

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kebutuhan pangan menjadi hal yang sangat penting bagi mayoritas penduduk Indonesia karena beras merupakan sumber makanan utama. Berdasarkan data Kerangka Sampel Area (KSA) dari BPS, produksi padi nasional dalam bentuk Gabah Kering Giling (GKG) pada tahun 2018 mencapai 59,2 juta ton, mengalami penurunan sebesar 7,76% menjadi 54,60 juta ton pada tahun 2019. Pada tahun 2020, produksi sedikit meningkat sebesar 0,08% menjadi 54,65 juta ton, namun pada tahun 2021 mengalami penurunan sebesar 0,43%, menjadi 54,41 juta ton atau setara dengan 34,84 juta ton beras. Salah satu faktor yang dapat mempengaruhi penurunan dari produksi pangan tersebut adalah ketersediaan unsur hara di dalam tanah atau perawatan tanaman pangan yang kurang. Unsur hara tanah seperti nitrogen, fosfor, dan kalium sangat penting bagi pertumbuhan dan perkembangan tanaman padi. Namun ketersediaan unsur hara tersebut di dalam tanah dapat berbeda-beda tergantung lokasi dan jenis tanah (Umm E Farwa, 2020).

Unsur hara tanah adalah bagian terpenting dari pertanian. Sebelum melakukan penanaman, petani harus mengecek atau melihat kondisi tanah terlebih dahulu. Unsur hara tanah nitrogen, fosfor dan kalium merupakan parameter penting yang dapat menunjang proses pertumbuhan tanaman. Unsur nitrogen membantu pertumbuhan daun tanaman, unsur Fosfor berperan dalam pertumbuhan akar, perkembangan bunga dan buah, dan unsur kalium membantu fungsi tanaman secara keseluruhan (Umm E Farwa, 2020). Kesuburan tanah dapat diukur dari

kemampuan tanaman dalam menyediakan unsur hara tanah yang menunjang pertumbuhan dan produksi tanaman. Unsur hara tanah NPK dapat membantu dalam pertumbuhan tanaman sehingga dibutuhkan dalam jumlah yang banyak. Oleh karena itu, perlu dilakukan pengukuran kandungan unsur hara tanah untuk mengetahui seberapa baik status kesuburan tanah (M. tahir, 2021). Selain itu, penting bagi para petani untuk memahami hubungan antara temperatur, *humidity*, pH, *rainfall* dan unsur hara tanah nitrogen, posfor, dan kalium. Karena, antara unsur hara tanah NPK, temperatur, *humidity*, pH, *rainfall* memiliki hubungan yang sangat signifikan dan penting dalam konteks pertanian dan pertumbuhan tanaman. Selain itu, dalam pengelolaan tanah, pemahaman mengenai hubungan antara unsur hara tanah NPK, *temperature*, *humidity*, pH, dan *rainfall* dapat membantu petani dalam mengatur pemberian pupuk dan irigasi sesuai dengan kebutuhan tanaman, serta mengurangi penggunaan pupuk dan air secara efisien (Rudy Gunawan, 2019) (Siswanto, 2018).

Namun pemahaman petani mengenai unsur hara tanah masih kurang dan informasi mengenai unsur hara tanah tersebut masih terbatas. Namun seiring berkembangnya dunia pertanian, para petani di beberapa negara maju dilengkapi dengan laboratorium pengujian yang banyak digunakan untuk menentukan nilai NPK (Umm E Farwa, 2020).

Tetapi, para petani di negara berkembang khususnya Indonesia, masih belum mengetahui cara memanfaatkan informasi tersebut. Banyak produsen pupuk yang mencetak nilai nutrisi dari kandungan NPK pada kantong pupuk, namun petani belum mengetahui cara memanfaatkan nilai nutrisi dari kandungan NPK

(Umm E Farwa, 2020). Kurangkanya informasi dan pemahaman tentang nutrisi penting ini mengakibatkan hasil panen berkurang atau tidak memadai. Oleh karena itu, diperlukan strategi untuk membantu petani mengatasi permasalahan tersebut.

Penelitian sebelumnya yang dilakukan oleh (Umm E Farwa, 2020) berjudul “*Prediction of Soil Macronutrients Using Machine Learning Algorithm*”. Dalam penelitian ini, algoritma *machine learning* diterapkan untuk mengetahui hubungan antara variabel *independent* dan *dependent* untuk analisis data unsur hara tanah menggunakan *Linear regression, Ridge Regression, Bayesian Regression*. Variabel bebas (*independent*) meliputi kelembapan, suhu, pH tanah, dan kapasitas tukar kation (KTK), dan variabel terikat (*dependent*) meliputi nitrogen, fosfor, dan kalium (NPK). Penelitian ini menyimpulkan hubungan antara posfor, kalium, pH tanah, dan KTK. Nitrogen, kelembapan tanah, dan suhu digunakan untuk memperoleh kandungan NPK tanah menggunakan algoritma *machine learning*.

