

# Studi Klasifikasi Presisi Kesesuaian Lahan Pertanian Menggunakan Algoritma Extreme Gradient Boosting (XGBoost)

Bernadus Gunawan Sudarsono <sup>1,\*</sup>, Allan Desi Alexander <sup>1</sup>

<sup>1</sup>Informatika; Fakultas Ilmu Komputer; Universitas Bhayangkara Jakarta Raya; Jl. Perjuangan No.81, Marga Mulya, Bekasi Utara, Bekasi, Jawa Barat, 17143, Telp.(021) 88955882; e-mail: [bernadus.gs@dsn.ubharajaya.ac.id](mailto:bernadus.gs@dsn.ubharajaya.ac.id), [allan@ubharajaya.ac.id](mailto:allan@ubharajaya.ac.id)

\* Korespondensi: [bernadus.gs@dsn.ubharajaya.ac.id](mailto:bernadus.gs@dsn.ubharajaya.ac.id)

Diterima: 20 Juli 2026; Review: 28 Juli 2026; Disetujui: 30 Juli 2026; Diterbitkan: 30 Juli 2026

---

## Abstract

*The application of artificial intelligence technology in the agricultural sector is the main foundation in the paradigm shift towards sustainable precision agriculture. This study presents a comprehensive analysis of the application of the Extreme Gradient Boosting (XGBoost) algorithm to predict the suitability of crop types based on soil chemical characteristics and macro-environmental conditions. Model evaluation was conducted using the benchmark dataset Crop Recommendation Dataset accessed through the Kaggle platform. This dataset has a perfect class balance with a total of 2,200 samples evenly divided into 22 agricultural commodities. The developed predictive model evaluates seven soil and climate biophysical parameters, namely nitrogen, phosphorus, potassium, air temperature, relative humidity, soil acidity (pH), and rainfall intensity. The test results show that the XGBoost algorithm ranks top in classification accuracy with values ranging from 99.31% to 99.77%, surpassing other ensemble algorithms such as Random Forest as well as traditional models such as Decision Tree and Naive Bayes. Feature contribution analysis demonstrates that climate parameters (rainfall and humidity) act as primary ecological filters at the macro-level, while soil macronutrient (NPK) ratios serve as secondary determinants at the crop-specific micro-level. Overall, this boosting-based ensemble approach offers high accuracy and robustness to data outliers, making it a highly reliable agronomic decision-making tool for supporting sustainable land productivity.*

**Keywords:** Precision Agriculture; Classification; Extreme Gradient Boosting (XGBoost), Feature Engineering

## Abstrak

Penerapan teknologi kecerdasan buatan dalam sektor pertanian menjadi fondasi utama dalam pergeseran paradigma menuju pertanian presisi yang berkelanjutan. Studi ini menyajikan analisis komprehensif mengenai penerapan algoritma Extreme Gradient Boosting (XGBoost) untuk memprediksi kesesuaian jenis tanaman berdasarkan karakteristik kimia tanah dan kondisi lingkungan makro. Evaluasi model dilakukan menggunakan dataset benchmark *Crop Recommendation Dataset* yang diakses melalui platform Kaggle. Dataset ini memiliki keseimbangan kelas yang sempurna dengan total 2.200 sampel yang terbagi rata ke dalam 22 komoditas pertanian. Model prediktif yang dikembangkan mengevaluasi tujuh parameter biofisik tanah dan iklim, yaitu nitrogen, fosfor, kalium, suhu udara, kelembapan relatif, derajat keasaman (pH) tanah, dan intensitas curah hujan. Hasil pengujian menunjukkan bahwa algoritma XGBoost menempati peringkat teratas dalam akurasi klasifikasi dengan nilai berkisar antara 99,31% hingga 99,77%, melampaui algoritma ensemble lainnya seperti Random Forest

serta model tradisional seperti Decision Tree dan Naive Bayes. Analisis kontribusi fitur membuktikan bahwa parameter iklim (curah hujan dan kelembapan) bertindak sebagai filter ekologis utama pada tingkat makro, sedangkan rasio makronutrien (NPK) tanah berfungsi sebagai penentu sekunder pada tingkat mikro spesifik jenis tanaman. Secara keseluruhan, pendekatan ensemble berbasis boosting ini menawarkan akurasi tinggi and ketahanan terhadap pencilaan data, menjadikannya instrumen keputusan agronomis yang sangat andal untuk mendukung produktivitas lahan secara berkelanjutan.

