

EXPLAINABLE MACHINE LEARNING UNTUK KLASIFIKASI KESESUAIAN LAHAN JAGUNG BERBASIS DATA TABULAR DI KABUPATEN SOPPENG

Ismail^{1*}, Andi Zulkifli Nusri², Sukma Ayu³

^{1,3}Program Studi Sistem Informasi, Universitas Lamappapoleonro

²Program Studi Teknik Informatika, Universitas Lamappapoleonro

E-mail: ismail@unipol.ac.id^{1*}, andizulkifli51@gmail.com², sukma.ayu05@gmail.com³

INFO ARTIKEL

Sejarah Artikel

Diterima : 08/06/2026

Direvisi : 16/06/2026

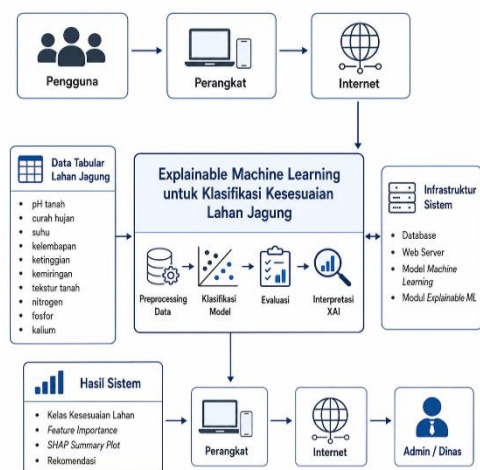
Diterbitkan : 17/06/2026

*Corresponding author

ismail@unipol.ac.id

DOI: 10.70247/jumistik.v5i1.318

GRAPHICAL ABSTRACT



ABSTRACT

Maize land suitability classification is important to support data-driven agricultural decision-making, particularly in selecting land with optimal physical, climatic, and nutrient conditions. This study aims to develop an Explainable Machine Learning model for classifying maize land suitability based on tabular data. The dataset consisted of 560 records from Soppeng Regency with ten input variables, including soil pH, rainfall, temperature, humidity, elevation, slope, soil texture, nitrogen, phosphorus, and potassium. The research stages included data preprocessing, feature encoding, numerical standardization, data splitting, model training, model evaluation, and model interpretation using Explainable Machine Learning. Five classification algorithms were evaluated: Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbor, and XGBoost. The results showed that Decision Tree achieved the best performance, with accuracy of 0.9911, precision of 0.9914, recall of 0.9911, and F1-score of 0.9911. Feature importance and SHAP analysis indicated that slope, soil pH, phosphorus, nitrogen, and potassium were the most influential factors.

Keywords: Classification; Decision Tree; Explainable Machine Learning; Land Suitability; Maize.

ABSTRAK

Klasifikasi kesesuaian lahan jagung penting untuk mendukung pengambilan keputusan pertanian berbasis data, terutama dalam menentukan lahan yang memiliki kondisi fisik, iklim, dan unsur hara yang sesuai. Penelitian ini bertujuan membangun model Explainable Machine Learning untuk klasifikasi kesesuaian lahan jagung berbasis data tabular. Dataset yang digunakan terdiri atas 560 data dari Kabupaten Soppeng dengan sepuluh variabel input, yaitu pH tanah, curah hujan, suhu, kelembapan, ketinggian, kemiringan lahan, tekstur tanah, nitrogen, fosfor, dan kalium. Tahapan penelitian meliputi preprocessing data, encoding fitur, standarisasi data numerik, pembagian data, pelatihan model, evaluasi model, dan interpretasi menggunakan Explainable Machine Learning. Algoritma yang diuji meliputi Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbor, dan XGBoost. Hasil pengujian menunjukkan bahwa Decision Tree memperoleh performa terbaik dengan accuracy 0,9911, precision 0,9914, recall 0,9911, dan F1-score 0,9911. Feature importance dan SHAP menunjukkan bahwa kemiringan lahan, pH tanah, fosfor, nitrogen, dan kalium menjadi faktor paling berpengaruh.

Kata kunci: Data Tabular; Decision Tree; Jagung; Kesesuaian Lahan; Machine Learning

© 2026 Penerbit STMIK Amika Soppeng. All rights reserved

PENDAHULUAN

Jagung merupakan salah satu komoditas pangan strategis yang memiliki peran penting dalam mendukung ketahanan pangan, memenuhi kebutuhan industri pakan, serta meningkatkan ekonomi masyarakat pertanian [1]. Sebagai tanaman pangan yang banyak dibudidayakan, jagung memiliki nilai ekonomi yang tinggi dan menjadi sumber pendapatan bagi petani di berbagai wilayah. Keberhasilan budidaya jagung tidak hanya ditentukan oleh teknik penanaman dan pemeliharaan, tetapi juga sangat dipengaruhi oleh kesesuaian lahan yang digunakan [2]. Lahan yang sesuai akan memberikan kondisi tumbuh yang lebih optimal, sedangkan lahan yang kurang sesuai dapat menyebabkan penurunan produktivitas, meningkatnya biaya budidaya, dan rendahnya efisiensi pemanfaatan sumber daya pertanian.

Kesesuaian lahan untuk tanaman jagung dapat dinilai berdasarkan berbagai parameter, seperti pH tanah, curah hujan, suhu, kelembapan, ketinggian tempat, kemiringan lahan, tekstur tanah, serta kandungan unsur hara seperti nitrogen, fosfor, dan kalium. Parameter-parameter tersebut saling berhubungan dan berpengaruh terhadap kemampuan lahan dalam mendukung pertumbuhan tanaman [3]. Dalam praktiknya, penilaian kesesuaian lahan sering kali tidak sederhana karena melibatkan banyak variabel dengan karakteristik yang berbeda. Oleh karena itu, diperlukan pendekatan analisis yang mampu mengolah data secara sistematis, objektif, dan konsisten agar hasil penilaian kesesuaian lahan dapat digunakan sebagai dasar pengambilan keputusan yang lebih tepat.

