penelitian terdahulu sebagai acuan dalam melakukan penelitian. Selanjutnya pada latar belakang terdapat rumusan masalah, tujuan, dan manfaat dari penelitian yang dilakukan. Agar penelitian lebih terarah pada bab ini juga diberikan batasan masalah dan juga sistematika penulisan dari penelitian.

2. BAB II KAJIAN PUSTAKA

Bab II kajian pustaka berisi tentang teori penunjang yang digunakan dalam penyelesaian rumusan masalah. Teori penunjang merupakan teori-teori yang berkaitan dengan penelitian dan digunakan sebagai referensi. Teori penunjang didapatkan dari beberapa sumber, seperti jurnal, buku, dan artikel.

3. BAB III METODE USULAN

Bab III metode penelitian berisi tentang rancangan usulan penelitian yang akan dilakukan. Isi dari bab ini berhubungan dengan rumusan masalah pada

kajian pustaka. Langkah-langkah yang disusun dilakukan untuk menjawab rumusan masalah.

4. BAB IV HASIL DAN PEMBAHASAN

Bab IV hasil dan pembahasan berisi tentang uraian analisis dan pembahasan yang berkaitan tentang pengambilan data, pengolahan data dengan menggunakan algoritma klasifikasi untuk mendapatkan hasil akurasi dari algoritma yang digunakan serta menjelaskan hasil dari penelitian yang telah dilakukan.

5. BAB V KESIMPULAN DAN SARAN

Bab V kesimpulan dan saran berisi tentang kesimpulan dari penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.