**Kata kunci:** Pertanian Presisi; Klasifikasi; *Extreme Gradient Boosting* (XGBoost); Rekayasa Fitur.

## **1. Pendahuluan**

Sektor pertanian memegang peranan yang sangat vital dalam menyokong ketahanan pangan global, stabilitas ekonomi, dan pembangunan berkelanjutan (Nabiha et al., 2026). Namun, sektor ini kian dihadapkan pada tantangan yang semakin berat, seperti perubahan iklim yang tidak menentu, fluktuasi cuaca ekstrem, serta laju konversi lahan produktif menjadi kawasan non-pertanian. Ketidakpastian faktor lingkungan dan penurunan kualitas hara tanah berpotensi besar menurunkan produktivitas tanaman, yang pada gilirannya dapat memicu kerugian ekonomi signifikan bagi para petani akibat kegagalan panen (Woro Estiningtyas, 2025). Oleh sebab itu, transisi dari pola pertanian konvensional yang berbasis pada intuisi atau kebiasaan empiris menuju paradigma pertanian presisi (*precision agriculture*) berbasis data (*data-driven*) menjadi sebuah urgensi global guna memastikan efisiensi sumber daya dan keberlanjutan hasil pertanian (Garnida, 2026).

Dalam kerangka pertanian presisi, penentuan jenis komoditas tanaman yang paling sesuai dengan kondisi fisik dan kimiawi lahan merupakan langkah awal yang paling krusial. Rekomendasi kesesuaian tanaman secara ilmiah disusun dengan menganalisis interaksi dinamis antara parameter kesuburan kimiawi tanah—yaitu unsur makronutrien Nitrogen (N), Fosfor (P), dan Kalium (K) serta derajat keasaman (pH) tanah—dengan parameter iklim makro seperti suhu udara harian, kelembapan relatif, dan intensitas curah hujan harian (Han et al., 2021). Kompleksitas hubungan non-linear antar-variabel biofisik dan iklim ini menuntut adanya instrumen analisis yang cerdas dan mampu memberikan keputusan secara cepat, akurat, dan andal di tingkat lapangan (Q. Wang & Wang, 2026).

Penerapan teknologi kecerdasan buatan, khususnya algoritma pembelajaran mesin (*machine learning*), telah membuka peluang besar untuk memetakan pola kesesuaian lahan secara otomatis dan akurat (L. Wang et al., 2021). Model-model prediktif berbasis pohon keputusan (*tree-based models*), seperti Random Forest (RF) dan Extreme Gradient Boosting (XGBoost) (Heinen et al., 2022), telah terbukti memberikan performa klasifikasi yang sangat andal pada data tabular terstruktur di sektor agronomis. Berbeda dengan algoritma klasifikasi tunggal seperti Decision Tree tradisional yang rentan terhadap masalah *overfitting* dan sensitif terhadap pencilaan (*outliers*), algoritma *ensemble* berbasis *boosting* seperti XGBoost bekerja secara sekuensial dengan meminimalkan sisa kerugian (*loss*) dari model sebelumnya melalui optimalisasi gradien orde kedua (Chen & Guestrin, 2016). Selain itu, integrasi fungsi

regularisasi L2 secara langsung pada fungsi objektif XGBoost terbukti efektif mengontrol kompleksitas pohon dan meningkatkan generalisasi model pada data baru (Zhu et al., 2025).

Meskipun efektivitas algoritma ini telah banyak dilaporkan, beberapa penelitian terdahulu sering kali hanya memfokuskan evaluasi pada metrik akurasi tunggal tanpa menganalisis secara komprehensif pengaruh rekayasa fitur agronomis, stabilitas model terhadap variabilitas sampel, serta signifikansi teoritis dari kontribusi masing-masing fitur predictor (Alegia & Cerna, 2026). Studi ini hadir untuk mengisi celah tersebut dengan melakukan analisis komprehensif penerapan model XGBoost pada *Crop Recommendation Dataset* berskala benchmark. Melalui pendekatan terstruktur yang memadukan rekayasa fitur agronomis, optimasi parameter berbasis kisi (*GridSearchCV*), dan skema validasi silang bertingkat (*stratified 5-fold cross-validation*), studi ini bertujuan memetakan batas keputusan klasifikasi kesesuaian tanaman secara presisi sekaligus mengkaji secara mendalam stabilitas dan keandalan model guna mendukung efisiensi pengambilan keputusan pertanian di era digital.

