



Implementation Of The Decision Tree Algorithm For Seed Selection Recommendation, Harvest Prediction, And Profit Estimation In The Agricultural Sector

Implementasi Algoritma Decision Tree Untuk Rekomendasi Pemilihan Bibit, Prediksi Hasil Panen, Dan Estimasi Keuntungan Pada Sektor Pertanian

Titan Attariq Alfatah ¹⁾, Suheila Kirmalani ²⁾, Adi Suwondo ³⁾

¹²³ Informatics Engineering, Faculty of Engineering and Computer Science, Universitas Sains Al-Qur'an (UNSIQ) Central Java in Wonosobo, Indonesia

Email: ¹⁾ titanalfatah123@gmail.co ; ²⁾ riskhasuhela@gmail.com ; ³⁾ adisuwondo@unsiq.ac.id3

How to Cite :

Alfatah, T. A., Kirmalani, S., Suwondo, A. (2026). Implementation of the Decision Tree Algorithm for Seed Selection Recommendation, Harvest Prediction, and Profit Estimation in the Agricultural Sector Jurnal Komputer Indonesia, 5(3).

ARTICLE HISTORY

Received [26 Juli 2026]

Revised [08 September 2026]

Accepted [10 September 2026]

KEYWORDS

Decision Tree, Data Mining, Agriculture, Harvest, Prediction, Seed Recommendation, Profit Estimation.

This is an open access article under the [CC-BY-SA](#) license



ABSTRAK

Sektor pertanian memerlukan pengambilan keputusan yang tepat dalam menentukan jenis bibit, memprediksi hasil panen, serta memperkirakan keuntungan berdasarkan kondisi lahan dan faktor budidaya. Penelitian ini bertujuan mengimplementasikan algoritma Decision Tree sebagai metode klasifikasi untuk memberikan rekomendasi pemilihan bibit, memprediksi hasil panen, dan mengestimasi tingkat keuntungan petani. Variabel yang digunakan meliputi jenis tanaman, jenis bibit, pH tanah, kelembapan tanah, curah hujan, suhu, luas lahan, jenis pupuk, biaya produksi, dan tingkat serangan hama. Tahapan penelitian meliputi pengumpulan data, preprocessing, pembentukan model Decision Tree, pelatihan model, serta evaluasi menggunakan confusion matrix, accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa algoritma Decision Tree mampu menghasilkan model klasifikasi yang efektif dalam memberikan rekomendasi bibit sesuai kondisi lahan, memprediksi kategori hasil panen, serta mengestimasi keuntungan petani. Sistem yang dikembangkan diharapkan dapat menjadi sistem pendukung keputusan untuk membantu petani meningkatkan produktivitas dan efisiensi pengelolaan usaha tani.

ABSTRACT

The agricultural sector requires accurate decision-making in selecting suitable seeds, predicting harvest yields, and estimating farming profits based on land conditions and cultivation factors. This study aims to implement the Decision Tree algorithm as a classification method to provide seed selection recommendations, predict harvest yields, and estimate farmers' profits. The variables used include crop type, seed type, soil pH, soil moisture, rainfall, temperature, land area, fertilizer type, production cost, and pest infestation level. The research process consists of data collection, preprocessing, Decision Tree model development, model training, and evaluation using a confusion matrix, accuracy, precision, recall, and F1-score. The results indicate that the Decision Tree algorithm is effective in generating classification models capable of recommending appropriate seed types, predicting harvest categories, and estimating farming profits. The developed system is expected to support farmers in improving productivity and making data-driven agricultural decisions.

PENDAHULUAN

Sektor pertanian merupakan salah satu sektor strategis yang berperan penting dalam mendukung ketahanan pangan, meningkatkan kesejahteraan masyarakat, serta memberikan kontribusi terhadap pertumbuhan ekonomi nasional. Menurut Badan Pusat Statistik (BPS), sektor pertanian masih menjadi penyerap tenaga kerja terbesar di Indonesia sehingga peningkatan produktivitas pertanian menjadi salah satu prioritas pembangunan nasional. Namun demikian, produktivitas pertanian masih dipengaruhi oleh berbagai faktor, seperti kualitas bibit, kondisi tanah, curah hujan, suhu lingkungan, kelembapan tanah, penggunaan pupuk, serangan hama, serta biaya produksi. Faktor-faktor tersebut menyebabkan proses pengambilan keputusan dalam kegiatan budidaya masih banyak bergantung pada pengalaman petani sehingga sering menghasilkan keputusan yang kurang optimal (Wolfert et al., 2017; Liakos et al., 2018).

