

Sistem Rekomendasi Pupuk Berdasarkan Kondisi dan Jenis Tanah Menggunakan *LightGBM*

Aryanti Aryanti^{*1}, Aknes Tasia Pratama², Dava Anugrah Limanda³, Hisanah Nakhwah Aulia Faruly⁴, Muhammad Juan Farza Rafly Alganiyu⁵,

^{1,2,3,4,5}Jurusan Teknik Elektro, Politeknik Negeri Sriwijaya, Palembang

e-mail: ^{*}¹aryanti@polsri.ac.id, ²aknestasia2279@gmail.com,

³davaanugrahlmanda@gmail.com, ⁴hisanah.nakhwah@gmail.com, ⁵juanfarza28@gmail.com

Abstrak

Pemupukan yang tidak tepat masih menjadi permasalahan dalam sektor pertanian karena dapat menurunkan produktivitas tanaman, meningkatkan biaya produksi, serta menyebabkan penggunaan pupuk yang kurang efisien. Oleh karena itu, diperlukan sistem rekomendasi pupuk yang dapat membantu petani menentukan jenis pupuk yang sesuai berdasarkan kondisi tanah. Penelitian ini mengembangkan sistem rekomendasi pupuk berbasis machine learning menggunakan algoritma Light Gradient Boosting Machine (LightGBM) untuk mendukung pertanian presisi. Dataset yang digunakan terdiri dari 10.000 data dengan delapan parameter input, yaitu jenis tanah (clay, silt, sandy, loamy), kelembaban tanah, pH tanah, kandungan karbon organik, serta kadar nitrogen, fosfor, dan kalium. Target prediksi terdiri dari tujuh kelas pupuk. Tahapan penelitian meliputi preprocessing (data cleaning, label encoding, dan normalisasi min-max), pembagian dataset dengan rasio 80:20, pelatihan model LightGBM dengan konfigurasi $n_estimators=500$ dan $learning_rate=0,1$, serta evaluasi menggunakan confusion matrix dan classification report. Hasil eksperimen menunjukkan akurasi sebesar 86,35% pada data pengujian. Analisis feature importance menunjukkan bahwa Soil Moisture merupakan fitur paling berpengaruh, diikuti Soil pH, Organic Carbon, dan Nitrogen Level. Model menunjukkan performa sangat baik pada kelas Urea dan DAP dengan f1-score mencapai 0,95. Hasil penelitian membuktikan bahwa LightGBM efektif untuk klasifikasi multi-kelas dalam sistem rekomendasi pupuk dan berpotensi mendukung pertanian presisi yang lebih efisien dan berkelanjutan.

Kata kunci—LightGBM, Rekomendasi, Pupuk, Tanah, Pertanian

1. PENDAHULUAN

Tanah merupakan media tumbuh utama sekaligus penyedia unsur hara esensial yang menentukan keberhasilan produksi pertanian [1]. Kualitas tanah yang optimal sangat bergantung pada ketersediaan unsur hara makro utama, yaitu Nitrogen (N) untuk pertumbuhan vegetatif dan pembentukan klorofil, Fosfor (P) untuk perkembangan akar dan bunga [2], serta Kalium (K) yang mendukung proses fotosintesis [3]. Namun, praktik pertanian intensif yang tidak diimbangi dengan manajemen pemupukan yang tepat sering kali menyebabkan penurunan kesuburan tanah dan ketidakseimbangan nutrisi. Penggunaan pupuk yang tidak berdasarkan kondisi spesifik lahan, seperti pemberian dosis berlebihan, tidak hanya menurunkan efisiensi biaya tetapi juga memicu kerusakan lingkungan yang serius, termasuk pencemaran air dan tanah akibat pencucian nitrogen (leaching) [4]. Untuk mengatasi permasalahan tersebut, berbagai penelitian telah mengembangkan metode untuk mengoptimalkan manajemen nutrisi tanaman. Salah satu pendekatan yang banyak diterapkan adalah Site-Specific Nutrient Management (SSNM), yaitu metode pemupukan yang disesuaikan dengan karakteristik spesifik tanah di lokasi tertentu untuk meningkatkan efisiensi dan produktivitas [5]. Seiring perkembangan teknologi, penerapan

machine learning mulai dimanfaatkan dalam sistem rekomendasi pertanian untuk membantu pengambilan keputusan yang lebih cepat dan akurat berdasarkan kondisi lingkungan dan tanah [6]. Berbagai pendekatan telah diterapkan untuk menjaga produktivitas dan kesuburan tanah secara berkelanjutan. Teknologi smart fertilizer dilaporkan mampu meningkatkan hasil dan kualitas tanaman [7], sedangkan kombinasi pupuk organik dan anorganik dapat meningkatkan keberlanjutan produksi pertanian dalam jangka panjang [8]. Selain itu, praktik pengelolaan lahan pada sistem agroforestri berkontribusi terhadap pemeliharaan kesuburan tanah [9], serta didukung oleh konsep smart agroforestry yang mengintegrasikan keberlanjutan lingkungan dan ekonomi masyarakat.

