

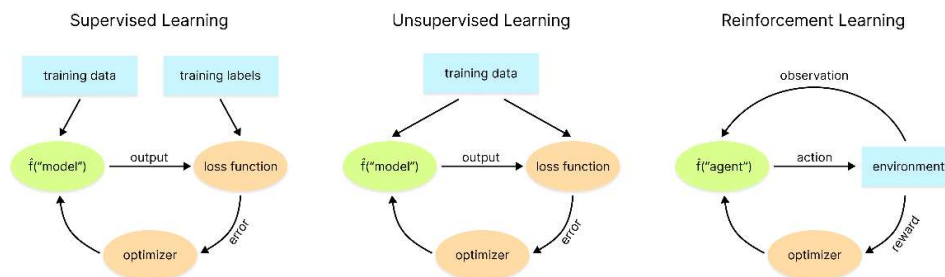
BAB 2

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Machine Learning

Machine Learning merupakan bagian dari *Artificial Intelligence* (AI) yang memanfaatkan data sebagai dasar bagi sistem dalam mengenali pola dan menghasilkan suatu keputusan atau prediksi. Pendekatan ini memungkinkan sistem memperoleh kemampuan prediktif melalui proses pembelajaran dari data yang tersedia, sehingga dapat digunakan untuk menangani permasalahan yang memiliki pola data kompleks (HASYDNA & DINATA, 2020). Secara umum, metode pembelajaran pada *Machine Learning* seperti Gambar 2.1 terdiri atas *Supervised Learning*, *Unsupervised Learning*, dan *Reinforcement Learning*.



Gambar 2. 1 Jenis *Machine Learning*

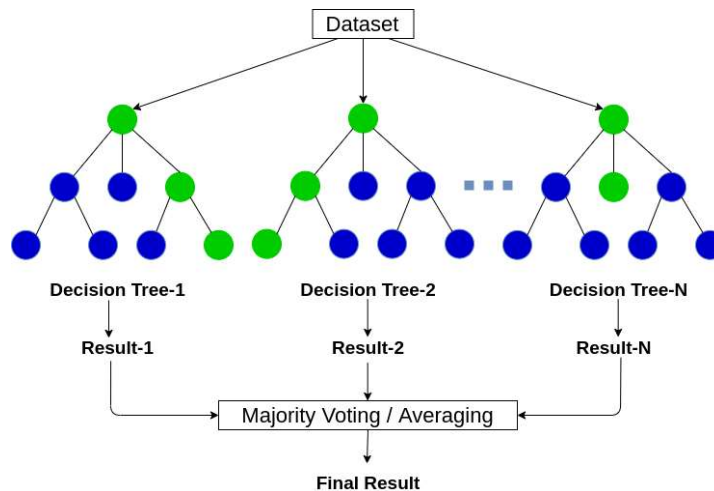
Sumber : <https://www.aisafetybook.com/textbook>

Penelitian ini menerapkan pendekatan *Supervised Learning* karena model dibangun menggunakan dataset yang telah memiliki variabel masukan dan nilai target. Variabel masukan berupa karakteristik lahan digunakan untuk mempelajari hubungan terhadap target berupa kebutuhan pupuk. Dengan demikian, model dapat memanfaatkan pola data historis untuk menghasilkan estimasi kebutuhan pupuk pada data baru.

2.1.2 Algoritma Random Forest

Random Forest merupakan algoritma *Machine Learning* berbasis kumpulan *decision tree* yang dapat digunakan untuk menyelesaikan permasalahan klasifikasi

maupun regresi. Algoritma ini membentuk beberapa pohon keputusan dari sampel data yang berbeda, kemudian menggabungkan hasil dari seluruh pohon untuk memperoleh keluaran akhir. Pendekatan tersebut dapat membantu model menghasilkan prediksi yang lebih stabil serta mengurangi kecenderungan *overfitting* (Mahendra Alvanof & Kesuma Dinata, 2024).