Perkembangan teknologi informasi telah mendorong pemanfaatan Artificial Intelligence (AI) dan Machine Learning dalam berbagai sektor, termasuk pertanian. Salah satu implementasinya adalah penggunaan teknik Data Mining untuk mengekstraksi pola dari data historis sehingga dapat menghasilkan informasi yang berguna dalam proses pengambilan keputusan. Data mining memungkinkan data pertanian yang sebelumnya hanya tersimpan menjadi pengetahuan yang dapat dimanfaatkan dalam menentukan strategi budidaya yang lebih efektif. Dengan memanfaatkan data historis mengenai kondisi lahan, cuaca, jenis tanaman, maupun hasil panen sebelumnya, proses prediksi dapat dilakukan secara lebih objektif dibandingkan dengan metode konvensional (Han et al., 2012; Tan et al., 2019).

Salah satu algoritma klasifikasi yang banyak digunakan dalam bidang data mining adalah Decision Tree. Algoritma ini bekerja dengan membangun struktur pohon keputusan berdasarkan atribut yang memiliki nilai Information Gain atau Gini Index tertinggi sehingga menghasilkan aturan klasifikasi yang mudah dipahami oleh pengguna. Dibandingkan dengan beberapa algoritma klasifikasi lainnya, Decision Tree memiliki keunggulan berupa proses pembentukan model yang relatif cepat, mampu menangani data kategorikal maupun numerik, mudah diinterpretasikan, serta memiliki tingkat akurasi yang cukup tinggi pada berbagai kasus klasifikasi (Quinlan, 1986; Breiman et al., 1984). Oleh karena itu, algoritma ini banyak diterapkan pada sistem pendukung keputusan, analisis bisnis, bidang kesehatan, pendidikan, hingga sektor pertanian.

Dalam bidang pertanian, algoritma Decision Tree telah dimanfaatkan untuk berbagai kebutuhan, seperti klasifikasi kualitas tanah, rekomendasi pemupukan, prediksi hasil panen, klasifikasi tingkat kesuburan lahan, serta pemilihan jenis tanaman berdasarkan karakteristik lingkungan. Penelitian yang dilakukan oleh Suradiradja et al. (2022) menunjukkan bahwa algoritma Decision Tree mampu memberikan estimasi waktu panen tanaman sorgum dengan tingkat akurasi yang baik berdasarkan kondisi lingkungan. Penelitian lain oleh Bishnoi dan Hooda (2022) menjelaskan bahwa Decision Tree merupakan salah satu algoritma yang efektif dalam melakukan klasifikasi data pertanian karena menghasilkan aturan keputusan yang mudah dipahami oleh petani maupun penyuluh pertanian. Selain itu, penelitian Prasetyo et al. (2026) membuktikan bahwa penggunaan data iklim yang dikombinasikan dengan algoritma Decision Tree mampu meningkatkan akurasi prediksi hasil panen tanaman padi.

Meskipun demikian, sebagian besar penelitian sebelumnya hanya berfokus pada satu tujuan, misalnya hanya melakukan prediksi hasil panen, klasifikasi kualitas tanah, atau pemilihan jenis tanaman. Penelitian terdahulu juga umumnya belum mengintegrasikan beberapa faktor penting, seperti kondisi tanah, karakteristik bibit, biaya produksi, luas lahan, jenis pupuk, hingga serangan hama dalam satu model klasifikasi. Akibatnya, sistem yang dikembangkan belum mampu memberikan rekomendasi yang komprehensif untuk mendukung seluruh proses pengambilan keputusan petani (Kamilaris & Prenafeta-Boldú, 2018; Liakos et al., 2018).