## 2. Metode Penelitian

### Sumber dan Karakteristik Dataset

Penelitian ini memanfaatkan data sekunder terstruktur yang bersumber dari kumpulan data (dataset) rekomendasi tanaman yang mencakup simulasi kondisi lahan pertanian riil dengan menggabungkan serta mengaugmentasi data historis curah hujan, indikator iklim, dan rasio pupuk di wilayah India. Data ini mencakup 2.200 baris dengan delapan kolom, yang terdiri atas tujuh fitur numerik sebagai variabel prediktor dan satu kolom kategorikal sebagai variabel target (label) (Ingle, 2026).

**Tabel 1. Deskripsi Variabel Prediktor Dataset Kesesuaian Lahan**

| Simbol Fitur       | Nama Parameter | Satuan Ukur | Representasi Agronomis                                     | Rata-Rata (Mean) |
|--------------------|----------------|-------------|------------------------------------------------------------|------------------|
| <b>N</b>           | Nitrogen       | mg/kg       | Rasio nitrogen tanah untuk merangsang fase vegetatif daun. | 50,55            |
| <b>P</b>           | Fosfor         | mg/kg       | Rasio fosfor untuk pembelahan sel, akar, dan pembungaan.   | 53,36            |
| <b>K</b>           | Kalium         | mg/kg       | Rasio kalium untuk sirkulasi water dan ketahanan penyakit. | 48,15            |
| <b>temperature</b> | Suhu Udara     | °C          | Suhu rata-rata harian selama siklus pertumbuhan.           | 25,62            |

|                 |             |    |                                                                |        |
|-----------------|-------------|----|----------------------------------------------------------------|--------|
| <b>humidity</b> | Kelembapan  | %  | Kelembapan relatif udara yang memengaruhi evapotranspirasi.    | 71,48  |
| <b>ph</b>       | pH Tanah    | -  | Indikator keasaman/kebasaan tanah pembatas serapan unsur hara. | 6,47   |
| <b>rainfall</b> | Curah Hujan | mm | Akumulasi curah hujan harian pembatas ketersediaan air.        | 103,46 |

Sumber:Hasil Penelitian (2026)

Pada tabel 1 menjelaskan Variabel target mencakup 22 komoditas tanaman pangan, hortikultura, serat, dan perkebunan yang masing-masing diwakili secara seimbang oleh 100 sampel data. Keseimbangan kelas ini sangat krusial dalam pemodelan statistik karena mengeliminasi bias algoritma terhadap kelas mayoritas dan memastikan stabilitas metrik evaluasi seperti presisi dan sensitivitas di seluruh kategori tanaman.

### Rekayasa Fitur (*Feature Engineering*)

Guna memperkaya pola hubungan non-linear yang dapat dipelajari oleh model prediktif, dilakukan rekayasa fitur dengan membentuk empat variabel baru berdasarkan landasan teoretis agronomis. Pertama, variabel  $NPK\_mean$  dirumuskan melalui rata-rata aritmatika dari tiga makronutrien utama,  $NPK\_mean = (N+P+K)/3$ , untuk merepresentasikan tingkat kesuburan kimiawi tanah secara agregat. Kedua, dibentuk Indikator Cekaman Suhu-Kelembapan (*Temperature-Humidity Index* atau THI) yang dihitung dengan formula  $THI = (temperature \times humidity)/100$ , yang berguna untuk mengidentifikasi tingkat evapotranspirasi potensial pada wilayah pertumbuhan tanaman. Ketiga, dibuat fitur kategorikal  $ph\_category$  dengan mengelompokkan nilai pH menjadi tiga kelas utama, yaitu tanah asam ( $pH < 5,5$ ), netral ( $5,5 - 7,5$ ), dan alkalis ( $pH > 7,5$ ), untuk membantu model mengisolasi preferensi kimiawi spesifik tanaman. Keempat, fitur  $rainfall\_level$  dibentuk dengan mengklasifikasikan curah hujan ke dalam empat tingkat intensitas, yaitu rendah (0–50 mm), sedang (50–100 mm), tinggi (100–200 mm), dan sangat tinggi ( $>200$  mm), guna mempermudah pemisahan spasial komoditas basah dan kering.