Gambar 2. 2 Proses kerja *Random Forest*

Sumber : <https://medium.com/@jainvidip/understanding-random-forest>

Pada Gambar 2.2 proses kerja *Random Forest* membentuk beberapa *decision tree* menggunakan sampel data yang diperoleh melalui proses *bootstrap sampling*. Pada setiap pembentukan pohon, sebagian fitur dipilih secara acak untuk menentukan percabangan. Proses tersebut dilakukan hingga terbentuk sejumlah pohon yang selanjutnya digunakan secara bersama-sama untuk menghasilkan prediksi.

2.1.3 Preprocessing Data

Preprocessing merupakan tahap persiapan dataset sebelum data digunakan dalam pembentukan model. Tahap ini dilakukan untuk memastikan bahwa data yang digunakan memiliki format yang sesuai, konsisten, dan dapat diterima oleh algoritma. Pada penelitian ini, proses *preprocessing* mencakup pembersihan data, pemilihan fitur, transformasi variabel kategorikal menggunakan *One-Hot Encoding*, serta pembagian dataset menjadi data pelatihan dan pengujian. dapat dilihat pada Gambar 2.3

- *Data Cleaning*

Data cleaning dilakukan untuk memeriksa kondisi dataset sebelum memasuki tahap pemodelan. Pemeriksaan mencakup keberadaan nilai yang kosong, data yang berulang, kesalahan penulisan, serta ketidaksesuaian format antar variabel. Tahap ini diperlukan agar data yang masuk ke proses pelatihan memiliki struktur yang konsisten dan tidak menimbulkan gangguan terhadap proses pembelajaran model.

- *Feature Importance*

Digunakan untuk melihat kontribusi relatif masing-masing fitur terhadap proses prediksi yang dilakukan oleh model *Random Forest*. Nilai *feature importance* menunjukkan kontribusi relatif masing-masing fitur dalam proses pembentukan prediksi model. Dalam penelitian ini, analisis *Feature Importance* digunakan untuk mengetahui kontribusi variabel seperti luas lahan, jumlah pohon, usia tanaman, jenis tanah, jenis pupuk, dan kondisi daun terhadap prediksi kebutuhan pupuk.

- *One-Hot Encoding*

One-Hot Encoding merupakan salah satu teknik transformasi data kategorikal menjadi data numerik dengan memberikan nilai bilangan bulat (integer) pada setiap kategori. Proses ini diperlukan karena sebagian besar algoritma *Machine Learning*, termasuk *Random Forest Regression*, memerlukan data masukan dalam bentuk numerik. Setiap kategori akan diberikan kode numerik yang berbeda tanpa mengubah jumlah kategori yang ada. Karena *Random Forest* merupakan algoritma berbasis pohon keputusan (*tree-based*), penggunaan *One-Hot Encoding* dinilai sesuai dan tidak menyebabkan peningkatan jumlah fitur.

- *Train-Test Split*

Train-test split digunakan untuk memisahkan dataset menjadi dua kelompok dengan fungsi yang berbeda. Data pelatihan digunakan oleh model untuk mempelajari pola hubungan antara variabel masukan dan target, sedangkan data pengujian digunakan untuk melihat kemampuan model ketika diberikan data yang tidak digunakan selama proses pembelajaran.

Pada penelitian ini, dataset dibagi dengan perbandingan 80% untuk pelatihan dan 20% untuk pengujian.