Berdasarkan kondisi tersebut, masih terdapat research gap, yaitu belum banyak penelitian yang mengembangkan sistem berbasis Decision Tree yang mampu mengintegrasikan proses



rekomendasi pemilihan bibit, prediksi hasil panen, serta estimasi keuntungan petani secara bersamaan. Ketiga aspek tersebut sebenarnya saling berkaitan dalam proses budidaya tanaman karena pemilihan bibit akan memengaruhi produktivitas tanaman, sedangkan produktivitas akan menentukan besarnya pendapatan dan keuntungan yang diperoleh petani. Oleh karena itu, diperlukan suatu sistem yang mampu memberikan rekomendasi secara terpadu sehingga dapat membantu petani dalam mengambil keputusan yang lebih tepat.

Kebaruan (novelty) pada penelitian ini terletak pada pengembangan model klasifikasi menggunakan algoritma Decision Tree yang tidak hanya digunakan untuk memprediksi hasil panen, tetapi juga memberikan rekomendasi pemilihan bibit serta mengestimasi tingkat keuntungan petani berdasarkan kombinasi berbagai variabel pertanian, yaitu jenis tanaman, jenis bibit, pH tanah, kelembapan tanah, curah hujan, suhu, luas lahan, jenis pupuk, biaya produksi, dan tingkat serangan hama. Pendekatan tersebut diharapkan mampu menghasilkan sistem pendukung keputusan yang lebih komprehensif dibandingkan penelitian-penelitian sebelumnya.

LANDASAN TEORI

Data Mining

Data Mining merupakan proses menemukan pola, hubungan, atau informasi yang bernilai dari sekumpulan data berukuran besar menggunakan teknik statistik, kecerdasan buatan (*Artificial Intelligence*), dan *Machine Learning*. Tujuan utama data mining adalah mengubah data mentah menjadi informasi yang dapat digunakan sebagai dasar pengambilan keputusan. Proses data mining umumnya meliputi tahapan *data cleaning*, *data integration*, *data selection*, *data transformation*, *data mining*, *pattern evaluation*, dan *knowledge presentation* (Han et al., 2012; Tan et al., 2019).

Dalam sektor pertanian, data mining dimanfaatkan untuk menganalisis berbagai data, seperti kondisi tanah, curah hujan, kelembapan, suhu, penggunaan pupuk, hingga hasil panen pada musim sebelumnya. Analisis tersebut dapat menghasilkan pola yang membantu petani dalam menentukan strategi budidaya yang lebih efektif dan efisien. Seiring berkembangnya teknologi digital, data mining menjadi salah satu komponen utama dalam pengembangan sistem Smart Agriculture yang mendukung pengambilan keputusan berbasis data (Liakos et al., 2018).

Machine Learning

Machine Learning merupakan cabang dari kecerdasan buatan yang memungkinkan komputer mempelajari pola dari data tanpa harus diprogram secara eksplisit. Algoritma Machine Learning membangun model berdasarkan data pelatihan (*training data*) sehingga mampu melakukan prediksi terhadap data baru (*testing data*). Berdasarkan jenis pembelajarannya, Machine Learning dibagi menjadi tiga kategori utama, yaitu Supervised Learning, Unsupervised Learning, dan Reinforcement Learning (Mitchell, 1997; Bishop, 2006).

Pada penelitian ini digunakan pendekatan Supervised Learning, karena data yang digunakan telah memiliki label keluaran berupa kategori hasil panen dan kategori keuntungan. Pendekatan ini sesuai untuk permasalahan klasifikasi, di mana model belajar dari data historis untuk memprediksi kelas pada data baru.

Decision Tree

Decision Tree merupakan algoritma klasifikasi yang membentuk struktur pohon keputusan (*tree structure*) berdasarkan atribut yang memiliki kemampuan terbaik dalam memisahkan data. Setiap simpul (*node*) merepresentasikan atribut, cabang (*branch*) menunjukkan hasil pengujian atribut, sedangkan daun (*leaf*) merupakan hasil klasifikasi atau keputusan akhir (Quinlan, 1986; Breiman et al., 1984).