Permasalahan utama dalam penentuan kesesuaian lahan jagung adalah masih terbatasnya pemanfaatan metode berbasis data dalam proses klasifikasi lahan. Pada banyak kasus, penilaian kesesuaian lahan masih dilakukan secara manual, berbasis pengalaman, atau menggunakan pendekatan konvensional yang belum sepenuhnya mampu menangani kompleksitas data multivariabel. Padahal, data karakteristik lahan umumnya tersedia dalam bentuk tabular dan memiliki potensi besar untuk dianalisis menggunakan pendekatan Machine Learning. Data tabular tersebut dapat merepresentasikan kondisi lahan secara lebih terstruktur, sehingga dapat digunakan untuk membangun model klasifikasi yang mampu mengelompokkan lahan ke dalam kategori kesesuaian tertentu.

Machine Learning menjadi salah satu pendekatan yang relevan untuk digunakan dalam klasifikasi kesesuaian lahan karena mampu mempelajari pola dari data dan menghasilkan prediksi berdasarkan hubungan antarvariabel. Beberapa algoritma *Machine Learning* telah banyak digunakan dalam bidang pertanian untuk mendukung proses prediksi, klasifikasi, dan rekomendasi berbasis data [4]. Namun, penggunaan *Machine Learning* tidak hanya perlu

dilihat dari sisi akurasi model. Model yang memiliki performa tinggi tetap dapat menimbulkan persoalan apabila hasil keputusannya sulit dijelaskan, terutama ketika model tersebut digunakan oleh pengguna non-teknis seperti petani, penyuluh pertanian, atau pengambil kebijakan. Kondisi ini menunjukkan bahwa transparansi model menjadi aspek penting dalam penerapan *Machine Learning* di bidang pertanian [5].

Urgensi penelitian ini terletak pada kebutuhan terhadap model klasifikasi kesesuaian lahan jagung yang tidak hanya akurat, tetapi juga transparan dan dapat dijelaskan. Dalam konteks pengambilan keputusan pertanian, pengguna tidak cukup hanya mengetahui bahwa suatu lahan termasuk sangat sesuai, sesuai, cukup sesuai, atau tidak sesuai. Pengguna juga perlu memahami faktor-faktor apa saja yang paling memengaruhi hasil klasifikasi tersebut. Informasi semacam ini penting agar rekomendasi yang dihasilkan oleh model dapat diterima, dipahami, dan digunakan secara lebih bijak dalam perencanaan budidaya jagung. Oleh karena itu, pendekatan *Explainable Machine Learning* menjadi penting untuk menjembatani kemampuan prediktif model dengan kebutuhan interpretasi yang mudah dipahami oleh manusia.

Gap penelitian ini terletak pada masih terbatasnya penelitian yang secara khusus mengintegrasikan klasifikasi kesesuaian lahan jagung berbasis data tabular dengan pendekatan *Explainable Machine Learning*. Penelitian terdahulu pada bidang *Machine Learning* pertanian umumnya lebih berfokus pada peningkatan performa model, seperti *accuracy*, *precision*, *recall*, dan *F1-score* [6]. Meskipun metrik tersebut penting, hasil klasifikasi yang hanya menampilkan nilai performa belum cukup untuk menjelaskan alasan di balik keputusan model. Akibatnya, model berpotensi dipandang sebagai sistem yang bersifat *black box* dan kurang memberikan pemahaman terhadap faktor dominan yang menentukan kesesuaian lahan [7].

Sebagai solusi terhadap permasalahan tersebut, penelitian ini mengusulkan penerapan *Explainable Machine Learning* untuk klasifikasi kesesuaian lahan jagung berbasis data tabular [8]. Penelitian ini dirancang dengan tahapan pengumpulan data karakteristik lahan, preprocessing data, penerapan beberapa algoritma *Machine Learning*, evaluasi performa model, pemilihan model terbaik, serta interpretasi model menggunakan pendekatan *Explainable Machine Learning* [9]. Melalui pendekatan ini, model tidak hanya diarahkan untuk menghasilkan prediksi kelas kesesuaian lahan, tetapi juga untuk menjelaskan kontribusi setiap variabel terhadap keputusan klasifikasi. Dengan demikian, hasil penelitian diharapkan mampu memberikan informasi yang lebih bermakna dan dapat digunakan dalam proses pengambilan keputusan pertanian [10].

Tujuan penelitian ini adalah membangun model klasifikasi kesesuaian lahan jagung berbasis data tabular, membandingkan performa beberapa algoritma *Machine Learning*, menentukan model

terbaik berdasarkan metrik evaluasi, serta menginterpretasikan faktor dominan yang memengaruhi hasil klasifikasi menggunakan pendekatan *Explainable Machine Learning*. Kontribusi penelitian ini adalah menghadirkan model klasifikasi yang akurat, transparan, dan dapat dijelaskan, sehingga dapat menjadi dasar pengembangan sistem pendukung keputusan pertanian berbasis data. Secara praktis, hasil penelitian ini diharapkan dapat membantu petani, penyuluh pertanian, pemerintah daerah, dan pihak terkait dalam memahami potensi lahan jagung secara lebih objektif. Secara ilmiah, penelitian ini diharapkan dapat memperkaya kajian *Machine Learning* di bidang pertanian, khususnya pada penerapan *Explainable Machine Learning* untuk analisis kesesuaian lahan berbasis data tabular.