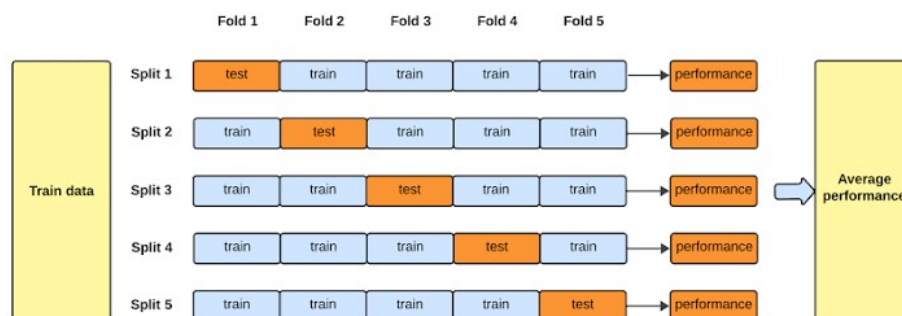


Gambar 2. 3 Preprocessing data

Sumber: <https://binus.ac.id/bagaimana-melakukan-preprocessing-data>

2.1.4 Cross Validation

Berdasarkan Gambar 2.4 *Cross Validation* digunakan untuk mengetahui kestabilan kinerja model ketika diterapkan pada bagian data yang berbeda. Pada metode *K-Fold Cross Validation*, dataset dibagi menjadi beberapa kelompok (*fold*). Setiap kelompok secara bergantian digunakan sebagai data validasi, sedangkan kelompok lainnya digunakan untuk pelatihan. Seluruh proses dilakukan sampai setiap *fold* memperoleh kesempatan sebagai data validasi, kemudian hasil evaluasi dari seluruh percobaan dirata-ratakan. (Azis dkk., 2020). Penelitian ini menggunakan 5-Fold Cross Validation untuk mengevaluasi kestabilan performa *Random Forest Regression* sebelum dilakukan optimasi hiperparameter.

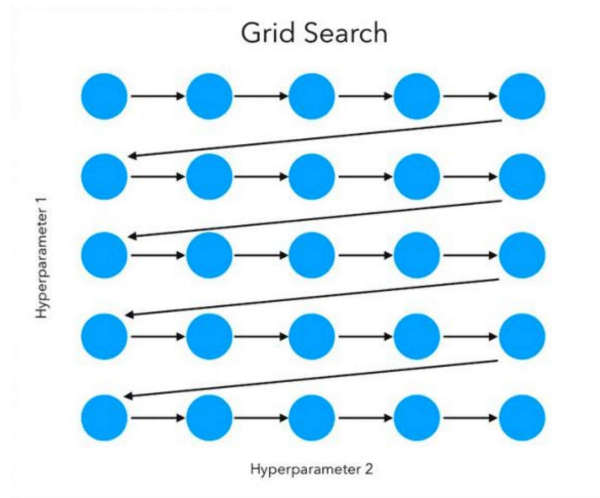


Gambar 2. 4 Cross Validation

Sumber: <https://towardsdatascience.com/how-to-cross-validation/>

2.1.5 GridSearchCv

GridSearchCV seperti Gambar 2.5 digunakan untuk mencari kombinasi hiperparameter yang memberikan performa model paling baik. Proses pencarian dilakukan dengan menguji beberapa kombinasi parameter yang telah ditentukan dan mengevaluasinya menggunakan *Cross Validation*. Pada penelitian ini, parameter yang diuji meliputi *n_estimators*, *max_depth*, *min_samples_split*, dan *min_samples_leaf*. Kombinasi dengan performa terbaik kemudian digunakan sebagai konfigurasi model final sebelum diterapkan pada sistem berbasis web.



Gambar 2. 5 *GridSearchCV*

Sumber: <https://medium.com/gridsearchcv-library-d28b3f8f17be>

2.1.6 Evaluasi Kinerja

Dalam penelitian ini, teknik evaluasi yang digunakan untuk mengukur performa algoritma *Random Forest* dalam memprediksi kebutuhan pupuk pada lahan kelapa sawit di antaranya:

1. *Root Mean Squared Error* (RMSE)

RMSE mengukur rata-rata kesalahan prediksi dengan memberi bobot lebih besar pada *Error* yang lebih besar, sehingga lebih sensitif terhadap *outlier*. Rumus RMSE

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - Y_n)^2}$$

Keterangan:

- N = Jumlah sampel data
- Y_i = Nilai aktual
- Y_n = Nilai prediksi

Semakin kecil nilai RMSE, maka semakin baik model dalam melakukan prediksi.

2. Mean Absolute Error (MAE)

MAE mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi tanpa memperhitungkan arah kesalahan (positif atau negatif). Rumus MAE adalah:

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - Y_n|$$

Keterangan:

- N = Jumlah sampel data
- Y_i = Nilai aktual
- Y_n = Nilai prediksi

Semakin kecil nilai MAE, maka semakin baik akurasi model.