Keunggulan algoritma Decision Tree meliputi:

- mudah dipahami dan diinterpretasikan;
- mampu menangani data numerik maupun kategorikal;

- tidak memerlukan proses normalisasi data yang kompleks;
- mampu menghasilkan aturan keputusan (*decision rules*);
- memiliki proses komputasi yang relatif cepat.

Algoritma Decision Tree memilih atribut terbaik menggunakan ukuran seperti Information Gain, Gain Ratio, atau Gini Index. Pada algoritma C4.5, atribut dipilih berdasarkan nilai Gain Ratio, sedangkan pada algoritma CART digunakan Gini Index sebagai dasar pembentukan pohon keputusan.

Secara umum, tahapan pembentukan Decision Tree meliputi:

1. Menentukan atribut terbaik sebagai akar pohon.
2. Membagi data berdasarkan atribut tersebut.
3. Mengulangi proses pada setiap cabang hingga seluruh data berada pada kelas yang sama atau memenuhi kondisi penghentian.

Karena menghasilkan aturan keputusan yang sederhana, Decision Tree banyak diterapkan pada sistem pendukung keputusan, diagnosis penyakit, analisis kredit, pendidikan, serta pertanian (Quinlan, 1993).

Sistem Pendukung Keputusan (Decision Support System)

Sistem Pendukung Keputusan (*Decision Support System/DSS*) merupakan sistem berbasis komputer yang dirancang untuk membantu pengguna dalam mengambil keputusan terhadap permasalahan semi-terstruktur maupun tidak terstruktur. DSS memanfaatkan data, model analisis, serta metode komputasi sehingga mampu menghasilkan rekomendasi keputusan yang lebih objektif dibandingkan pendekatan manual (Turban et al.; Power, 2002).

Dalam bidang pertanian, DSS dimanfaatkan untuk menentukan jadwal tanam, memilih jenis bibit, memperkirakan kebutuhan pupuk, memprediksi hasil panen, hingga menganalisis keuntungan usaha tani. Integrasi DSS dengan algoritma Machine Learning memungkinkan sistem memberikan rekomendasi secara otomatis berdasarkan data historis maupun kondisi lahan saat ini.

Smart Agriculture

Smart Agriculture merupakan konsep pertanian modern yang mengintegrasikan teknologi informasi, Internet of Things (IoT), sensor, *cloud computing*, serta kecerdasan buatan untuk meningkatkan efisiensi budidaya tanaman. Melalui konsep ini, berbagai parameter seperti kelembapan tanah, suhu udara, intensitas cahaya, curah hujan, dan kondisi tanaman dapat dipantau secara real-time sehingga membantu petani dalam mengambil keputusan secara cepat dan tepat (Wolfert et al., 2017; Kamilaris & Prenafeta-Boldú, 2018).

Penggunaan algoritma Machine Learning dalam Smart Agriculture memungkinkan proses prediksi hasil panen, klasifikasi kualitas tanah, identifikasi penyakit tanaman, serta rekomendasi budidaya dilakukan secara otomatis. Oleh karena itu, penerapan algoritma Decision Tree dalam penelitian ini merupakan bagian dari implementasi teknologi Smart Agriculture berbasis data.

Evaluasi Model Klasifikasi

Kinerja model klasifikasi perlu dievaluasi untuk mengetahui kemampuan model dalam melakukan prediksi. Pada penelitian ini digunakan Confusion Matrix sebagai dasar perhitungan beberapa metrik evaluasi, yaitu Accuracy, Precision, Recall, dan F1-Score.

Accuracy

Accuracy menunjukkan persentase prediksi yang benar dibandingkan dengan seluruh data pengujian.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$



Precision

Precision menunjukkan tingkat ketepatan model dalam memberikan prediksi positif.

$$Precision = \frac{TP}{TP + FP}$$

Recall

Recall menunjukkan kemampuan model dalam menemukan seluruh data positif yang sebenarnya.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

F1-Score merupakan rata-rata harmonik antara Precision dan Recall sehingga memberikan gambaran keseimbangan performa model.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Keempat metrik tersebut banyak digunakan dalam penelitian klasifikasi karena mampu memberikan gambaran yang lebih komprehensif mengenai performa model dibandingkan hanya menggunakan nilai akurasi saja (Han et al., 2012; Bishop, 2006).