Berdasarkan tinjauan literatur yang dilakukan pada penelitian sebelumnya, penelitian ini menggunakan algoritma pembelajaran mesin dan termasuk pembelajaran terbimbing (Trivusi, 2022). Pembelajaran yang diawasi mencakup berbagai jenis teknik algoritma, seperti regresi linier, *K-nearest neighbours*, *support vector machine (SVM)*, *Naive Bayes*, *random forest*, dan *neural network* (Algoritma, 2022). Penelitian ini menggunakan metode *random forest* yang digunakan untuk memprediksi unsur hara tanah pada lahan pertanian berdasarkan variabel *independent* (bebas) yang diantaranya temperatur, *humidity*, pH dan *rainfall*. Algoritma *Random forest* dapat dijalankan pada proses klasifikasi sekaligus regresi dan mampu menghasilkan prediksi dengan tingkat akurasi tinggi

(Riki Supriyadi, 2020). Selain itu, algoritma *Random forest* memiliki kelebihan diantaranya kuat terhadap data *outlier*, risiko *overfitting* rendah, serta dapat menghasilkan prediksi dengan tingkat akurasi tinggi (Algoritma, 2022) (Trivusi, 2022). Namun, algoritma *Random forest* juga memiliki beberapa kelemahan diantaranya waktu komputasi pada dataset berskala besar relatif lambat, serta cenderung bias saat berhadapan dengan variabel kategorikal (Trivusi, 2022). Mengatasi hal tersebut dapat disesuaikan dengan penelitian yang sedang dilakukan, karena pada penelitian ini untuk dataset yang digunakan tidak terlalu besar, jadi dalam melakukan prediksi atau waktu komputasinya tidak akan relatif lambat, serta yang diutamakan dalam penelitian yang dilakukan ini, yaitu lebih untuk melihat performa kinerja dari algoritma *Random forest* dalam melakukan prediksi dengan menghasilkan nilai akurasi dan hasil evaluasi yang baik.

Dalam Konteks prediksi unsur hara tanah, algoritma *Random forest* dapat memberikan akurasi tertinggi. Hal tersebut di dapatkan berdasarkan dari penelitian sebelumnya yang dilakukan oleh Madhuri Shripathi Rao, Arushi Singh, N.V. Subba Reddy dan Dinesh U Acharya dengan judul “*Crop Prediction Using Machine Learning*”. Penelitian ini dilakukan bertujuan untuk menemukan model terbaik untuk prediski tanaman, yang dapat membantu para petani dalam memutuskan jenis tanaman yang akan ditanami berdasarkan kondisi iklim dan nutrisi atau unsur hara tanah. Dalam penelitian ini menggunakan beberapa metode yang meliputi metode *K -Nearest Neighbor*, *Decision Tree*, dan *Random forest Classifier*. Pada penelitian ini menyatakan bahwa algoritma *Random forest* memberikan hasil akurasi tertinggi

diantara ketiga metode dengan akurasi sebesar 99,32% (Madhuri Shripathi Rao, 2021).

Dengan demikian, dalam penelitian ini akan dilakukan penerapan salah satu metode *machine learning* yaitu algoritma *Random forest* untuk memprediksi unsur hara tanah berdasarkan temperatur, *humidity*, pH dan *rainfall* untuk mendapatkan hasil akurasi terbaik.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan maka rumusan masalah pada penelitian ini adalah bagaimana performa kinerja dan ketepatan dari algoritma *Random forest* dalam memprediksi unsur hara tanah?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini yaitu menerapkan model prediksi unsur hara tanah yang akurat menggunakan algoritma *Random forest*. Melakukan pengukuran kinerja performa algoritma *Random forest* dalam memprediksi unsur hara tanah agar hasil dari pengukuran kinerja tersebut dapat membantu para petani dalam menentukan unsur hara tanah atau mengelola tanah. Serta, meningkatkan pemahaman tentang penggunaan algoritma *Random forest* dalam memprediksi unsur hara tanah.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini yaitu dalam bidang keilmuan salah satunya penelitian dapat memperluas pemahaman khususnya tentang prediksi unsur hara tanah dengan menggunakan pendekatan berbasis data, serta efektivitas penggunaan

algoritma *Random forest* dalam memprediksi unsur hara tanah yang dapat menjadi salah satu acuan bagi penelitian lanjutan. Selain itu, penelitian ini tidak hanya dapat bermanfaat bagi para petani, tetapi memiliki potensi untuk para praktisi pertanian dan pengelola lingkungan dalam pengambilan keputusan salah satunya untuk manajemen tanah.