3. Koefisien Determinasi (R^2 Score)

R^2 mengukur seberapa baik model dapat menjelaskan variasi dalam data aktual. Rumus R^2 adalah:

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_i - Y_n)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2}$$

keterangan:

- $\bar{Y}$ = Rata-rata nilai aktual Nilai R^2

berkisar antara 0 hingga 1:

- Jika R^2 mendekati 1, berarti model sangat baik dalam memprediksi.
- Jika R^2 mendekati 0, berarti model kurang mampu menjelaskan variabilitas data.

Dalam penelitian ini, ketiga metrik tersebut digunakan secara bersamaan agar performa model dapat dievaluasi dari berbagai aspek, yaitu besar kesalahan rata-rata (MAE), sensitivitas terhadap kesalahan besar (RMSE), serta kemampuan model menjelaskan variasi data (R^2).

2.1.7 Kelapa Sawit

Kelapa sawit (*Elaeis guineensis* Jacq) seperti yang di tunjukkan pada Gambar 2.6 adalah tanaman tropis bernilai ekonomi tinggi yang menjadi komoditas utama di Indonesia tanaman ini menghasilkan minyak kelapa sawit mentah untuk berbagai industri pangan dan non-pangan, serta minyak inti sawit (*Palm Kernel Oil*/PKO) (Rosmegawati, 2021). Dalam pemeliharaannya dibutuhkan 2 jenis unsur hara yakni unsur hara makro dan unsur hara mikro . Untuk menunjang keperluan zat hara tersebut dapat dilakukan pemupukan agar menghasilkan kelapa sawit yang berkualitas dan hasil maksimal.



Gambar 2. 6 Pohon kelapa sawit

Sumber : <https://images.app.goo.gl/6GcjkmL79ZeoRULPA>

2.1.8 Pemupukan

Berdasarkan gambar 2.7 pemupukan merupakan proses penggunaan pupuk untuk memenuhi kebutuhan zat hara agar mendukung perkembangan yang optimal, kondisi di masyarakat penggunaan pupuk tidak didasari dengan porsi standar (Ega Aprilia., 2020). Pemilik lahan melakukan perkiraan kebutuhan pupuk pada lahan milik mereka menggunakan dasar pengalaman pribadi seperti jika kondisi daun menguning maka penambahan urea 0,5 kg/pohon atau berdasarkan buah yang kecil maka membutuhkan KCL maka perlu penambahan 100-150 kg/ha/tahun. Lalu

berdasarkan kenaikan harga pupuk dan stok di pasaran akibatnya pemilik dapat mengurangi $\frac{3}{4}$ sampai $\frac{1}{2}$ dosis pupuk.



Gambar 2. 7 Pemupukan

Sumber : <https://www.elaeis.co/berita/baca>

2.1.9 Pandas

Pandas seperti yang di tunjukkan pada Gambar 2.8 adalah pustaka *Python* yang menyediakan struktur data dan alat analisis data tingkat tinggi. Kemampuan *Pandas* dalam menangani data berukuran besar dan kompleks menjadikannya alat penting dalam analisis data. Dalam penelitian ini, *Pandas* digunakan untuk melakukan proses preprocessing meliputi penggabungan data, pembersihan data, manipulasi data, dan persiapan dataset sebelum proses pelatihan model. yang dikumpulkan melalui hasil penggabungan data KUD Budi Sawit dan *Google Form* untuk pelatihan model prediksi.



Gambar 2. 8 Pandas

Sumber : https://en.m.wikipedia.org/wiki/File:Pandas_logo.svg

2.1.10 Scikit-learn

Scikit-learn pada Gambar 2.9 adalah pustaka *Python* yang menyediakan berbagai alat untuk pembelajaran mesin, termasuk algoritma untuk klasifikasi, regresi, dan klustering. *Scikit-learn* menawarkan antarmuka yang konsisten dan dokumentasi yang komprehensif, memudahkan implementasi dan evaluasi model pembelajaran mesin. Dalam penelitian ini, *Scikit-learn* digunakan untuk membangun dan mengevaluasi model prediksi pupuk berdasarkan data yang telah diproses.