TINJAUAN PUSTAKA

Kesesuaian Lahan Jagung

Kesesuaian lahan jagung merupakan proses penilaian terhadap kemampuan suatu lahan dalam mendukung pertumbuhan dan produktivitas tanaman jagung berdasarkan karakteristik lingkungan, tanah, iklim, topografi, dan unsur hara [11]. Penilaian ini penting karena tanaman jagung membutuhkan kondisi tumbuh tertentu agar dapat berkembang secara optimal. Lahan yang memiliki karakteristik sesuai akan memberikan peluang lebih besar terhadap peningkatan hasil produksi, sedangkan lahan yang kurang sesuai dapat menyebabkan pertumbuhan tanaman tidak maksimal, meningkatnya biaya budidaya, serta rendahnya efisiensi pemanfaatan lahan. Oleh karena itu, analisis kesesuaian lahan menjadi bagian penting dalam perencanaan pertanian yang berorientasi pada produktivitas dan keberlanjutan [12].

Dalam menentukan kesesuaian lahan jagung, beberapa parameter yang umum digunakan meliputi pH tanah, curah hujan, suhu, kelembapan, ketinggian tempat, kemiringan lahan, tekstur tanah, serta kandungan unsur hara seperti nitrogen, fosfor, dan kalium. Parameter-parameter tersebut memiliki pengaruh yang saling berkaitan terhadap pertumbuhan tanaman. pH tanah berhubungan dengan ketersediaan unsur hara, curah hujan dan suhu memengaruhi proses fisiologis tanaman, kelembapan berkaitan dengan ketersediaan air, sedangkan kemiringan lahan dapat memengaruhi risiko erosi dan pengelolaan lahan. Dengan mempertimbangkan berbagai parameter tersebut, kesesuaian lahan dapat diklasifikasikan ke dalam beberapa kategori, seperti sangat sesuai, sesuai, cukup sesuai, dan tidak sesuai [13].

Data Tabular dalam Analisis Pertanian

Data tabular merupakan bentuk data yang tersusun dalam baris dan kolom, di mana setiap baris mewakili satu objek atau sampel pengamatan, sedangkan setiap kolom merepresentasikan atribut atau variabel tertentu. Dalam konteks analisis pertanian, data tabular dapat digunakan untuk

menyimpan informasi karakteristik lahan, seperti kondisi tanah, iklim, topografi, unsur hara, dan kelas kesesuaian lahan. Struktur data yang sistematis membuat data tabular mudah dikelola, dibersihkan, dianalisis, dan digunakan sebagai input dalam pemodelan *Machine Learning* [14].

Penggunaan data tabular dalam penelitian kesesuaian lahan jagung memiliki keunggulan karena mampu merepresentasikan banyak parameter lahan secara terstruktur. Setiap atribut dalam data dapat diperlakukan sebagai fitur yang digunakan oleh model untuk mempelajari pola hubungan antara karakteristik lahan dan kelas kesesuaian. Data numerik, seperti pH tanah, suhu, curah hujan, ketinggian, dan kandungan unsur hara, dapat diproses secara langsung atau melalui tahapan normalisasi. Sementara itu, data kategorik seperti tekstur tanah atau kelas kesesuaian lahan dapat diolah melalui proses *encoding* agar dapat dikenali oleh algoritma *Machine Learning*. Dengan demikian, data tabular menjadi bentuk data yang relevan dan efisien untuk mendukung klasifikasi kesesuaian lahan berbasis komputasi cerdas [15].

Machine Learning untuk Klasifikasi Kesesuaian Lahan

Machine Learning merupakan pendekatan komputasi yang memungkinkan sistem mempelajari pola dari data dan menggunakan pola tersebut untuk menghasilkan prediksi atau klasifikasi. Dalam klasifikasi kesesuaian lahan jagung, *Machine Learning* digunakan untuk mengolah data karakteristik lahan dan mengelompokkannya ke dalam kelas kesesuaian tertentu. Pendekatan ini dinilai relevan karena mampu menangani data multivariabel dan menemukan hubungan antarfitur yang tidak selalu mudah dikenali melalui analisis manual. Dengan memanfaatkan *Machine Learning*, proses penentuan kesesuaian lahan dapat dilakukan secara lebih objektif, cepat, dan konsisten [16].

Beberapa algoritma *Machine Learning* yang sering digunakan dalam tugas klasifikasi antara lain *Decision Tree*, *Random Forest*, *Support Vector Machine*, *K-Nearest Neighbor*, *Gradient Boosting*, dan *XGBoost*. *Decision Tree* memiliki keunggulan karena struktur keputusannya relatif mudah dipahami, sedangkan *Random Forest* menggabungkan banyak pohon keputusan untuk meningkatkan stabilitas dan performa model. *Support Vector Machine* bekerja dengan mencari batas pemisah terbaik antar kelas, sementara *K-Nearest Neighbor* melakukan klasifikasi berdasarkan kedekatan karakteristik antar data. *Gradient Boosting* dan *XGBoost* dikenal sebagai algoritma *ensemble* yang mampu menghasilkan performa tinggi melalui kombinasi beberapa model lemah menjadi model yang lebih kuat. Perbandingan beberapa algoritma tersebut diperlukan untuk menentukan model terbaik dalam mengklasifikasikan kesesuaian lahan jagung berbasis data tabular.