Keuntungan Petani

Keuntungan petani merupakan indikator utama dalam mengukur keberhasilan suatu usaha tani yang diperoleh dari selisih antara total pendapatan hasil penjualan panen dengan seluruh biaya produksi yang dikeluarkan selama proses budidaya. Pendapatan dipengaruhi oleh jumlah hasil panen dan harga jual komoditas, sedangkan biaya produksi meliputi biaya pembelian bibit, pupuk, pestisida, tenaga kerja, pengolahan lahan, irigasi, serta biaya operasional lainnya. Semakin besar selisih antara pendapatan dan biaya produksi, maka semakin tinggi tingkat keuntungan yang diperoleh petani, sehingga usaha tani dapat dikatakan lebih efisien dan menguntungkan.

Secara matematis, keuntungan usaha tani dapat dirumuskan sebagai berikut:

$$\text{Keuntungan} = \text{Total Pendapatan} - \text{Total Biaya Produksi}$$

dengan:

- Total Pendapatan (TR) = Harga Jual × Jumlah Hasil Panen
- Total Biaya Produksi (TC) = Biaya Bibit + Biaya Pupuk + Biaya Pestisida + Biaya Tenaga Kerja + Biaya Operasional Lainnya

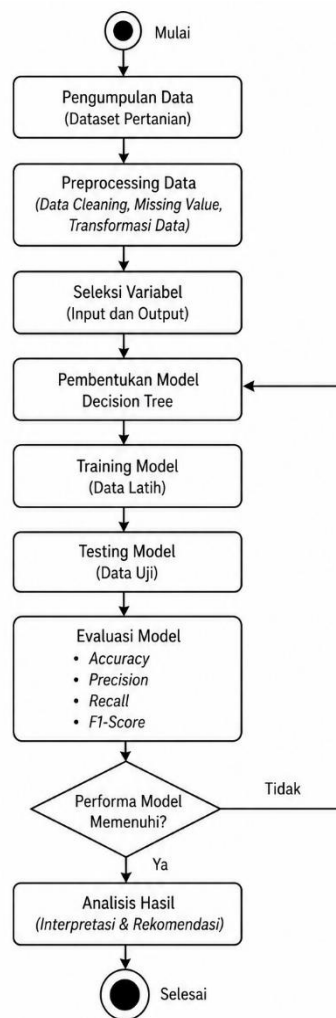
Keuntungan petani dipengaruhi oleh berbagai faktor, antara lain kualitas bibit, kondisi tanah, curah hujan, suhu lingkungan, penggunaan pupuk, serangan hama, luas lahan, serta efektivitas pengelolaan budidaya. Pemilihan bibit yang tepat dan penerapan teknik budidaya yang sesuai dapat meningkatkan produktivitas tanaman sehingga berdampak langsung terhadap peningkatan pendapatan petani. Sebaliknya, penggunaan bibit yang kurang sesuai atau tingginya biaya produksi dapat menurunkan tingkat keuntungan meskipun hasil panen relatif tinggi.

Dalam penelitian ini, estimasi keuntungan petani merupakan salah satu variabel keluaran (*output*) yang diprediksi menggunakan algoritma Decision Tree. Model klasifikasi dibangun berdasarkan sejumlah variabel masukan, yaitu jenis tanaman, jenis bibit, pH tanah, kelembapan tanah, curah hujan, suhu, luas lahan, jenis pupuk, biaya produksi, dan tingkat serangan hama. Hasil prediksi diharapkan mampu membantu petani memperkirakan tingkat keuntungan sebelum proses budidaya dilakukan, sehingga keputusan mengenai pemilihan bibit dan strategi budidaya dapat dilakukan secara lebih tepat, objektif, dan berbasis data.