Gambar 2. 9 Scikit-learn

Sumber : <https://en.wikipedia.org/wiki/Scikit-learn>

2.1.11 Flask

Flask pada Gambar 2.10 adalah kerangka kerja web mikro untuk *Python* yang memungkinkan pengembangan aplikasi web dengan cepat dan sederhana. Flask menyediakan fitur- fitur dasar seperti *routing* URL (Syafi'i dkk., 2025) Pada penelitian ini *Flask* digunakan sebagai backend yang menghubungkan model *Random Forest* yang telah disimpan menggunakan Joblib dengan antarmuka web sehingga pengguna dapat memperoleh hasil prediksi secara real-time.



Gambar 2. 10 Flask

Sumber : <https://docs.appseed.us/content/what-is/flask/>

2.1.12 Pupuk 12:6:22:3+TE

Pupuk 12:6:22:3+TE seperti gambar 2.11 merupakan pupuk majemuk yang mengandung unsur hara nitrogen (N), fosfor (P), kalium (K), dan magnesium (Mg), serta unsur mikro atau Trace Elements (TE). Angka 12:6:22:3 menunjukkan komposisi kandungan hara utama dalam pupuk tersebut, yaitu 12% nitrogen, 6% fosfor dalam bentuk P_2O_5 , 22% kalium dalam bentuk K_2O , dan 3% magnesium dalam bentuk MgO. Kandungan kalium yang relatif tinggi menjadikan pupuk ini dapat digunakan untuk memenuhi kebutuhan unsur hara tanaman kelapa sawit, terutama dalam mendukung pertumbuhan dan pembentukan hasil tanaman.

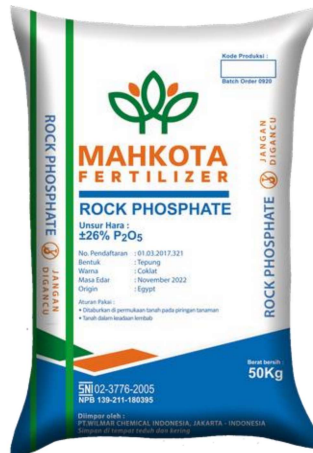


Gambar 2. 11 12:6:22:3+TE

Sumber: <https://saprotan-utama.com/product/npk-sawit-12-6-22-3te/>

2.1.13 Rock Phosphate (RP)

Rock Phosphate (RP) seperti gambar 2.12 merupakan pupuk yang berasal dari batuan fosfat dan digunakan sebagai sumber unsur hara fosfor (P). Fosfor berperan dalam berbagai proses pertumbuhan tanaman, termasuk perkembangan sistem perakaran dan metabolisme energi. Pada tanaman kelapa sawit, fosfor diperlukan untuk mendukung pertumbuhan akar dan perkembangan tanaman sehingga ketersediaannya perlu diperhatikan dalam pengelolaan kesuburan tanah.

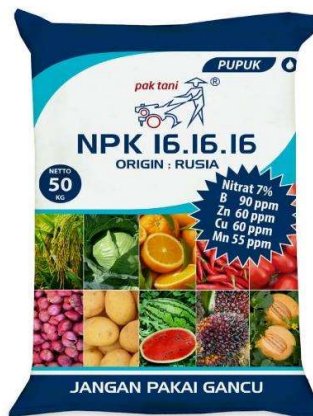


Gambar 2. 12 Rock Phosphate (RP)

Sumber: <https://pupukmahkota.co.id/produk/pupuk-tunggal-makro>

2.1.14 Pupuk NPK

Pupuk NPK seperti gambar 2.13 merupakan pupuk majemuk yang menyediakan tiga unsur hara makro utama, yaitu nitrogen (N), fosfor (P), dan kalium (K). Nitrogen berperan dalam mendukung pertumbuhan vegetatif tanaman, fosfor berperan dalam perkembangan akar dan proses metabolisme, sedangkan kalium berperan dalam berbagai proses fisiologis tanaman serta mendukung pembentukan dan pengisian hasil. Komposisi unsur hara pada pupuk NPK dapat berbeda berdasarkan formulasi produk yang digunakan.



Gambar 2. 13 NPK

Sumber: <https://saprotan-utama.com/product/npk-16-16-16-biru/>

2.1.15 Pupuk Organik

Pupuk organik seperti gambar 2.14 merupakan pupuk yang berasal dari bahan-bahan organik yang telah mengalami proses pengolahan atau dekomposisi. Penggunaan pupuk organik dapat membantu meningkatkan kandungan bahan organik tanah serta memperbaiki kondisi fisik, kimia, dan biologis tanah. Pada perkebunan kelapa sawit, pupuk organik dapat digunakan sebagai salah satu sumber bahan organik untuk membantu mempertahankan kondisi tanah yang mendukung pertumbuhan tanaman.