Evaluasi model *Machine Learning* dalam penelitian klasifikasi umumnya dilakukan

menggunakan metrik seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*. *Accuracy* digunakan untuk mengukur ketepatan model secara keseluruhan, *precision* menunjukkan ketepatan prediksi pada suatu kelas, *recall* menggambarkan kemampuan model dalam mengenali data pada kelas tertentu, sedangkan *F1-score* menunjukkan keseimbangan antara *precision* dan *recall*. *Confusion matrix* digunakan untuk melihat pola kesalahan klasifikasi antar kelas. Penggunaan beberapa metrik evaluasi diperlukan agar performa model dapat dianalisis secara lebih menyeluruh, terutama apabila distribusi kelas dalam dataset tidak seimbang [17].

Explainable Machine Learning

Explainable Machine Learning merupakan pendekatan yang digunakan untuk menjelaskan bagaimana model Machine Learning menghasilkan keputusan. Pendekatan ini menjadi penting karena beberapa model dengan performa tinggi sering kali bersifat kompleks dan sulit dipahami oleh pengguna, sehingga dikenal sebagai model *black box*. Dalam konteks pertanian, keterbukaan model menjadi hal yang penting karena hasil klasifikasi tidak hanya digunakan oleh peneliti atau pengembang sistem, tetapi juga oleh petani, penyuluh pertanian, pemerintah daerah, dan pengambil kebijakan. Oleh karena itu, model yang baik tidak cukup hanya menghasilkan prediksi yang akurat, tetapi juga perlu memberikan penjelasan yang dapat dipahami mengenai alasan di balik prediksi tersebut [18].

Metode *Explainable Machine Learning* seperti *SHAP*, *LIME*, *feature importance*, dan *permutation importance* dapat digunakan untuk mengidentifikasi kontribusi setiap fitur terhadap hasil klasifikasi. *SHAP* dapat menjelaskan pengaruh masing-masing variabel terhadap prediksi model secara global maupun lokal, sedangkan *LIME* dapat memberikan penjelasan terhadap prediksi pada sampel tertentu. *Feature importance* membantu menunjukkan fitur yang paling dominan dalam proses klasifikasi, sementara *permutation importance* mengukur perubahan performa model ketika suatu fitur diacak atau dihilangkan. Dalam penelitian kesesuaian lahan jagung, pendekatan ini dapat digunakan untuk mengetahui apakah faktor seperti pH tanah, curah hujan, suhu, unsur hara, atau kemiringan lahan memiliki pengaruh besar terhadap kelas kesesuaian lahan [19].

Penerapan *Explainable Machine Learning* dalam klasifikasi kesesuaian lahan jagung memberikan nilai tambah karena model tidak hanya menghasilkan kelas prediksi, tetapi juga menjelaskan faktor dominan yang memengaruhi keputusan tersebut. Penjelasan ini dapat meningkatkan kepercayaan pengguna terhadap hasil model dan membantu proses pengambilan keputusan pertanian secara lebih transparan. Dengan demikian, *Explainable Machine Learning* menjadi pendekatan yang relevan untuk menggabungkan kekuatan prediktif *Machine*

Learning dengan kebutuhan interpretasi dalam praktik pertanian berbasis data [20].

State Of The Art

Adapun perbedaan (kebaharuan) penelitian ini dapat dilihat pada tabel 1 berikut:

Tabel 1. Penelitian terdahulu

No	Judul	Metode	Hasil	Perbedaan dengan Penelitian Sekarang
1	<i>Evaluation of Maximum Entropy (Maxent) Machine Learning Model to Assess Relationships between Climate and Corn Suitability</i> [5]	Maximum Entropy (Maxent) berbasis variabel iklim	Hasil penelitian menunjukkan bahwa Maxent dapat digunakan untuk memodelkan kesesuaian lahan tanaman jagung berdasarkan variabel	penelitian sekarang menggunakan data tabular multivariabel serta menambahkan Explainable Machine Learning untuk menjelaskan faktor dominan
2	<i>Enhancing Crop Recommendation Systems with Explainable Artificial Intelligence: A Study on Agricultural Decision-Making</i> [19]	Machine Learning dan Explainable Artificial Intelligence	XAI dapat meningkatkan transparansi model dan membantu menjelaskan alasan di balik rekomendasi pertanian	penelitian sekarang berfokus pada klasifikasi kesesuaian lahan jagung berbasis data tabular
3	<i>Cropland Suitability Prediction Method Based on Biophysical Variables from Copernicus Data and Machine Learning</i> [15]	Machine Learning berbasis variabel biofisik dari data Copernicus	Hasil penelitian menunjukkan bahwa Machine Learning dapat menjadi alternatif analisis kesesuaian lahan berbasis GIS.	penelitian sekarang menggunakan data tabular kesesuaian lahan jagung

Sumber: Hasil telaah penelitian terdahulu, 2026

METODOLOGI PENELITIAN

Jenis Data dan Lokasi Penelitian

Penelitian ini menggunakan pendekatan kuantitatif berbasis eksperimen *Machine Learning*. Jenis data yang digunakan adalah data tabular yang memuat karakteristik lahan untuk tanaman jagung. Data tabular dipilih karena mampu menyajikan berbagai parameter lahan secara terstruktur dalam bentuk baris dan kolom, sehingga dapat diproses secara sistematis oleh algoritma Machine Learning. Setiap baris data merepresentasikan satu objek atau sampel lahan, sedangkan setiap kolom

merepresentasikan atribut lahan yang digunakan sebagai variabel input dalam proses klasifikasi.