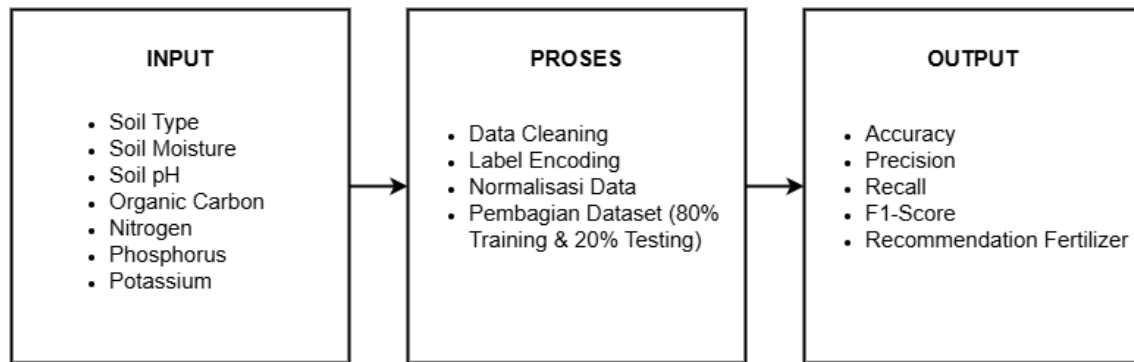
Penerapan machine learning pada sistem rekomendasi pupuk dan klasifikasi kondisi tanah telah banyak dilakukan menggunakan berbagai algoritma. Aryanti dan Iryani [6] memanfaatkan algoritma Decision Tree untuk mendukung rekomendasi pada sektor pertanian. Algoritma Decision Tree memiliki keunggulan dalam kemudahan interpretasi, namun rentan mengalami overfitting ketika dihadapkan pada data yang kompleks. Penelitian lain menggunakan Random Forest dan Support Vector Machine (SVM) untuk mengolah data pertanian multivariat dan meningkatkan akurasi klasifikasi [10], [11]. Meskipun mampu memberikan performa yang baik, Random Forest cenderung menghasilkan model yang lebih kompleks dan membutuhkan sumber daya komputasi yang lebih besar, sedangkan SVM sensitif terhadap pemilihan parameter dan kurang efisien pada dataset berukuran besar. Selain itu, algoritma boosting seperti XGBoost telah menunjukkan kemampuan yang baik dalam memodelkan hubungan non-linear yang kompleks, namun memerlukan proses tuning parameter yang relatif lebih rumit dan biaya komputasi yang lebih tinggi [11].

Meskipun berbagai algoritma machine learning telah diterapkan pada bidang pertanian, masih terdapat beberapa keterbatasan yang perlu diperhatikan. Sebagian penelitian lebih berfokus pada rekomendasi tanaman atau benih, sementara penelitian yang secara khusus membahas rekomendasi pupuk berdasarkan kombinasi parameter jenis tanah, kelembaban, pH, karbon organik, serta kandungan Nitrogen, Fosfor, dan Kalium masih relatif terbatas [6], [12]. Selain itu, sebagian besar penelitian hanya mengevaluasi satu algoritma sehingga belum memberikan gambaran yang komprehensif mengenai efektivitas berbagai metode machine learning dalam menghasilkan rekomendasi pupuk. Oleh karena itu, diperlukan penelitian yang tidak hanya membangun sistem rekomendasi pupuk berbasis machine learning, tetapi juga membandingkan performa beberapa algoritma untuk memperoleh model yang lebih akurat dan efisien.

Berdasarkan permasalahan dan gap penelitian tersebut, penelitian ini mengusulkan sistem rekomendasi pupuk berdasarkan kondisi dan jenis tanah menggunakan algoritma Light Gradient Boosting Machine (LightGBM) dan AdaBoost. LightGBM dipilih karena memiliki keunggulan dalam kecepatan pelatihan, efisiensi penggunaan memori, serta kemampuan menangani dataset berskala besar dengan akurasi yang tinggi [13]. Sementara itu, AdaBoost digunakan sebagai algoritma pembanding karena merupakan salah satu metode boosting yang banyak digunakan dalam permasalahan klasifikasi dan memiliki performa yang kompetitif pada berbagai penelitian machine learning [14], [11]. Sistem yang dikembangkan memanfaatkan parameter jenis tanah, kelembaban tanah, pH tanah, kandungan karbon organik, serta kadar Nitrogen, Fosfor, dan Kalium untuk menghasilkan rekomendasi jenis pupuk yang sesuai. Melalui perbandingan kedua algoritma tersebut, penelitian ini bertujuan untuk memperoleh model terbaik dalam menghasilkan rekomendasi pupuk yang akurat sehingga dapat mendukung penerapan pertanian presisi yang lebih efisien dan berkelanjutan [15].