Gambar 2. 14 Bio-organik

Sumber: <https://encrypted-tbn0.gstatic.com/images?q=tbn>

2.1.16 Urea

Urea seperti gambar 2.15 merupakan pupuk yang digunakan sebagai sumber unsur nitrogen (N). Nitrogen merupakan salah satu unsur hara makro yang dibutuhkan tanaman dalam jumlah relatif besar dan berperan dalam pertumbuhan vegetatif serta pembentukan jaringan tanaman. Pada tanaman kelapa sawit, pemberian pupuk nitrogen perlu disesuaikan dengan kebutuhan tanaman dan kondisi lahan agar pertumbuhan tanaman dapat berlangsung secara optimal.



Gambar 2. 15 Urea

Sumber: <https://makrochemindo.com/urea-carbamide/>

2.1.17 High Grade Fertilizer Borate (HGFB)

High Grade Fertilizer Borate (HGFB) seperti gambar 2.16 merupakan pupuk yang digunakan sebagai sumber unsur boron (B). Boron termasuk unsur hara mikro yang dibutuhkan tanaman dalam jumlah relatif kecil, tetapi tetap memiliki peranan penting dalam pertumbuhan dan perkembangan tanaman. Oleh karena itu, HGFB digunakan untuk membantu memenuhi kebutuhan unsur boron pada tanaman kelapa sawit sesuai dengan kondisi dan kebutuhan tanaman.



Gambar 2. 16 High Grade Fertilizer Borate

Sumber: <https://www.goautama.com/product/pupuk-borat-mahkota-25kg>

2.1.18 Dolomit

Dolomit seperti gambar 2.17 merupakan bahan yang mengandung kalsium (Ca) dan magnesium (Mg) serta dapat digunakan untuk membantu memperbaiki kondisi tanah yang memiliki tingkat kemasaman tinggi. Pemberian dolomit dapat membantu meningkatkan pH tanah sekaligus menyediakan unsur kalsium dan magnesium bagi tanaman. Dengan demikian, penggunaan dolomit tidak hanya berfungsi sebagai sumber unsur hara, tetapi juga sebagai bahan pembenah tanah.



Gambar 2. 17 Dolomit

Sumber: <https://saraswantifertilizer.com>

2.1.19 KCL

KCL atau kalium klorida merupakan pupuk yang digunakan sebagai sumber unsur kalium (K). Kalium merupakan salah satu unsur hara makro yang dibutuhkan tanaman kelapa sawit dan berperan dalam berbagai proses fisiologis tanaman. Ketersediaan kalium yang cukup diperlukan untuk mendukung pertumbuhan serta produktivitas tanaman kelapa sawit.

2.2 Penelitian Terdahulu

Penelitian mengenai prediksi hasil panen padi yang berjudul Implementasi Algoritma *Random Forest Regression* untuk memprediksi Hasil Panen Padi di Desa Minanga. Variabel yang digunakan meliputi luas lahan, jumlah bibit, jenis pupuk, curah hujan, hama dan gulma, pengendalian hama, serta sistem tanam jajar legowo. Model dievaluasi menggunakan RMSE, MAPE, dan koefisien determinasi (R^2). Hasil penelitian menunjukkan akurasi sebesar 95,11%, nilai MAPE sebesar 4,884%, RMSE sebesar 0,250, dan nilai R^2 sebesar 0,99. Namun, penelitian ini masih memiliki keterbatasan pada proses optimasi parameter model yang belum dilakukan secara mendalam(Nur dkk., 2023).