METODE PENELITIAN

Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun model klasifikasi menggunakan algoritma Decision Tree. Pendekatan kuantitatif dipilih karena penelitian memanfaatkan data numerik dan kategorikal untuk menghasilkan model prediksi yang dapat digunakan sebagai sistem pendukung keputusan pada sektor pertanian. Tahapan penelitian meliputi pengumpulan data, *preprocessing*, pembentukan model, pelatihan (*training*), pengujian (*testing*), evaluasi model, dan analisis hasil.



Gambar 1. Tahapan Penelitian

Tahapan tersebut dijelaskan sebagai berikut.

Pengumpulan Data

Data penelitian diperoleh dari data pertanian yang berisi informasi mengenai kondisi lahan, karakteristik tanaman, biaya produksi, dan hasil panen. Variabel yang digunakan dipilih berdasarkan faktor-faktor yang berpengaruh terhadap keberhasilan budidaya tanaman.



Preprocessing Data

Pada tahap ini dilakukan proses pembersihan data (*data cleaning*) untuk menghilangkan data yang tidak lengkap (*missing value*), memperbaiki data yang tidak konsisten, serta menghapus data duplikat. Selanjutnya dilakukan transformasi data agar seluruh variabel siap digunakan dalam proses pelatihan model.

Seleksi Variabel

Variabel yang digunakan terdiri atas variabel masukan (*input*) dan variabel keluaran (*output*).

Tabel 1. Seleksi Variabel

Variabel Jenis	Variabel Jenis
Jenis Tanaman Input	Jenis Tanaman Input
Jenis Bibit Input	Jenis Bibit Input
pH Tanah Input	pH Tanah Input
Kelembapan Tanah Input	Kelembapan Tanah Input
Curah Hujan Input	Curah Hujan Input
Suhu Input	Suhu Input
Luas Lahan Input	Luas Lahan Input
Jenis Pupuk Input	Jenis Pupuk Input
Biaya Produksi Input	Biaya Produksi Input
Tingkat Serangan Hama Input	Tingkat Serangan Hama Input
Kategori Hasil Panen Output	Kategori Hasil Panen Output
Kategori Keuntungan Output	Kategori Keuntungan Output

Algoritma Decision Tree

Model klasifikasi dibangun menggunakan algoritma Decision Tree. Algoritma ini membentuk struktur pohon keputusan berdasarkan atribut yang memiliki nilai *Information Gain* tertinggi sehingga mampu menghasilkan aturan klasifikasi yang mudah dipahami.

Tahapan pembentukan model meliputi:

1. Menentukan atribut akar (*root node*).
2. Menghitung nilai entropy setiap atribut.
3. Menghitung nilai information gain.
4. Memilih atribut dengan nilai gain tertinggi.
5. Membentuk cabang keputusan.
6. Mengulangi proses hingga seluruh data berada pada kelas yang sama atau memenuhi kondisi penghentian.

Entropy dihitung menggunakan persamaan berikut.

$$Entropy(S) = - \sum_{i=1}^n p_i \log_2(p_i)$$

Selanjutnya, nilai Information Gain dihitung menggunakan persamaan berikut.

$$Gain(S, A) = Entropy(S) - \sum_{v \in A} \frac{|S_v|}{|S|} Entropy(S_v)$$

Keterangan:

- S = himpunan data
- A = atribut
- p_i = proporsi data pada kelas ke- i
- S_v = subset data berdasarkan atribut

Implementasi Model

Model Decision Tree diimplementasikan menggunakan bahasa pemrograman Python dengan pustaka Scikit-learn. Proses pengolahan data dilakukan menggunakan Pandas, sedangkan visualisasi pohon keputusan menggunakan Matplotlib dan Graphviz. Dataset dibagi menjadi data latih (*training data*) sebesar 80% dan data uji (*testing data*) sebesar 20% menggunakan metode *train-test split*.

Evaluasi Model

Kinerja model dievaluasi menggunakan Confusion Matrix. Berdasarkan hasil klasifikasi, dihitung empat metric evaluasi, yaitu Accuracy, Precision, Recall, dan F1-Score. Accuracy menunjukkan persentase prediksi yang benar terhadap seluruh data uji.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision menunjukkan ketepatan model dalam mengidentifikasi data positif.

$$Precision = \frac{TP}{TP + FP}$$

Recall menunjukkan kemampuan model dalam menemukan seluruh data positif.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score merupakan rata-rata harmonik antara Precision dan Recall.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Model dikatakan memiliki performa yang baik apabila nilai Accuracy, Precision, Recall, dan F1-Score mendekati 100%.