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan eksperimen berbasis pembelajaran mesin. Tujuan penelitian ini adalah untuk membangun sistem rekomendasi pupuk yang menggunakan algoritma LightGBM berdasarkan parameter kondisi dan jenis tanah. Tahapan penelitian termasuk identifikasi masalah, pengumpulan data, preprocessing, pembagian data, pelatihan model, dan evaluasi model. Tahapan Penelitian dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data (INPUT)

Pada titik ini, data dikumpulkan dari kumpulan data pertanian yang sudah ada sebelumnya, yang menunjukkan hubungan antara kondisi tanah dan jenis pupuk [16]. Parameter yang digunakan dalam studi ini termasuk jenis tanah (*clay, silt, sandy, loamy*), kelembaban tanah, pH tanah, kandungan karbon organik, serta unsur hara seperti nitrogen, fosfor, dan kalium. Label yang digunakan adalah jenis pupuk seperti NPK, Urea, Kompos, MOP/KCL, SSP, DAP dan Zinc Sulphate [17]. Dataset bersumber dari platform Kaggle dengan total 10.000 data.

Tabel 1. Atribut dan Deskripsi Dataset Fertilizer Recommendation

Atribut	Deskripsi	Tipe Data	Satuan / Contoh
Soil_Type	Jenis tanah (clay, silt, sandy, loamy)	Kategorikal	Clay, Sandy, dll.
Soil_Moisture	Kelembaban tanah	Numerik	%
Soil_pH	Tingkat keasaman tanah	Numerik	0 - 14
Organic_Carbon	Kandungan karbon organik tanah	Numerik	%
Nitrogen (N)	Kadar unsur nitrogen dalam tanah	Numerik	mg/kg
Phosphorus (P)	Kadar unsur fosfor dalam tanah	Numerik	mg/kg
Potassium (K)	Kadar unsur kalium dalam tanah	Numerik	mg/kg
Recommended_Fertilizer	Label jenis pupuk (target/output)	Kategorikal	NPK, Urea, dll.

Tahap input merupakan data masukan yang digunakan sebagai parameter dalam sistem rekomendasi pupuk. Terdapat tujuh parameter kondisi tanah yang menjadi input, yaitu:

- *Soil Type* (Jenis Tanah): merupakan kategori jenis tanah yang terdiri atas clay (tanah liat), silt (tanah lanau), sandy (tanah berpasir), dan loamy (tanah lempung). Jenis tanah menentukan kapasitas retensi air dan ketersediaan unsur hara bagi tanaman.
- *Soil Moisture* (Kelembaban Tanah): merupakan kandungan air di dalam tanah yang dinyatakan dalam satuan persen (%). Kelembaban tanah berpengaruh langsung terhadap kelarutan dan penyerapan pupuk oleh akar tanaman.
- *Soil pH* (Keasaman Tanah): merupakan tingkat keasaman tanah yang diukur pada skala 0 hingga 14. Nilai pH tanah sangat memengaruhi ketersediaan unsur hara dan aktivitas mikroorganisme di dalam tanah.
- *Organic Carbon* (Karbon Organik): merupakan kandungan karbon organik dalam tanah yang dinyatakan dalam persen (%). Karbon organik mencerminkan tingkat kesuburan tanah dan kemampuan tanah dalam menyimpan unsur hara.
- *Nitrogen* (N): merupakan kadar unsur nitrogen dalam tanah yang diukur dalam satuan mg/kg. Nitrogen berperan penting dalam pertumbuhan vegetatif tanaman dan pembentukan klorofil.
- *Phosphorus* (P): merupakan kadar unsur fosfor dalam tanah yang diukur dalam satuan mg/kg. Fosfor berperan dalam perkembangan akar, pembungaan, serta pembentukan biji tanaman.

2.2 Pra-Proses

Pra-proses data adalah salah satu tahap paling penting dalam pipeline pembelajaran mesin. Performa model akan sangat dipengaruhi oleh kualitas data yang baik. Praproses penelitian ini terdiri dari tiga tahap utama.

2.2.1 Pembersihan Data (Data Cleaning)

Pada tahap ini dilakukan identifikasi dan penanganan data yang tidak valid, meliputi: (a) penghapusan nilai kosong (*missing value*) yang dapat menyebabkan bias pada model; (b) penghapusan data duplikat yang dapat mempengaruhi distribusi kelas; (c) penanganan nilai *outlier* menggunakan metode Interquartile Range (IQR).