Data penelitian dapat diperoleh dari data primer dan data sekunder. Data primer dapat dikumpulkan melalui observasi lapangan, pengukuran langsung kondisi lahan, wawancara dengan petani atau penyuluh pertanian, serta dokumentasi lokasi penelitian. Sementara itu, data sekunder dapat diperoleh dari instansi terkait, seperti Dinas Pertanian, Badan Pusat Statistik, serta sumber data publik lain yang relevan. Variabel yang digunakan dalam penelitian ini meliputi pH tanah, curah hujan, suhu, kelembapan, ketinggian tempat, kemiringan lahan, tekstur tanah, nitrogen, fosfor, kalium, dan kelas kesesuaian lahan sebagai variabel target.

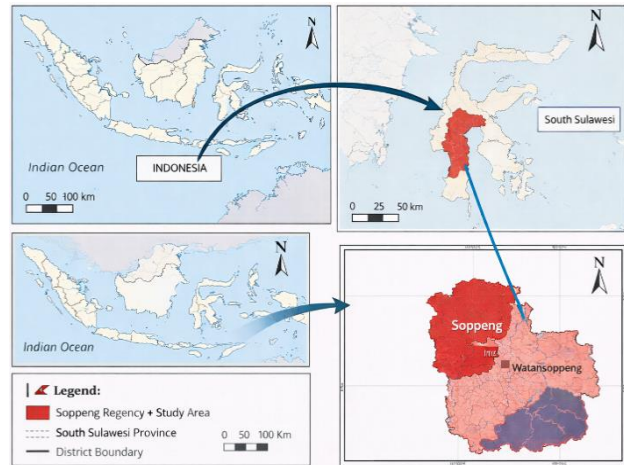
Tabel 2. Jenis Data dan Sumber Data Penelitian

N o.	Jenis Data	Variabel/Informasi	Sumber Data	Keterangan
1	Data tanah	pH tanah, tekstur tanah, nitrogen, fosfor, dan kalium	Observasi lapangan, Dinas Pertanian	Digunakan sebagai variabel input untuk menilai karakteristik kesuburan tanah
2	Data iklim	Curah hujan, suhu, dan kelembapan	dataset publik yang relevan	Digunakan untuk mengetahui kondisi lingkungan yang memengaruhi pertumbuhan jagung
3	Data topografi	Ketinggian tempat dan kemiringan lahan	observasi lapangan, dan dataset publik	Digunakan untuk menilai kondisi fisik lahan terhadap kesesuaian budidaya jagung
4	Data administrasi wilayah	Profil Kabupaten Soppeng, kecamatan, desa/kelurahan, dan batas wilayah	BPS Kabupaten Soppeng	Digunakan untuk mendeskripsikan lokasi dan ruang lingkup wilayah penelitian
5	Data kelas kesesuaian lahan	Sangat sesuai, sesuai, cukup sesuai, dan tidak sesuai	Pedoman kesesuaian lahan, validasi ahli,	Digunakan sebagai variabel target dalam proses klasifikasi Machine Learning

Sumber: Hasil Pengumpulan Data penelitian, 2026

Lokasi penelitian ini berada di Kabupaten Soppeng, Provinsi Sulawesi Selatan. Kabupaten Soppeng dipilih sebagai lokasi penelitian karena memiliki potensi pertanian dan variasi karakteristik lahan yang relevan

untuk pengembangan tanaman jagung. Secara geografis, wilayah Kabupaten Soppeng memiliki kondisi topografi yang beragam, mulai dari dataran hingga perbukitan. Keragaman kondisi wilayah tersebut menjadikan Kabupaten Soppeng relevan sebagai lokasi kajian klasifikasi kesesuaian lahan jagung berbasis data tabular. Berikut peta lokasi penelitian :

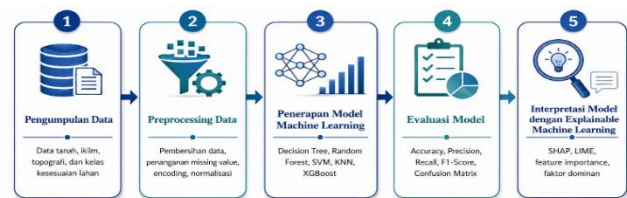


Gambar 1. Lokasi Penelitian

(Sumber : Website Pemerintahan Kabupaten Soppeng, 2026)

Tahapan Penelitian

Tahapan penelitian dirancang secara sistematis agar proses pemodelan dapat dilakukan secara terukur, transparan, dan dapat direplikasi. Secara umum, penelitian ini terdiri atas lima tahapan utama, yaitu pengumpulan data, preprocessing data, penerapan model *Machine Learning*, evaluasi model, serta interpretasi model menggunakan pendekatan *Explainable Machine Learning* [21]. Secara umum tahapan penelitian dapat dilihat pada gambar 2 berikut:



Gambar 2. Tahapan Penelitian

(Sumber : Hasil Rancangan Penelitian, 2026)

Gambar 2 tersebut menjelaskan alur tahapan penelitian yang dilakukan secara sistematis mulai dari pengumpulan data, yaitu data tanah, iklim, topografi, dan kelas kesesuaian lahan; dilanjutkan dengan *preprocessing* data untuk membersihkan data, menangani nilai kosong, melakukan *encoding*, dan normalisasi [22]. Selanjutnya, data yang telah siap digunakan diproses melalui penerapan model *Machine Learning* dengan beberapa algoritma seperti *Decision Tree*, *Random Forest*, *SVM*, *KNN*, dan *XGBoost* [23]. Setiap model kemudian dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix* untuk menentukan model

terbaik. Tahap akhir adalah interpretasi model menggunakan *Explainable Machine Learning* melalui metode seperti *SHAP*, *LIME*, dan *feature importance* guna mengetahui faktor dominan yang memengaruhi klasifikasi kesesuaian lahan jagung, sehingga hasil model lebih transparan dan dapat digunakan sebagai dasar pengambilan keputusan pertanian berbasis data [24].