Dalam penelitian yang lain dengan judul Analisis Produktivitas Menggunakan *Machine Learning Random Forest* di Kabupaten Batang Tahun 2018–2022 melakukan analisis produktivitas padi menggunakan data citra Sentinel-2 dan algoritma *Random Forest*. Penelitian memanfaatkan indeks vegetasi seperti NDVI, EVI, dan LSWI untuk mengidentifikasi produktivitas lahan pertanian. Hasil penelitian menunjukkan tingkat akurasi klasifikasi sebesar 91% dengan nilai Kappa sebesar 0,82. Penelitian ini berhasil menunjukkan bahwa teknologi penginderaan jauh dapat dimanfaatkan untuk mendukung pengelolaan sektor pertanian secara efektif. Akan tetapi, penelitian masih berfokus pada produktivitas tanaman padi dan belum membahas kebutuhan pupuk secara spesifik(Masdian dkk., 2023).

Penelitian selanjutnya berjudul Analisis Prediksi Nutrisi Tanaman Kelapa Sawit Menggunakan Metode *Random Forest Regression*. Penelitian ini mengembangkan sistem prediksi kandungan nutrisi tanaman kelapa sawit menggunakan citra drone multispektral dan algoritma *Random Forest Regression*. Variabel yang digunakan meliputi NDVI, *band Red, Green, Blue, Near Infrared* (NIR), serta kandungan Nitrogen (N), Fosfor (P), Kalium (K), dan Magnesium (Mg). Hasil penelitian menunjukkan nilai R^2 sebesar 0,641 untuk Nitrogen, 0,601 untuk Fosfor, 0,558 untuk Kalium, dan 0,765 untuk Magnesium. Penelitian ini memberikan kontribusi dalam pemantauan nutrisi tanaman kelapa sawit secara cepat dan efisien. Meskipun demikian, penelitian hanya berfokus pada prediksi kandungan nutrisi dan belum mampu memprediksi jumlah kebutuhan pupuk yang harus diberikan kepada tanaman(Putra dkk., 2025).

Pada penelitian berikutnya berjudul Implementasi Sistem Rekomendasi Pupuk Berbasis *Random Forest* dalam Aplikasi *E-Commerce* pertanian mengembangkan sistem rekomendasi pupuk menggunakan algoritma *Random Forest* berdasarkan jenis tanah, jenis tanaman, dan jenis pupuk. Penelitian ini bertujuan membantu petani menentukan pupuk yang sesuai untuk meningkatkan efisiensi pemupukan. Hasil penelitian menunjukkan bahwa sistem mampu memberikan rekomendasi pupuk sesuai dengan kondisi pertanian pengguna. Penelitian ini memiliki nilai praktis karena dapat langsung dimanfaatkan oleh petani dalam menentukan jenis pupuk yang tepat. Namun, penelitian hanya menghasilkan rekomendasi jenis pupuk dan belum menentukan jumlah kebutuhan pupuk yang diperlukan oleh tanaman (Marbun dkk., 2025).

Penelitian terakhir yang relevan berjudul Sistem Penjadwalan Pemupukan Kelapa Sawit Menggunakan Metode *Random Forest Regression*. Penelitian ini mengembangkan sistem penjadwalan pemupukan berbasis aplikasi *mobile* untuk membantu petani menentukan waktu pemupukan yang tepat. Variabel yang digunakan meliputi umur tanaman, pH tanah, riwayat pemupukan, frekuensi pemupukan, dan curah hujan. Hasil penelitian menunjukkan nilai MAE sebesar 4,07 hari, RMSE sebesar 14,89 hari, dan nilai R^2 sebesar 0,75. Penelitian ini berhasil membantu proses penjadwalan pemupukan secara lebih terukur dan berbasis data. Akan tetapi, penelitian masih berfokus pada prediksi waktu pemupukan dan belum membahas prediksi jumlah kebutuhan pupuk yang diperlukan tanaman kelapa sawit (Darma dkk., 2026)