2.2.2 Transformasi Data (Label Encoding)

Algoritma LightGBM memerlukan input dalam bentuk numerik. Oleh karena itu, fitur kategorikal dikonversi menggunakan metode *Label Encoding*. Tabel 2 berikut menunjukkan contoh transformasi yang dilakukan.

Tabel 2. Hasil Label Encoding pada Fitur Kategorikal

Fitur	Nilai Kategori	Nilai Numerik (Encoded)
Soil_Type	Clay	0
Soil_Type	Loamy	1
Soil_Type	Sandy	2
Soil_Type	Silt	3
Fertilizer	DAP	0
Fertilizer	Kompos	1
Fertilizer	MOP/KCL	2
Fertilizer	NPK	3
Fertilizer	SSP	4
Fertilizer	Urea	5

2.2.3 Normalisasi Data

Normalisasi dilakukan menggunakan metode Min-Max Scaling untuk menyeragamkan rentang nilai fitur numerik (seperti pH, Nitrogen, Phosphorus, dan Potassium) ke dalam interval [0, 1]. Normalisasi ini memastikan tidak ada fitur yang mendominasi proses pembelajaran akibat perbedaan skala yang besar.

2.2.4 Pembagian Dataset

Untuk memastikan bahwa model yang dibangun mampu melakukan generalisasi yang efektif terhadap data baru, dataset dalam penelitian ini dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian, yang merupakan tahapan penting dalam proses pengembangan model pembelajaran mesin.

Delapan puluh persen data pelatihan digunakan untuk mengajarkan model bagaimana pola hubungan antara parameter kondisi tanah dan jenis pupuk yang disarankan. Sementara itu, dua puluh persen data pengujian digunakan untuk menguji kinerja model terhadap data yang belum pernah dilihat sebelumnya.

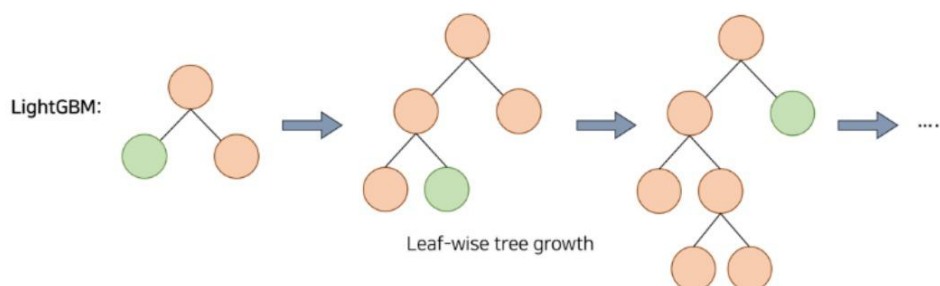
Rasio 80:20 didasarkan pada keseimbangan antara data yang cukup untuk pelatihan model dan data yang representatif untuk pengujian, sehingga hasil evaluasi dapat lebih akurat menunjukkan kinerja model [18].

Tabel 3. Pembagian Dataset Training dan Testing

Subset Data	Proporsi	Fungsi
Data Training	80%	Melatih model LightGBM
Data Testing	20%	Menguji performa model

2.4 Algoritma LightGBM

Algoritma peningkatan gradient adalah Light Gradient Boosting Machine (LightGBM) yang dikembangkan oleh Microsoft Research [19]. Berdasarkan Gambar 2, LightGBM membangun pohon keputusan menggunakan pendekatan *leaf-wise tree growth*, yaitu memperluas daun dengan nilai gain terbesar pada setiap iterasi sehingga mampu meningkatkan akurasi model secara efisien. LightGBM menggunakan dua teknik utama yang membedakannya dari algoritma boosting lainnya, keduanya, One-Side Gradient-based Sampling (GOSS) dan Exclusive Feature Bundling (EFB). Teknik-teknik ini meningkatkan kecepatan dan efisiensi LightGBM dalam memproses dataset yang sangat besar tanpa mengurangi akurasi [20].



Gambar 2. Leaf-wise Tree Growth pada LightGBM

Keunggulan LightGBM yang menjadi alasan pemilihan algoritma ini dalam penelitian antara lain: (a) kecepatan pelatihan yang tinggi dibandingkan algoritma gradient boosting konvensional seperti XGBoost; (b) penggunaan memori yang lebih efisien berkat teknik histogram-based splitting; (c) kemampuan menangani fitur kategorikal secara langsung; (d) performa tinggi pada dataset multi-kelas seperti prediksi jenis pupuk; dan (e) kemampuan menangani ketidakseimbangan kelas (class imbalance) dengan baik.

2.5 Output

Output diperoleh dari evaluasi model yang dilakukan. Hal ini dilakukan untuk mengetahui seberapa baik model LightGBM dapat mengklasifikasikan jenis pupuk dengan benar berdasarkan parameter tanah yang diberikan. Accuracy, Precision, Recall, F1-Score, dan *Confusion matrix* adalah metrik evaluasi yang digunakan [21].