HASIL DAN PEMBAHASAN

Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini adalah `dataset_lahan_jagung.csv` yang disusun dalam format CSV/data tabular. Dataset ini terdiri atas 560 data dengan 15 kolom, yang mencakup identitas sampel, informasi wilayah, variabel karakteristik lahan, serta kelas kesesuaian lahan sebagai target klasifikasi. Penelitian ini menggunakan 10 variabel input utama yang merepresentasikan karakteristik tanah, iklim, topografi, dan unsur hara, yaitu pH tanah, curah hujan, suhu, kelembapan, ketinggian tempat, kemiringan lahan, tekstur tanah, nitrogen, fosfor, dan kalium. Data penelitian diperoleh dari Dinas Tanaman Pangan, Holtikultura dan Observasi Lapangan, serta wawancara dengan petani dan penyuluh pertanian.

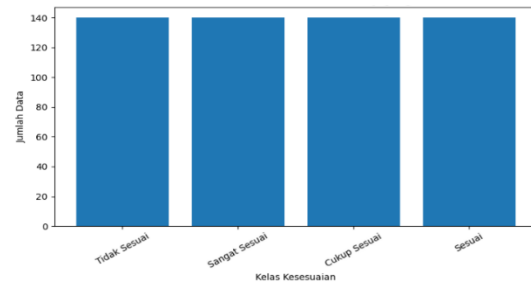
Data penelitian berlokasi di Kabupaten Soppeng dengan cakupan 8 kecamatan. Variabel target yang digunakan adalah `kelas_kesesuaian`, yang terdiri atas 4 kelas, yaitu Tidak Sesuai, Cukup Sesuai, Sesuai, dan Sangat Sesuai. Kelas target tersebut digunakan sebagai label dalam proses klasifikasi menggunakan algoritma Machine Learning. Berdasarkan hasil pemeriksaan awal, dataset tidak memiliki missing value dan tidak ditemukan data duplikat, sehingga data dinilai layak untuk dilanjutkan ke tahap preprocessing dan pemodelan. Ringkasan dataset dapat dilihat pada tabel dibawah ini:

Tabel 3. Ringkasan Dataset

No.	Komponen Dataset	Keterangan
1	Nama dataset	<code>dataset_lahan_jagung.csv</code>
2	Jumlah data	560
3	Jumlah kolom	15
4	Jumlah variabel input	10
5	Kabupaten	Soppeng
6	Jumlah kecamatan	8
7	Variabel target	<code>kelas_kesesuaian</code>
8	Jumlah kelas target	4
9	Kelas target	Tidak Sesuai, Sangat Sesuai, Cukup Sesuai, Sesuai
10	Missing value	0
11	Data duplikat	0
12	Format data	CSV/data tabular

Sumber: Hasil pengolahan dataset penelitian, 2026.

Berdasarkan tabel 3, dataset memiliki struktur yang sesuai untuk digunakan dalam penelitian klasifikasi kesesuaian lahan jagung berbasis data tabular. Tidak adanya missing value dan data duplikat menunjukkan bahwa dataset memiliki kualitas awal yang baik. Selain itu, jumlah data yang memadai dan keberadaan variabel input yang relevan mendukung proses pelatihan model Machine Learning serta interpretasi faktor dominan menggunakan pendekatan *Explainable Machine Learning*.



Gambar 3. Distribusi Kelas

(Sumber : Hasil Proses data dengan Python, 2026)

Gambar 3 menunjukkan distribusi kelas kesesuaian lahan jagung pada dataset yang digunakan dalam penelitian. Dataset terdiri atas empat kelas, yaitu Tidak Sesuai, Sangat Sesuai, Cukup Sesuai, dan Sesuai, dengan masing-masing kelas memiliki jumlah data sebanyak 140 data. Distribusi yang seimbang ini menunjukkan bahwa dataset tidak mengalami ketimpangan kelas, sehingga proses pelatihan model *Machine Learning* dapat dilakukan secara lebih adil terhadap seluruh kategori kesesuaian lahan.

Hasil Preprocessing Data

Tahapan *preprocessing* data dilakukan untuk memastikan dataset memiliki kualitas yang baik sebelum digunakan dalam proses pelatihan model Machine Learning. *Preprocessing* menjadi tahap penting karena kualitas data sangat memengaruhi kemampuan model dalam mempelajari pola dan menghasilkan klasifikasi yang tepat.

Tahap pertama *preprocessing* adalah pemeriksaan kualitas awal dataset, yaitu dengan mengecek keberadaan *missing value* dan data duplikat. Berdasarkan hasil pemeriksaan, dataset tidak memiliki *missing value* dan tidak ditemukan data duplikat.

Tabel 5. Pemeriksaan Kualitas Awal Dataset

No.	Komponen Pemeriksaan	Sebelum Penanganan	Setelah Penanganan	Keterangan
1	Missing value	0	0	Tidak ditemukan data kosong
2	Data duplikat	0	0	Tidak

ditemukan
data
ganda

Sumber: Hasil preprocessing dataset penelitian, 2026.