### Pembagian Data dan Validasi Silang

Dataset dibagi menjadi subset pelatihan (*training set*) sebesar 80% (1.760 sampel) and subset pengujian (*testing set*) sebesar 20% (440 sampel). Pembagian ini menggunakan metode stratifikasi (*stratified split*) yang menjamin representasi merata dari ke-22 kelas tanaman pada kedua subset. Untuk menjamin kekuatan model terhadap variabilitas sampel dan mencegah masalah kebocoran data (*data leakage*), diterapkan skema validasi silang 5-lipat (*5-fold cross-validation*) selama fase pelatihan. Melalui metode ini, data pelatihan dibagi kembali menjadi lima subset acak secara bergantian, di mana empat subset digunakan untuk pelatihan dan satu

subset digunakan sebagai data validasi internal, sehingga menghasilkan estimasi performa model yang stabil dan objektif.

### Model dan Formulasi XGBoost

XGBoost merupakan algoritma tingkat lanjut berbasis *Gradient Boosted Decision Trees* (GBDT) yang mengoptimalkan fungsi kerugian melalui penambahan pohon keputusan baru secara iterative (Rahmattun Illahiyah et al., 2026). Secara matematis, prediksi akhir untuk sampel  $x_i$  pada iterasi ke- $t$  dinyatakan sebagai kombinasi aditif (Rizkallah, 2025).

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i)$$

Dimana  $\hat{y}_i^{(t-1)}$  adalah prediksi kumulatif dari langkah sebelumnya,  $f_t(x_i)$  adalah fungsi pemeta dari pohon keputusan baru pada Langkah ke- $t$ , dan  $\eta$  melambangkan laju pembelajaran (learning rate).

Fungsi objektif terregularisasi yang diminimalkan pada setiap langkah pelatihan dirumuskan sebagai berikut.

$$\text{Obj}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t)$$

Dimana  $l$  adalah fungsi kerugian diferensial yang mengukur selisih antara label actual  $y_i$  dan prediksi  $\hat{y}_i$ . Suku  $\Omega(f_t)$  mempresentasikan penalti regularisasi untuk mengontrol kompleksitas arsitektur pohon guna meredam overfitting, dengan formulasi (Zhu et al., 2025);

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

Pada suku tersebut,  $T$  menunjukkan jumlah daun pada pohon,  $w_j$  melambangkan bobot prediksi pada daun ke- $j$ ,  $\gamma$  adalah parameter pengendali kompleksitas pembelahan simpul, dan  $\lambda$  adalah parameter regularisasi L2 pada bobot daun. Aproksimasi orde kedua dari deret Taylor digunakan untuk mengoptimalkan fungsi objektif tersebut dengan cepat .

$$\text{Obj}^{(t)} \approx \sum_{i=1}^n \left[ l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

Dengan  $g_i$  dan  $h_i$  didefinisikan sebagai turunan pertama (*gradient*) dan kedua (*hessian*) dari fungsi kerugian terhadap prediksi langkah sebelumnya.

$$g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}} \quad \text{dan} \quad h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial (\hat{y}_i^{(t-1)})^2}$$

Bobot optimal  $w_j^*$  untuk daun ke- $j$  pada struktur pohon yang terbentuk dapat dihitung secara langsung melalui persamaan (Pramudhyta & Rohman, 2024):

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda}$$

Di mana  $I_j$  adalah himpunan indeks sampel yang dialokasikan ke dalam daun ke- $j$ .

Optimasi parameter model dilakukan menggunakan pencarian kisi terstandarisasi (*GridSearchCV*). Ruang parameter pencarian mencakup kedalaman maksimum pohon (*max\_depth* bernilai 3 hingga 8), jumlah estimator (*n\_estimators* bernilai 100 hingga 500), laju pembelajaran (*learning\_rate* bernilai 0,01 hingga 0,3), rasio sub-sampel kolom (*colsample\_bytree* bernilai 0,6 hingga 1,0), dan parameter kontrol pembelahan simpul (*gamma* bernilai 0 hingga 0,5). Struktur dasar pengklasifikasi dikonfigurasi menggunakan fungsi objektif *multi:softprob* untuk menghasilkan keluaran berupa probabilitas klasifikasi pada masing-masing dari 22 kelas target, dengan metrik evaluasi internal berupa log-loss multi-kelas (*mlogloss*) (Bukowski et al., 2024).