Tabel 4. Metrik Evaluasi Model yang Digunakan

Metrik	Formula	Keterangan
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$	Proporsi prediksi benar dari total prediksi
Precision	$TP/(TP+FP)$	Ketepatan prediksi positif model
Recall	$TP/(TP+FN)$	Kemampuan model mendeteksi kelas positif
F1-Score	$2 \times (P \times R) / (P + R)$	Harmonisasi antara Precision dan Recall
<i>Confusion matrix</i>	-	Matriks visualisasi kesalahan klasifikasi per kelas

Model dinyatakan baik dan siap digunakan apabila nilai akurasi mencapai minimal 85% pada data testing. Apabila belum memenuhi ambang batas tersebut, dilakukan tuning hyperparameter secara iteratif hingga performa yang diinginkan tercapai.

2.6 Lingkungan Implementasi

Model digunakan menggunakan Python pada platform Google Colaboratory (Colab). Tabel 5 berikut merangkum spesifikasi lingkungan implementasi yang digunakan.

Tabel 5. Spesifikasi Lingkungan Implementasi Penelitian

Komponen	Spesifikasi
Platform	Google Colaboratory (Colab)
Bahasa Pemrograman	Python 3.x
Library Utama	LightGBM, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn
Sumber Dataset	Kaggle - Fertilizer Recommendation Dataset
Metode Validasi	Hold-Out (80% Training, 20% Testing)
Hardware (Colab)	CPU/GPU Runtime Google Cloud

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pelatihan Model

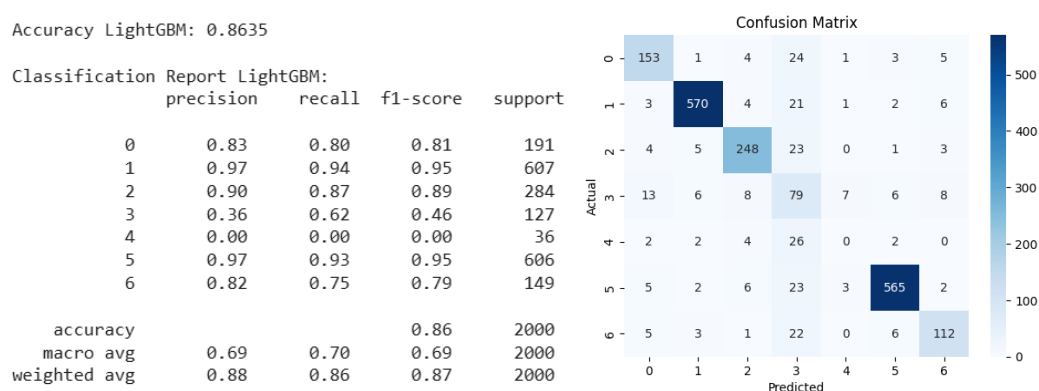
Penelitian ini mengimplementasikan algoritma Light Gradient Boosting Machine (LightGBM) sebagai model utama dalam sistem rekomendasi pemupukan berdasarkan parameter kondisi tanah. Variabel yang digunakan meliputi *Soil Type*, *Soil Moisture*, *Soil pH*, *Organic Carbon*, *Nitrogen Level*, *Phosphorus Level*, dan *Potassium Level* sebagai fitur input, serta *Recommended Fertilizer* sebagai variabel target.

Tahap awal yang dilakukan adalah *preprocessing* data, yang mencakup pemilihan fitur yang relevan serta transformasi data kategorikal menggunakan metode *Label Encoding*. Selanjutnya, dataset dibagi menjadi data latih dan data uji dengan perbandingan 80:20 menggunakan metode *train-test split* dengan nilai *random state* sebesar 42. Model LightGBM dikonfigurasi dengan parameter *n_estimators* sebesar 500, *learning_rate* sebesar 0,1, serta *random_state* sebesar 42.

Berdasarkan hasil pengujian, model LightGBM memperoleh nilai akurasi sebesar 86,35%. Nilai tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan jenis pupuk berdasarkan kondisi tanah, sekaligus melampaui ambang batas minimum akurasi yang ditetapkan sebesar 85%.

3.2 Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan *confusion matrix* dan *classification report* yang mencakup metrik *precision*, *recall*, dan *f1-score* seperti pada Gambar 3.