1.5 Batasan Masalah

Berdasarkan permasalahan yang didapatkan, agar penelitian tidak keluar dari topik maka di tentukan batasan masalah pada penelitian ini, yaitu :

1. Algoritma yang diterapkan dalam penelitian ini adalah algoritma *Random forest*.
2. Dataset yang digunakan untuk penelitian ini yaitu di peroleh dari *kaggle* (<https://www.kaggle.com/datasets/chitrakumari25/smart-agricultural-production-optimizing-engine>).
3. Data yang digunakan merupakan data *temperature, humidity, pH, rainfall*, unsur hara tanah (Nitrogen, Posfor, dan Kalium).
4. Tanaman yang di prediksi dalam studi ini yaitu tanaman pangan.
5. Variabel *Dependent* (terikat) dari penelitain ini terdiri dari variabel Nitrogen, Pospor, Kalium.
6. Variabel *Independent* (tidak terikat) terdiri *temperature, humidity, pH, rainfall*.
7. Penelitian ini lebih perfokus terhadap pengukuran performa yang dilakukan dengan menggunakan dataset yang telah di tentukan antara variabel

dependent (terikat) dan variabel *independent* (tidak terikat) serta menggunakan model *Random Forest*.

8. Mengukur performa kinerja dari metode *Random forest* dengan menggunakan *R-squared* (R²) dan *Mean Absolute Percentage Error* (MAPE).

1.6 Metodologi Penelitian

Metodologi penelitian meliputi pembahasan terkait waktu dan tempat penelitian, tahapan penelitian, pendekatan penelitian, jenis penelitian, variabel penelitian, serta objek penelitian. Tahapan metode penelitian dalam penelitian ini adalah sebagai berikut :

1. Studi Literatur

Pada tahap studi literatur dilakukan pencarian terkait penelitian sebelumnya mengenai prediksi dan algoritma yang digunakan yang mampu mendukung penelitian yang akan dilakukan secara teoritis. Sumber literatur yang digunakan diantaranya dari jurnal, skripsi terdahulu, dan sumber Pustaka dari internet.

2. Pengumpulan dan Analisis Data

Pada tahap ini melakukan pengambilan data yang diperlukan yaitu data dari penelitian sebelumnya. Data yang diperlukan dalam penelitian ini merupakan data temperatur, *humidity*, pH, *rainfall* dan unsur hara tanah (Nitrogen, Posfor, Kalium) untuk dijadikan sampel dalam menentukan unsur hara tanah Nitrogen, Pospor, dan Kalium dengan mengimplementasikan algoritma *Random forest* untuk dapat memperoleh hasil akurasi yang terbaik.

3. Pengolahan Data

Pada tahap ini dilakukan pengolahan data untuk memperoleh data yang berkualitas. Serta data dengan bentuk yang mudah di pahami sesuai dengan kebutuhan penelitian.

4. *Modelling*

Pada tahapan ini akan dilakukan tahap pemodelan dengan menggunakan algoritma *Random forest* untuk memprediksi unsur hara tanah dengan memperoleh hasil akurasi terbaik.

5. Evaluasi

Pada tahap ini dilakukan proses evaluasi dengan tujuan untuk menguji performa model dari algoritma *Random forest* untuk mendapatkan hasil akurasi dari algoritma dalam melakukan prediksi.

6. Kesimpulan

Tahapan ini merupakan tahapan akhir pada penelitian yaitu berupa hasil prediksi dengan mengimplementasikan algoritma *Machine Learning* yaitu *Random forest*.

1.7 Sistematika Penulisan

Adapun aturan dan sistematika penulisan yang digunakan dalam penelitian ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab pendahuluan berisi latar belakang umum seperti gambaran secara garis besar tentang isi laporan, rumusan masalah, batasan masalah pada penelitian, tujuan penelitian, manfaat penelitian yang diperoleh, metode penelitian serta sistematika penulisan laporan tugas akhir.

BAB II LANDASAN TEORI

Bab landasan teori meliputi teori-teori yang saling berhubungan dengan penelitian seperti konsep serta metode dan algoritma yang terkait dengan penelitian ini. Pada bab ini juga berisi penjelasan dari penelitian sebelumnya yang relevan dengan penelitian ini.

BAB III METODOLOGI

Bab metodologi berisi uraian tentang metode yang digunakan dalam melakukan penelitian, mulai dari studi literatur, identifikasi masalah, pengumpulan data, data *understanding*, data *preprocessing*, pembagian data (*data training* dan *data testing*), pemodelan, evaluasi, ancaman terhadap validitas, kesimpulan dan saran.

BAB IV HASIL DAN PEMBAHASAN

Bab hasil dan pembahasan berisi pemaparan hasil serta pembahasan terhadap perancangan pada bab sebelumnya, yaitu bagaimana alur sebuah data apabila dilakukan pengolahan data, pemodelan serta evaluasi data dengan menggunakan algoritma *Random forest* serta menentukan keakuratan dan performa model *Random forest* dalam memprediksi unsur hara tanah.

BAB V KESIMPULAN DAN SARAN

Bab ini merupakan bab terakhir berisi kesimpulan dan saran dari hasil penelitian, serta merupakan garis besar dari metode penelitian yang telah dilakukan. Kesimpulan adalah hasil akhir dari penelitian yang dilakukan, sedangkan saran berisi tentang rekomendasi sesuai dengan keterbatasan yang ada pada penelitian yang dilakukan.