### **3. Hasil dan Pembahasan**

#### **Karakteristik Ekologis dan Sinergi Nutrisi Tanaman**

Analisis statistik deskriptif dan eksplorasi data mengungkapkan adanya pengelompokan ekologis yang sangat spesifik antar-varietas tanaman. Tanah netral merupakan media tumbuh yang paling dominan di dalam dataset, ditunjukkan oleh nilai tengah pH yang berkumpul stabil pada kisaran 6,0 hingga 7,5. Meskipun demikian, terdapat perbedaan karakteristik kebutuhan air dan nutrisi yang tajam di antara komoditas tersebut. Padi mendominasi kategori tanaman hidrofilik basah dengan rata-rata kebutuhan curah hujan mencapai 236,18 mm, disusul erat oleh kelapa dengan kebutuhan rata-rata 175,69 mm. Sebaliknya, tanaman buah seperti anggur dan delima tergolong ke dalam varietas xerofit dengan toleransi kekeringan tinggi, di mana rata-rata curah hujan yang dibutuhkan masing-masing hanya sebesar 69,61 mm dan 107,53 mm.

Uji korelasi Pearson menunjukkan pola hubungan linear yang lemah antar-variabel lingkungan, yang membuktikan signifikansi individual dari setiap parameter fisik dalam ruang keputusan prediktif. Pola korelasi kuat hanya ditemukan antara variabel Fosfor (P) dan Kalium (K) dengan nilai koefisien mencapai 0,74. Korelasi positif yang mencolok ini secara biologis didorong oleh kesamaan profil nutrisi pada tanaman apel dan anggur. Kedua komoditas buah tersebut menuntut pasokan kalium yang sangat masif (rata-rata K mencapai 199,89 mg/kg untuk apel dan 200,11 mg/kg untuk anggur) serta fosfor yang tinggi (rata-rata P mencapai 134,22 mg/kg untuk apel dan 132,53 mg/kg untuk anggur).

Pada spektrum nutrisi lainnya, tanaman katun menunjukkan kebutuhan nitrogen yang sangat tinggi (rata-rata 117,77 mg/kg) guna menyokong pertumbuhan biomassa seratnya. Sebaliknya, kelompok legum atau kacang-kacangan seperti lentil (N hara rata-rata 18,77 mg/kg), kacang hijau (N hara rata-rata 20,99 mg/kg), dan kacang merah (N hara rata-rata 20,75 mg/kg) menunjukkan kebutuhan nitrogen eksternal yang sangat rendah. Karakteristik ini berkaitan erat dengan simbiosis mutualisme antara akar legum dan bakteri pemfiksasi nitrogen

tanah, sehingga kebutuhan hara nitrogen dapat dipenuhi secara mandiri tanpa ketergantungan pada pemupukan kimia sintetis.

#### Perbandingan Performa Model Klasifikasi.

**Tabel 2: Perbandingan Performa Algoritma Klasifikasi Multi-Kelas**

| Algoritma Model | Akurasi Uji      |   | Presisi Makro | Sensitivitas (Recall) | F1-Score Makro |
|-----------------|------------------|---|---------------|-----------------------|----------------|
| XGBoost         | 99,09%<br>99,45% | - | 0,99          | 0,99                  | 0,99           |
| Random Forest   | 99,32%<br>99,81% | - | 0,99          | 0,99                  | 0,99           |
| Naive Bayes     | 90,09%<br>99,09% | - | 0,91          | 0,91                  | 0,91           |
| Decision Tree   | 90,01%<br>98,73% | - | 0,90          | 0,90                  | 0,90           |

Sumber: Hasil Penelitian (2026)

Pada tabel 2 menjelaskan Evaluasi kinerja model dijalankan dengan membandingkan performa algoritma XGBoost terhadap model Random Forest, Naive Bayes, dan Decision Tree. Seluruh model diuji pada dataset pengujian yang sama untuk menjamin validitas perbandingan.

Model XGBoost secara konsisten menduduki peringkat teratas dalam metrik klasifikasi keseluruhan. Keberhasilan algoritma ini berakar pada kemampuan mekanisme *boosting* yang bekerja secara adaptif untuk mengurangi kesalahan klasifikasi pada area batas keputusan yang tipis. Berbeda dengan Random Forest yang membangun pohon keputusan secara independen (paralel), XGBoost meminimalkan bias model secara sekuensial dengan merancang pohon baru yang mengoreksi sisa kerugian dari iterasi sebelumnya. Selain itu, penambahan fungsi penalti L2 pada daun pohon terbukti meredam volatilitas data, sehingga performa model tetap stabil tanpa mengalami penurunan kualitas klasifikasi (*overfitting*) pada subset pengujian.