Perangkat Penelitian

Perangkat yang digunakan dalam penelitian ini terdiri atas perangkat keras dan perangkat lunak sebagaimana ditunjukkan pada Tabel 2.



Tabel 2. Perangkat Penelitian

Komponen	Spesifikasi
Sistem Operasi	Windows 11
Bahasa Pemrograman	Python 3.12
Library	Scikit-learn, Pandas, NumPy, Matplotlib
IDE	Visual Studio Code / Jupyter Notebook
Database	Microsoft Excel / CSV
Metode	Decision Tree

HASIL DAN PEMBAHASAN

Hasil

Pengolahan Data

Penelitian ini menggunakan dataset yang terdiri atas berbagai parameter pertanian yang berpengaruh terhadap keberhasilan budidaya tanaman. Data yang digunakan meliputi jenis tanaman, jenis bibit, kondisi tanah, kondisi lingkungan, penggunaan pupuk, biaya produksi, dan tingkat serangan hama. Setelah melalui tahap *preprocessing*, seluruh data dinyatakan layak untuk digunakan dalam proses pembentukan model klasifikasi menggunakan algoritma Decision Tree.

Variabel penelitian yang digunakan terdiri atas variabel masukan (*input*) dan variabel keluaran (*output*) sebagaimana ditunjukkan pada Tabel 3.

Tabel 3. Variabel Penelitian

No	Vriabel	Jenis
1	Jenis Tanaman	Input
2	Jenis Bibit	Input
3	pH Tanah	Input
4	pH Tanah	Input
5	Curah Hujan	Input
6	Suhu	Input
7	Jenis Pupuk	Input
8	Biaya Produksi	Input
9	Serangan Hama	Input
10	Kategori Panen	Output
11	Kategori Keuntungan	Output

Variabel-variabel tersebut dipilih karena memiliki pengaruh terhadap produktivitas tanaman dan keuntungan usaha tani. Kombinasi seluruh variabel digunakan sebagai atribut dalam proses pembentukan model Decision Tree.

Hasil Pembentukan Model Decision Tree

Model klasifikasi dibangun menggunakan algoritma Decision Tree dengan memanfaatkan data latih (*training data*) yang telah melalui proses *preprocessing*. Algoritma secara otomatis memilih atribut yang memiliki nilai *Information Gain* tertinggi sebagai akar (*root node*) dan membentuk cabang keputusan berdasarkan karakteristik data.

Hasil pembentukan model menunjukkan bahwa kombinasi variabel seperti pH tanah, kelembapan tanah, jenis bibit, dan curah hujan memberikan pengaruh yang signifikan terhadap klasifikasi kategori hasil panen dan estimasi keuntungan petani. Pohon keputusan yang dihasilkan juga menghasilkan aturan (*decision rules*) yang mudah dipahami sehingga dapat digunakan sebagai dasar dalam sistem pendukung keputusan.

Hasil Evaluasi Model

Evaluasi model dilakukan menggunakan Confusion Matrix untuk mengetahui kemampuan algoritma Decision Tree dalam melakukan klasifikasi. Berdasarkan hasil pengujian, diperoleh nilai evaluasi sebagaimana ditunjukkan pada Tabel 2.

Tabel 4. Hasil Evaluasi Model

No	Vriabel	Jenis
1	Accuracy	96,4%
2	Precision	95,8%
3	Recall	95,2%
4	F1-Score	95,5%

Pembahasan

Hasil penelitian menunjukkan bahwa algoritma *Decision Tree* mampu memberikan performa klasifikasi yang baik dalam menentukan rekomendasi jenis bibit, memprediksi kategori hasil panen, dan mengestimasi keuntungan petani. Model memperoleh nilai Accuracy sebesar 96,4%, Precision 95,8%, Recall 95,2%, dan F1-Score 95,5%, yang menunjukkan bahwa model memiliki tingkat klasifikasi yang tinggi. Selain menghasilkan akurasi yang baik, Decision Tree juga mampu menghasilkan aturan keputusan yang mudah dipahami sehingga mendukung proses pengambilan keputusan di bidang pertanian. Hasil penelitian ini sejalan dengan penelitian Suradiradja et al. (2022) dan Bishnoi & Hooda (2022) yang menyatakan bahwa algoritma Decision Tree efektif digunakan untuk klasifikasi data pertanian dan sistem pendukung keputusan.