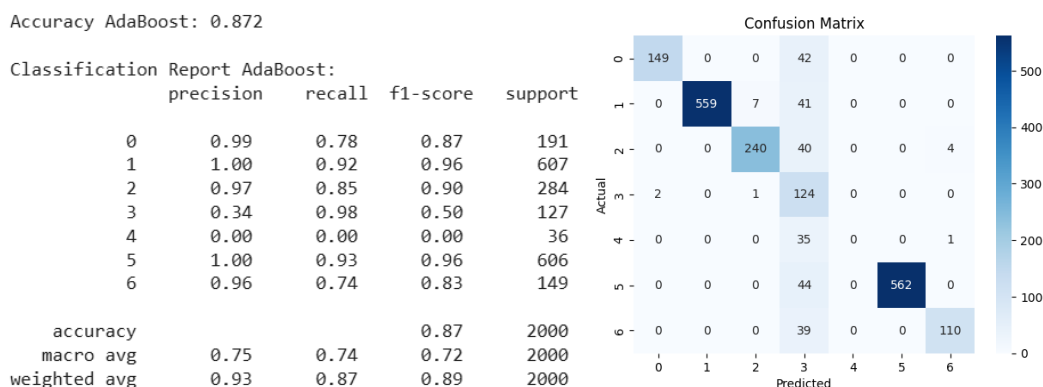
Gambar 3. Clasification Report dan *Confusion matrix* Model LightGBM

Berdasarkan *classification report* pada gambar 3, sebagian besar kelas memiliki nilai *precision* dan *recall* yang tinggi, terutama pada kelas 1 (Urea) dan kelas 5 (DAP) dengan *f1-score* mencapai 0,95. Kelas 2, 0, dan 6 juga menunjukkan performa yang cukup baik dengan *f1-score* masing-masing sebesar 0,89, 0,81, dan 0,79. Namun, performa model menurun pada kelas 3 dengan *f1-score* sebesar 0,46, serta kelas 4 yang memiliki nilai *precision*, *recall*, dan *f1-score* sebesar 0,00, yang menunjukkan bahwa model tidak mampu mengenali kelas tersebut.

Hasil ini sejalan dengan *confusion matrix*, di mana prediksi benar dominan berada pada diagonal utama, khususnya pada kelas 1 dan kelas 5. Sebaliknya, pada kelas 3 dan 4 terlihat

adanya kesalahan klasifikasi yang cukup signifikan. Nilai *macro average f1-score* sebesar 0,69 lebih rendah dibandingkan *weighted average f1-score* sebesar 0,87, yang mengindikasikan adanya ketidakseimbangan distribusi data antar kelas.

Sebagai pembandingan, evaluasi juga dilakukan terhadap model *AdaBoost* dengan hasil yang ditampilkan pada Gambar 4.



Gambar 4. Clasification Report dan *Confusion matrix* Model AdaBoost

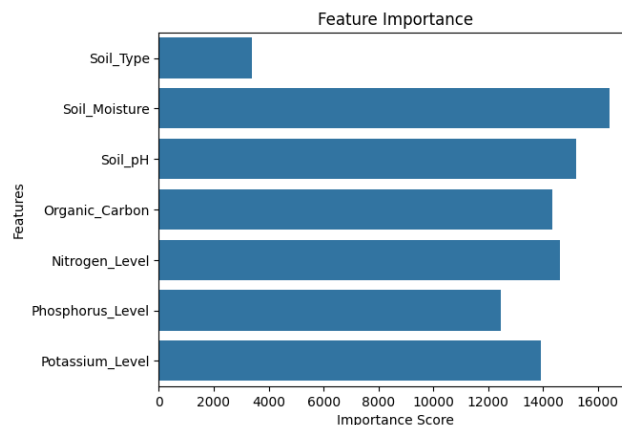
Berdasarkan *classification report* AdaBoost pada gambar 4, model mencapai *accuracy* sebesar 0,87, dengan performa sangat baik pada kelas 1 (Urea) dan kelas 5 (DAP) yang masing-masing memiliki *f1-score* sebesar 0,96. Kelas 2 juga menunjukkan performa yang baik dengan *f1-score* sebesar 0,90, sementara kelas 0 dan kelas 6 memiliki *f1-score* masing-masing 0,87 dan 0,83. Namun, serupa dengan LightGBM, model AdaBoost juga mengalami kesulitan dalam mengenali kelas 3 dengan *f1-score* hanya 0,50, serta kelas 4 yang memiliki nilai *precision*, *recall*, dan *f1-score* sebesar 0,00, menunjukkan model tidak mampu mengenali kelas tersebut sama sekali.

Hasil ini sejalan dengan *confusion matrix* AdaBoost, di mana prediksi benar terkonsentrasi pada diagonal utama, khususnya pada kelas 1 (559) dan kelas 5 (562). Namun, terlihat kesalahan klasifikasi yang cukup dominan pada kolom kelas 3, di mana banyak sampel dari kelas lain (kelas 0, 1, 2, 5, dan 6) salah diprediksi sebagai kelas 3. Nilai *macro average f1-score* sebesar 0,72 lebih rendah dibandingkan *weighted average f1-score* sebesar 0,89, yang mengindikasikan adanya ketidakseimbangan distribusi data antar kelas, konsisten dengan temuan pada model LightGBM.