Analisis performa per-kelas mengindikasikan tingkat klasifikasi sempurna (Presisi, Sensitivitas, dan F1-Score sebesar 1,00) untuk komoditas apel, anggur, buncis, kelapa, dan pisang. Hal ini dimungkinkan karena komoditas tersebut menempati ruang fitur yang sangat terisolasi dengan nilai ekstrem yang khas, seperti kebutuhan kalium tinggi pada anggur dan iklim basah bercurah hujan tinggi pada kelapa. Sebaliknya, komoditas padi menunjukkan

fluktuasi akurasi kelas yang sedikit lebih rendah pada beberapa pengujian (presisi sekitar 0,76 dan sensitivitas 0,81). Penurunan minor ini disebabkan adanya kemiripan pola spasial parameter curah hujan dan kelembapan yang tinggi antara padi dengan tanaman tropis basah lainnya seperti goni (*jute*) dan kelapa, sehingga menyebabkan beberapa sampel mengalami salah klasifikasi tipis pada area perbatasan kelas.

#### **Kontribusi Fitur dan Struktur Keputusan**

Melalui ekstraksi pentingnya fitur (*feature importance*), kontribusi relatif dari masing-masing parameter fisik tanah dan iklim dapat dievaluasi secara terukur. Temuan model menunjukkan bahwa parameter iklim, khususnya curah hujan dan tingkat kelembapan relatif, memberikan kontribusi akumulatif terbesar dalam pembentukan simpul keputusan utama. Pola curah hujan bertindak sebagai filter ekologis makro yang memisahkan dataset ke dalam kelompok besar (misalnya mengisolasi padi dan kelapa dari varietas xerofit).

Setelah penyaringan makro iklim terjadi, XGBoost menggunakan rasio unsur Nitrogen (N) dan Kalium (K) sebagai fitur penentu mikro untuk menetapkan prediksi komoditas akhir. Sebagai contoh, pada area dengan curah hujan sedang dan kelembapan tinggi, model mengandalkan fitur Nitrogen untuk mendiferensiasi antara katun (kebutuhan N hara tinggi) dengan tanaman kacang-kacangan (kebutuhan N hara sangat rendah). Di sisi lain, fitur suhu udara dan pH tanah menunjukkan skor kontribusi yang lebih rendah. Rendahnya skor kontribusi ini bukan berarti kedua parameter tersebut tidak penting secara agronomis, melainkan karena nilai suhu dan pH tanah dalam dataset ini sebagian besar terdistribusi secara seragam pada kondisi optimal, sehingga kurang memiliki daya pemisah biner yang kuat dibandingkan dengan variabel curah hujan dan nitrogen.

#### **4. Kesimpulan**

Sistem rekomendasi komoditas pertanian berbasis kecerdasan buatan terbukti dapat diimplementasikan dengan tingkat akurasi dan keandalan yang sangat tinggi menggunakan algoritma Extreme Gradient Boosting (XGBoost). Evaluasi kinerja model pada *dataset* yang digunakan membuktikan dominasi dari algoritma *ensemble* berbasis *boosting* ini dengan raihan akurasi rata-rata di atas 99%, bersaing ketat dengan metode Random Forest serta secara signifikan melampaui Naive Bayes dan Decision Tree. Hasil analisis agronomis menunjukkan bahwa parameter iklim berupa curah hujan dan kelembapan berperan sebagai filter ekologis makro, sedangkan rasio makronutrien kimia tanah (NPK) bertindak sebagai penentu mikro yang memisahkan klasifikasi jenis komoditas secara presisi. Sinergi nutrisi yang kuat terlihat pada tingginya korelasi antara fosfor dan kalium ( $\sim 0,74$ ) akibat kebutuhan ekstrem berpasangan pada tanaman buah bernilai tinggi seperti apel dan anggur.

Secara praktis, model prediktif XGBoost ini menawarkan solusi rekomendasi pertanian yang mandiri dan tidak memerlukan ketergantungan pada uji laboratorium fisik yang mahal. Untuk jangka panjang, sistem rekomendasi ini dapat diintegrasikan dengan perangkat keras mikrokontroler berbasis sensor nutrisi tanah NPK waktu-nyata serta sensor iklim lingkungan

untuk menangkap fluktuasi parameter biofisik secara langsung. Dengan menyandingkan model prediktif berbasis cloud ini bersama integrasi API cuaca eksternal dan pergerakan harga komoditas pasar, petani dapat memperoleh sistem pendukung keputusan yang tidak hanya adaptif terhadap dinamika alam, tetapi juga dirancang untuk memaksimalkan margin keuntungan finansial usaha tani secara berkelanjutan.