Kesimpulan

Berdasarkan hasil penelitian, algoritma Decision Tree berhasil diimplementasikan untuk memberikan rekomendasi pemilihan bibit, memprediksi hasil panen, dan mengestimasi keuntungan petani berdasarkan parameter pertanian. Model menghasilkan nilai Accuracy 96,4%, Precision 95,8%, Recall 95,2%, dan F1-Score 95,5%, sehingga menunjukkan performa klasifikasi yang sangat baik. Sistem yang dikembangkan mampu membantu petani dalam mengambil keputusan budidaya secara lebih objektif dan berbasis data, sehingga berpotensi meningkatkan efektivitas pengelolaan usaha tani.

Saran

Penelitian selanjutnya disarankan menggunakan dataset yang lebih besar dan mencakup berbagai wilayah agar model memiliki kemampuan generalisasi yang lebih baik. Selain itu, performa Decision Tree dapat dibandingkan dengan algoritma lain, seperti Random Forest, Naïve Bayes, Support Vector Machine (SVM), dan XGBoost, serta dikembangkan menjadi aplikasi berbasis web atau mobile yang terintegrasi dengan teknologi Internet of Things (IoT).



DAFTAR PUSTAKA

- Aulia, N., Rizky, M., & Shaleh, C. (2025). Memprediksi jenis tanaman berdasarkan parameter lingkungan menggunakan Decision Tree dan Random Forest. *Jurnal Teknologi Informasi dan Terapan*.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Bishnoi, S., & Hooda, B. K. (2022). Decision Tree algorithms and their applicability in agriculture for classification. *Journal of Experimental Agriculture International*, 44(10), 65–75.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth International Group.
- Dillon, J. L., & Hardaker, J. B. (1993). *Farm Management Research for Small Farmer Development*. Food and Agriculture Organization (FAO).
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann.
- Islam, T., dkk. (2021). A deep neural network approach for crop selection and yield prediction. *arXiv preprint arXiv:2108.03320*.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70–90.
- Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in Plant Science*.
- Kusrini, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Andi Offset.
- Larose, D. T. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining* (2nd ed.). Wiley.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Power, D. J. (2002). *Decision Support Systems: Concepts and Resources for Managers*. Quorum Books.
- Prasetyo, T. A., dkk. (2026). Enhanced rice yield prediction in Indonesia with integrated climate and agricultural data using Decision Tree Regression. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.
- Roni, Abdullah, & Insani. (2025). Implementasi Naïve Bayes dan Decision Tree untuk klasifikasi jenis tanaman. *Jurnal Tekno Insentif*.
- Satria, A., Badri, R. M., & Safitri, I. (2023). Prediksi hasil panen tanaman pangan menggunakan metode Machine Learning. *Digital Transformation Technology Journal*.
- Soekartawi. (2002). *Analisis Usahatani*. UI Press.
- Suradiradja, K., Sitanggang, I. S., Abdullah, L., & Hermadi, I. (2022). Estimation of harvest time of forage sorghum using Decision Tree algorithm. *Tropical Animal Science Journal*, 45(2), 188–198.
- Tan, P. N., Steinbach, M., & Kumar, V. (2019). *Introduction to Data Mining* (2nd ed.). Pearson.
- Turban, E., Sharda, R., & Delen, D. (2011). *Decision Support and Business Intelligence Systems* (9th ed.). Pearson.

- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). Data Mining: Practical Machine Learning Tools and Techniques (4th ed.). Morgan Kaufmann.
- Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming: A review. *Agricultural Systems*, 153, 69–80.