Secara keseluruhan, LightGBM menunjukkan performa yang lebih seimbang dibandingkan AdaBoost, terutama dilihat dari nilai *macro average f1-score* LightGBM (0,69) yang relatif lebih baik dalam menangani distribusi kelas dibandingkan AdaBoost (0,72), meskipun AdaBoost sedikit unggul dalam *accuracy* (0,87 vs 0,86).

3.3 Analisis Feature Importance

Analisis *feature importance* dilakukan untuk mengetahui kontribusi masing-masing fitur dalam proses prediksi model LightGBM seperti pada Gambar 5. Berdasarkan grafik *feature importance* pada Gambar 5, fitur yang paling berpengaruh dalam prediksi rekomendasi pupuk adalah *Soil Moisture* dengan skor importance tertinggi (sekitar 16.000 poin), diikuti oleh *Soil pH*, *Organic Carbon*, dan *Nitrogen Level*. Sementara itu, *Soil Type* memiliki kontribusi yang paling rendah dibandingkan fitur lainnya. Hasil ini menunjukkan bahwa kondisi fisik tanah, khususnya kelembaban dan keasaman, merupakan faktor penentu utama dalam pemilihan jenis pupuk yang tepat.



Gambar 5. Grafik Feature Importance Model LightGBM

3.4 Perbandingan Prediksi dan Nilai Aktual

Tabel 6 berikut menampilkan perbandingan antara nilai prediksi model dengan nilai aktual pada 10 sampel data uji, yang memperlihatkan kemampuan generalisasi model LightGBM terhadap data baru [22].

Tabel 6. Perbandingan Hasil Rekomendasi dan Nilai Aktual (10 Sampel)

No.	Soil_Moisture	Soil_pH	Nitrogen	Phosphorus	Actual	Predicted
1	16.36	8.09	136	51	Zinc Sulphate	Zinc Sulphate
2	14.80	6.72	152	12	DAP	DAP
3	12.63	5.39	155	13	DAP	DAP
4	36.83	7.91	29	86	Urea	Urea
5	44.27	6.94	26	76	Urea	Urea
6	19.55	6.09	75	77	MOP	MOP
7	33.17	7.10	121	25	DAP	DAP
8	34.38	5.72	70	17	DAP	DAP
9	24.51	5.99	129	51	MOP	MOP
10	11.15	7.88	84	70	MOP	MOP

Berdasarkan Tabel 6, seluruh 10 sampel yang ditampilkan berhasil diprediksi dengan benar oleh model LightGBM. Hal ini menguatkan temuan bahwa model memiliki kemampuan prediksi yang baik terutama pada kelas-kelas dengan distribusi data yang seimbang seperti DAP, Urea, dan MOP.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem rekomendasi pupuk berbasis *machine learning* menggunakan algoritma LightGBM dengan memanfaatkan dataset yang mencakup parameter kondisi tanah secara komprehensif. Berdasarkan hasil eksperimen dan evaluasi yang telah dilakukan, model LightGBM mencapai akurasi sebesar 86,35% pada data pengujian, melampaui ambang batas minimum 85% yang ditetapkan dalam penelitian ini, sehingga model dinilai layak diimplementasikan sebagai sistem pendukung keputusan pemupukan. Analisis *feature importance* menunjukkan bahwa *Soil Moisture* merupakan fitur yang paling berpengaruh dalam prediksi rekomendasi pupuk, diikuti oleh *Soil pH*, *Organic Carbon*, dan *Nitrogen Level*, yang mengindikasikan bahwa kondisi fisik dan kimia tanah merupakan faktor utama dalam penentuan jenis pupuk. Selain itu, model menunjukkan performa yang sangat baik pada kelas Urea dan DAP dengan *f1-score* sebesar 0,95. Namun, performa pada kelas dengan jumlah data yang lebih sedikit, seperti Zinc Sulphate dan Compost, masih relatif rendah, sehingga diperlukan penanganan *class imbalance* pada penelitian selanjutnya untuk meningkatkan kinerja model pada kelas minoritas.

5. SARAN

Berdasarkan hasil penelitian, penelitian selanjutnya disarankan untuk menerapkan teknik penanganan *class imbalance*, seperti *oversampling* atau *undersampling*, agar performa model pada kelas minoritas dapat meningkat. Selain itu, pengujian menggunakan algoritma *machine learning* lain, khususnya metode *ensemble learning*, perlu dilakukan untuk memperoleh model dengan kinerja yang lebih optimal. Pengembangan lebih lanjut juga dapat dilakukan dengan mengintegrasikan sistem rekomendasi pupuk dengan perangkat IoT berbasis sensor tanah sehingga rekomendasi dapat diberikan secara *real-time* dan lebih aplikatif.