#### **Daftar Pustaka**

- Alegia, J. G., & Cerna, P. D. (2026). From Prediction to Decision: An Explainable Machine Learning–Driven Framework for Precision Sugarcane Agriculture. *Proceedings of the 2026 International Conference on AI Decision-Making and Management*, 106–114. <https://doi.org/10.1145/3811839.3811856>
- Bukowski, M., Kurek, J., Świdorski, B., & Jegorowa, A. (2024). Custom Loss Functions in XGBoost Algorithm for Enhanced Critical Error Mitigation in Drill-Wear Analysis of Melamine-Faced Chipboard. *Sensors*, 24(4), 1092. <https://doi.org/10.3390/s24041092>
- Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Garnida, Y. (2026). *Precision Agriculture Technology: Accelerating Sustainable Food Security and Resource Efficiency Yudi Garnida*. 4(01), 41–50.
- Han, C., Chen, S., Yu, Y., Xu, Z., Zhu, B., Xu, X., & Wang, Z. (2021). Evaluation of Agricultural Land Suitability Based on RS, AHP, and MEA: A Case Study in Jilin Province, China. *Agriculture*, 11(4), 370. <https://doi.org/10.3390/agriculture11040370>
- Heinen, M., Mulder, H. M., Bakker, G., Wösten, J. H. M., Brouwer, F., Teuling, K., & Walvoort, D. J. J. (2022). The Dutch soil physical units map: BOFEK. *Geoderma*, 427, 116123. <https://doi.org/10.1016/j.geoderma.2022.116123>
- Ingle, A. (2026). *Crop Recommendation Dataset*. Kaggle.
- Nabiha, F. H., Octaviani, F. T., Ardhini, P., Hidayati, R., Destian, R., Permadi, S. Z., & Nuraya, A. S. (2026). Peran Sektor Pertanian Dalam Ketahanan Pangan Dan Perekonomian Nasional. *QOSIM: Jurnal Pendidikan, Sosial & Humaniora*, 4(3), 946–961.
- Pramudhyta, N. A., & Rohman, M. S. (2024). Perbandingan Optimasi Metode Grid Search dan Random Search dalam Algoritma XGBoost untuk Klasifikasi Stunting. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 8(1), 19. <https://doi.org/10.30865/mib.v8i1.6965>
- Rahmattun Illahiyah, Ginanjar Abdurrahman, & Daryanto. (2026). Implementation of the XGBoost Classifier Algorithm for Classifying Crowds of Hajj and Umrah Pilgrims. *MALCOM: Indonesian Journal of Machine and Computer Science*, 6(1), 106–115.
- Rizkallah, L. W. (2025). Enhancing the performance of gradient boosting trees on regression problems. *Journal of Big Data*, 12(1). <https://doi.org/10.1186/s40537-025-01071-3>
- Wang, L., Xu, M., Hou, Q., Wang, Z., Lan, Y., & Wang, S. (2021). Numerical verification on influence of multi-feature parameters to the downwash airflow field and operation effect of

- a six-rotor agricultural UAV in flight. *Computers and Electronics in Agriculture*, 190, 106425. <https://doi.org/10.1016/j.compag.2021.106425>
- Wang, Q., & Wang, T. (2026). Machine learning-driven sensor fusion for precision agriculture: current applications and future perspectives. *Sensor Review*, 1–26. <https://doi.org/10.1108/SR-08-2025-0613>
- Woro Estiningtyas. (2025). Kajian Dampak Perubahan Iklim pada Sektor Pertanian: Upaya Strategis Adaptasi untuk Mendukung Ketahanan Pangan. In *Prediksi Iklim untuk Ketahanan Pangan*. Penerbit BRIN. <https://doi.org/10.55981/brin.1244.c1386>
- Zhu, C., Huang, M., Ren, F., Li, X., Gu, J., Li, H., & He, M. (2025). Multivariate acoustic emissions precursors of rockburst from the perspective of early warning. *International Journal of Mining Science and Technology*, 35(5), 703–717. <https://doi.org/10.1016/j.ijmst.2025.04.002>