Tahap berikutnya adalah pemilihan fitur dan target. Variabel yang tidak digunakan sebagai fitur model, seperti *id_sampel*, kabupaten, kelas_kesesuaian, dan kelas_encoded, dipisahkan dari data input. Variabel target yang digunakan dalam proses klasifikasi adalah kelas_encoded, sedangkan kelas_kesesuaian digunakan sebagai label kategorik untuk memudahkan interpretasi hasil. Fitur yang digunakan dalam model terdiri atas variabel numerik dan kategorik. Variabel numerik meliputi pH tanah, curah hujan, suhu, kelembapan, ketinggian tempat, kemiringan lahan, nitrogen, fosfor, dan kalium. Sementara itu, variabel kategorik terdiri atas kecamatan dan tekstur tanah. Jenis Fitur yang Digunakan dalam Model dapat dilihat pada tabel 6 berikut:

Tabel 6. Jenis Fitur yang Digunakan dalam Model

No.	Jenis Fitur	Nama Fitur	Jumlah Fitur
1	Numerik	pH tanah, curah hujan, suhu, kelembapan, ketinggian, kemiringan, nitrogen, fosfor, kalium	9
2	Kategorik	kecamatan, tekstur tanah	2
3	Target	kelas_encoded	1

Sumber: Hasil preprocessing dataset penelitian, 2026

Pada tahap encoding, variabel target kelas_kesesuaian dikonversi ke dalam bentuk numerik agar dapat diproses oleh algoritma Machine Learning. Kelas "Tidak Sesuai" diberi nilai 0, "Cukup Sesuai" diberi nilai 1, "Sesuai" diberi nilai 2, dan "Sangat Sesuai" diberi nilai 3. Proses encoding ini tidak mengubah makna kategori, tetapi hanya mengubah bentuk representasi kelas agar sesuai dengan kebutuhan pemodelan klasifikasi.

Tabel 7. Hasil Encoding Kelas Target

No.	Kelas Encoded	Kelas Kesesuaian
1	0	Tidak Sesuai
2	1	Cukup Sesuai
3	2	Sesuai
4	3	Sangat Sesuai

Sumber: Hasil preprocessing dataset penelitian, 2026
Dataset kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20. Dari total 560 data, sebanyak 448 data digunakan sebagai data latih, sedangkan 112 data digunakan sebagai data uji. Pembagian data dilakukan secara stratified agar proporsi setiap kelas pada data latih dan data uji tetap seimbang.

Hasil pembagian menunjukkan bahwa setiap kelas memiliki 112 data pada data latih dan 28 data pada data uji. Komposisi ini penting agar model dapat mempelajari seluruh kelas secara adil dan proses evaluasi dapat menggambarkan kemampuan model terhadap semua kategori kesesuaian lahan. Distribusi Kelas pada Data Latih dan Data Uji diuraikan pada tabel berikut:

Tabel 8. Distribusi Kelas pada Data Latih dan Data Uji

No.	Kelas Encode d	Kelas Kesesuaian	Data Latih	Data Uji	Total
1	0	Tidak Sesuai	112	28	140
2	1	Cukup Sesuai	112	28	140
3	2	Sesuai	112	28	140
4	3	Sangat Sesuai	112	28	140
Total			448	112	560

Sumber: Hasil preprocessing dataset penelitian, 2026

Secara keseluruhan, hasil preprocessing menunjukkan bahwa dataset telah siap digunakan dalam proses pelatihan model Machine Learning. Tidak ditemukannya missing value dan data duplikat menunjukkan bahwa dataset memiliki kualitas awal yang baik. Proses *encoding* dan standarisasi berhasil mengubah seluruh fitur ke dalam bentuk yang dapat diproses oleh algoritma klasifikasi. Selain itu, pembagian data latih dan data uji yang seimbang pada setiap kelas mendukung proses pelatihan dan evaluasi model secara lebih objektif.

Tabel 10. Summary Hasil Preprocessing Data

No.	Tahapan Preprocessing	Hasil
1	Pemeriksaan missing value	0 missing value ditemukan
2	Penanganan missing value	0 missing value setelah penanganan
3	Pemeriksaan data duplikat	0 data duplikat ditemukan
4	Penghapusan data duplikat	0 data duplikat setelah penanganan
5	Encoding data kategorik	2 fitur kategorik diubah menggunakan OneHotEncoder
6	Normalisasi/standarisasi data numerik	9 fitur numerik distandarisasi menggunakan StandardScaler
7	Pembagian data latih dan data uji	448 data latih dan 112 data uji
8	Jumlah fitur setelah preprocessing	25 fitur digunakan dalam model

Sumber: Hasil preprocessing dataset penelitian, 2026

Hasil Pelatihan dan Analisis Model Terbaik Machine Learning

Tahap pelatihan model dilakukan setelah dataset melalui proses *preprocessing*, *encoding*, *standardisasi*, serta pembagian data menjadi data latih dan data uji. Pada tahap ini, sebanyak 448 data latih dengan 25 fitur hasil *preprocessing* digunakan untuk membangun model klasifikasi kesesuaian lahan jagung. Algoritma yang digunakan dalam penelitian ini terdiri atas *Decision Tree*, *Random Forest*, *Support Vector Machine*, *K-Nearest Neighbor*, dan *XGBoost*. Kelima algoritma tersebut dipilih karena memiliki karakteristik yang berbeda dalam mempelajari pola data tabular, sehingga dapat memberikan perbandingan performa yang lebih komprehensif.