DAFTAR PUSTAKA

- [1] M. Andronikus and M. U. Nganji, "Kajian Pengelolaan Lahan Pertanian di Desa Banggawatu, Kecamatan Tabundung, Kabupaten Sumba Timur," vol. 03, no. 2, pp. 109–118, 2025.
- [2] R. A. Ma'rifat and I. M. Suraharta, "Pengaruh Fosfor terhadap Pertumbuhan Akar dan Bunga Tanaman," vol. 2, pp. 306–312, 2024.
- [3] J. Rawat, N. Pandey, and J. Saxena, "Role of Potassium in Plant Photosynthesis, Transport, Growth and Yield," *Role Potassium Abiotic Stress*, pp. 1–14, 2022.
- [4] Y. Xing, Y. Xie, and X. Wang, "Enhancing soil health through balanced fertilization: a pathway to sustainable agriculture and food security," *Front. Microbiol.*, vol. 16, no. April, pp. 1–21, 2025.
- [5] H. H. Sarma et al., "Site Specific Nutrient Management (SSNM): Principles, Key Features and Its Potential Role in Soil, Crop Ecosystem and Climate Resilience Farming," *J. Adv. Biol. Biotechnol.*, vol. 27, no. 8, pp. 211–222, 2024.
- [6] A. Aryanti and N. Iryani, "A Smart Recommendation System for Crop Seed Selection Using Gradient Boosting Based on Environmental and Geospatial Data," vol. 9, no. 6, pp. 2965–2973, 2025.
- [7] Sutardi et al., "Smart Fertilizer: Ameliorant-Based NPK Coating Enhanced Rice Yield and Quality in Vertisol Paddy Soil," *J. Saudi Soc. Agric. Sci.*, vol. 24, no. 4, 2025.
- [8] J. Wang et al., "Long-Term Combined Application of Organic and Inorganic Fertilizers Increases Crop Yield Sustainability," *J. Integr. Agric.*, vol. 24, no. 1, pp. 290–305, 2025.
- [9] J. Chavez et al., "Impact of Farm Management on Soil Fertility in Agroforestry Systems in Bali, Indonesia," *Sustainability*, vol. 16, no. 18, 2024.
- [10] F. Teknik and U. Mulawarman, "Algoritma Swarm Intelligence untuk Data Berdimensi Tinggi pada Machine Learning: Review," 2024.
- [11] R. Maulidia, W. A. Arifin, and H. Syafr, "Perbandingan Algoritma Machine Learning untuk Prediksi," 2025.
- [12] B. Dey, J. Ferdous, and R. Ahmed, "Machine Learning Based Recommendation of Agricultural and Horticultural Crop Farming in India Under the Regime of NPK, Soil pH and Three Climatic Variables," *Heliyon*, 2024.

- [13] M. Hasan et al., "Ensemble Machine Learning-Based Recommendation System for Effective Prediction of Suitable Agricultural Crop Cultivation," *Front. Plant Sci.*, vol. 14, no. August, pp. 1–18, 2023.
- [14] K. A. Reddy et al., "Designing an Efficient Fertilizer Recommendation System with Bayesian Light GBM Processing and Mobile Net-Based Image Analysis," 2025.
- [15] V. Ralte et al., "Machine Learning-Based Fertilizer Recommendation System for Sustainable Crop Production," *Int. Sci. J. Eng. Manag.*, 2025.
- [16] V. Assistant Professor et al., "Machine Learning-Based Fertilizer Recommendation System for Sustainable Crop Production," 2025.
- [17] J. Deone, D. Khan, R. Afreen, and R. Mangrule, "Machine Learning Approaches for Crop Prediction and Fertilizer Recommendation Based on Soil Nutrientse," 2024.
- [18] T. Muthukrishnan, V. Pandi, and K. Thiruvankadam, "IoT-Enabled Smart Fertilizer Recommendation System Using Machine Learning," 2025.
- [19] A. S. Alfath and A. K. Wardhana, "Hypertension Risk Prediction Using Stacking Ensemble of CatBoost, XGBoost, and LightGBM," 2025.
- [20] D. Raihan and I. Lubis, "Optimalisasi Model Peramalan Penjualan dengan Pendekatan Algoritma Light Gradient Boosting Machine (LightGBM)," 2025.
- [21] R. M. Sumacot and J. R. Paragas, "IoT-Enabled K-Nearest Neighbors (KNN)-Based Soil Nutrient Recommendation System for Rice Cultivation," 2025.
- [22] G. Mamatha and J. S. Nayak, "A Novel Design and Implementation of Fertilizer Recommendation System Based on Hybrid Machine Learning Models Journal: SSRG International Journal of Electrical and Electronics," *SSRG Int. J. Electr. Electron. Eng.*, 2024.