Tabel 2. 1 Perbandingan variabel penelitian

Peneliti	Judul	Metode	Hasil
(Nur dkk., 2023)	Implementasi Algoritma <i>Random Forest Regression</i> untuk Memprediksi Hasil Panen Padi di Desa Minanga	<i>Random Forest Regression</i> dengan variabel luas lahan, jumlah bibit, jenis pupuk, curah hujan, hama, gulma, dan sistem tanam. Evaluasi menggunakan RMSE, MAPE, dan R^2 .	Model menghasilkan akurasi 95,11%, MAPE 4,884%, RMSE 0,250, dan R^2 sebesar 0,99. Penelitian belum menerapkan optimasi parameter model secara mendalam.
(Masdian dkk., 2023)	Analisis Produktivitas Menggunakan <i>Machine Learning Random Forest</i> di Kabupaten Batang Tahun 2018–2022	<i>Random Forest</i> menggunakan data citra Sentinel-2 dengan variabel NDVI, EVI, dan LSWI.	Penelitian memperoleh akurasi klasifikasi sebesar 91% dan nilai Kappa 0,82. Fokus penelitian masih pada produktivitas tanaman padi.
(Putra dkk., 2025)	Analisis Prediksi Nutrisi Tanaman Kelapa Sawit	<i>Random Forest Regression</i> menggunakan variabel NDVI, RGB, NIR, Nitrogen, Fosfor, Kalium, dan	Nilai R^2 sebesar 0,641 (N), 0,601 (P), 0,558 (K), dan 0,765 (Mg). Penelitian berfokus pada prediksi nutrisi, bukan jumlah kebutuhan pupuk.

Peneliti	Judul	Metode	Hasil
	Menggunakan Metode <i>Random Forest Regression</i>	Magnesium.	
(Marbun dkk., 2025)	Implementasi Sistem Rekomendasi Pupuk Berbasis Random <i>Forest</i> dalam Aplikasi <i>E-Commerce</i> Pertanian	<i>Random Forest</i> menggunakan variabel jenis tanah, jenis tanaman, dan jenis pupuk.	Sistem mampu memberikan rekomendasi jenis pupuk sesuai kondisi lahan, namun belum memprediksi jumlah kebutuhan pupuk.
(Darma dkk., 2026)	Sistem Penjadwalan Pemupukan Kelapa Sawit Menggunakan Metode <i>Random Forest Regression</i>	<i>Random Forest Regression</i> menggunakan variabel umur tanaman, pH tanah, riwayat pemupukan, frekuensi pemupukan, dan curah hujan. Evaluasi menggunakan MAE, RMSE, dan R^2 .	Model menghasilkan MAE sebesar 4,07 hari, RMSE sebesar 14,89 hari, dan R^2 sebesar 0,75. Penelitian berfokus pada penjadwalan waktu pemupukan.

Peneliti	Judul	Metode	Hasil
(Sudaris, 2026)	Prediksi Kebutuhan Pupuk pada Lahan Kelapa Sawit Menggunakan <i>Algoritma Random Forest</i>	<i>Random Forest</i> menggunakan variabel luas lahan, jumlah pohon, usia tanaman, jenis pupuk, jenis tanah, dan kondisi daun. Evaluasi menggunakan MAE, RMSE, dan R^2 <i>Score</i> , serta model diimplementasikan pada aplikasi berbasis web menggunakan Flask.	Model baseline menghasilkan MAE sebesar 65,06 kg, RMSE sebesar 213,10 kg, dan R^2 sebesar 0,865. Sementara model final hasil optimasi memperoleh MAE sebesar 66,17 kg, RMSE sebesar 213,02 kg, dan R^2 sebesar 0,865. Hasil <i>feature importance</i> menunjukkan bahwa jumlah pohon, luas lahan, dan jenis pupuk merupakan variabel dengan kontribusi terbesar terhadap prediksi kebutuhan pupuk.

Berdasarkan kajian terhadap beberapa penelitian terdahulu pada Tabel 2.1, algoritma *Random Forest* telah diterapkan dalam berbagai permasalahan di bidang pertanian, seperti prediksi hasil panen, analisis produktivitas tanaman, prediksi kandungan nutrisi, rekomendasi jenis pupuk, dan penjadwalan pemupukan. Hasil penelitian terdahulu menunjukkan bahwa *Random Forest* mampu digunakan pada data dengan karakteristik dan hubungan antarvariabel yang kompleks. Namun, penelitian yang secara khusus membahas prediksi jumlah kebutuhan pupuk pada lahan kelapa sawit masih terbatas. Oleh karena itu, penelitian ini berfokus pada pengembangan model prediksi kebutuhan pupuk berdasarkan karakteristik lahan dan tanaman.