Hasil pelatihan menunjukkan bahwa seluruh model berhasil dilatih dengan baik. Model *Random Forest* dan *XGBoost* memperoleh nilai *training accuracy* dan *training F1-score* sebesar 1,000000, yang menunjukkan bahwa kedua model mampu mempelajari pola data latih secara sangat optimal. Model *Decision Tree* juga menunjukkan hasil pelatihan yang sangat baik dengan nilai *training accuracy* sebesar 0,993304 dan *training F1-score* sebesar 0,993303. Sementara itu, *Support Vector Machine* memperoleh *training accuracy* sebesar 0,984375 dan *training F1-score* sebesar 0,984375, sedangkan *K-Nearest Neighbor* memperoleh *training accuracy* sebesar 0,959821 dan *training F1-score* sebesar 0,959615. Dari sisi waktu pelatihan, *K-Nearest Neighbor* memiliki waktu pelatihan tercepat, sedangkan *Random Forest* membutuhkan waktu pelatihan paling lama.

Tabel 11. Hasil Pelatihan Model Machine Learning

No.	Model	Jumlah Data Latih	Jumlah Fitur	Training Accuracy	Training F1-score	Waktu Pelatihan (detik)	Status
1	Decision Tree	448	25	0,993304	0,993303	0,013465	Berhasil dilatih
2	Random Forest	448	25	1,000000	1,000000	0,856967	Berhasil dilatih
3	SVM	448	25	0,984375	0,984375	0,040972	Berhasil dilatih
4	KNN	448	25	0,959821	0,959615	0,000792	Berhasil dilatih
5	XGBoost	448	25	1,000000	1,000000	0,350838	Berhasil dilatih

Sumber: Hasil pelatihan model penelitian, 2026

Model yang memiliki performa sangat tinggi pada data latih belum tentu memiliki kemampuan generalisasi yang sama baiknya pada data uji. Oleh karena itu, setiap model dievaluasi menggunakan data uji dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Dalam penelitian ini, *F1-score* digunakan sebagai dasar utama dalam menentukan model terbaik karena metrik ini mampu menggambarkan

keseimbangan antara *precision* dan *recall* pada kasus klasifikasi multikelas.

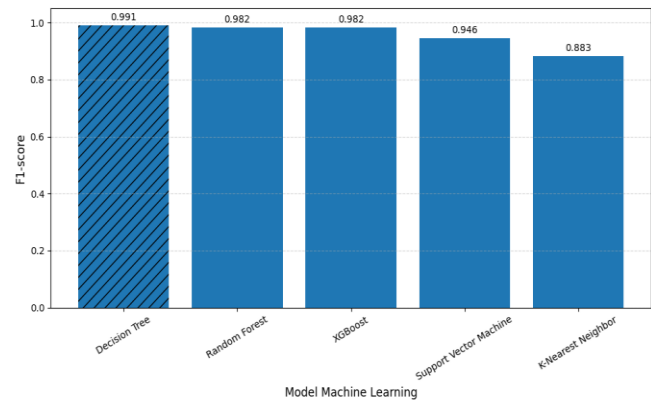
Hasil pengujian pada data uji menunjukkan bahwa model *Decision Tree* memperoleh performa terbaik dengan nilai *accuracy* sebesar 0,991071, *precision* sebesar 0,991379, *recall* sebesar 0,991071, dan *F1-score* sebesar 0,991069. Model *Random Forest* dan *XGBoost* memperoleh nilai *F1-score* yang sama, yaitu 0,982288, sedangkan *Support Vector Machine* memperoleh *F1-score* sebesar 0,946429. Model *K-Nearest Neighbor* memperoleh performa terendah dengan nilai *F1-score* sebesar 0,882996.

Tabel 12. Perbandingan Performa Model Machine Learning

No.	Model	Accuracy	Precision	Recall	F1-score
1	Decision Tree	0,991071	0,991379	0,991071	0,991069
2	Random Forest	0,982143	0,983333	0,982143	0,982288
3	XGBoost	0,982143	0,983333	0,982143	0,982288
4	Support Vector Machine	0,946429	0,946429	0,946429	0,946429
5	K-Nearest Neighbor	0,883929	0,888340	0,883929	0,882996

Sumber: Hasil pengujian model penelitian, 2026.

Selanjutnya visualisasi hasil Analisis Model Terbaik Berdasarkan *F1-score* dapat dilihat pada gambar berikut:



Gambar 4. Analisis Model Terbaik Berdasarkan *F1-score* (Sumber: Hasil pengujian model penelitian, 2026)

Berdasarkan Gambar 4, model *Decision Tree* memiliki nilai *F1-score* tertinggi dibandingkan model lainnya, yaitu sebesar 0,9911. *Random Forest* dan *XGBoost* berada pada posisi berikutnya dengan nilai *F1-score* sebesar 0,9823, diikuti oleh *Support Vector Machine* sebesar 0,9464, dan *K-Nearest Neighbor* sebesar 0,8830. Grafik tersebut memperjelas bahwa *Decision Tree* memiliki performa paling optimal dalam menjaga keseimbangan antara ketepatan prediksi dan kemampuan mengenali setiap kelas kesesuaian lahan jagung. Dengan demikian, *Decision Tree* dipilih sebagai model terbaik untuk dianalisis lebih lanjut.

Untuk melihat performa model terbaik secara lebih rinci, dilakukan analisis *classification report* pada model *Decision Tree*. Hasil *classification report* menunjukkan bahwa model mampu mengklasifikasikan sebagian besar kelas dengan sangat baik. Kelas Tidak Sesuai dan Sangat Sesuai memperoleh nilai *precision*, *recall*, dan *F1-score* sebesar 1,00, yang berarti seluruh data pada kedua

kelas tersebut berhasil dikenali secara tepat. Kelas Cukup Sesuai memperoleh *precision* sebesar 1,00, *recall* sebesar 0,96, dan *F1-score* sebesar 0,98, sedangkan kelas Sesuai memperoleh *precision* sebesar 0,97, *recall* sebesar 1,00, dan *F1-score* sebesar 0,98.

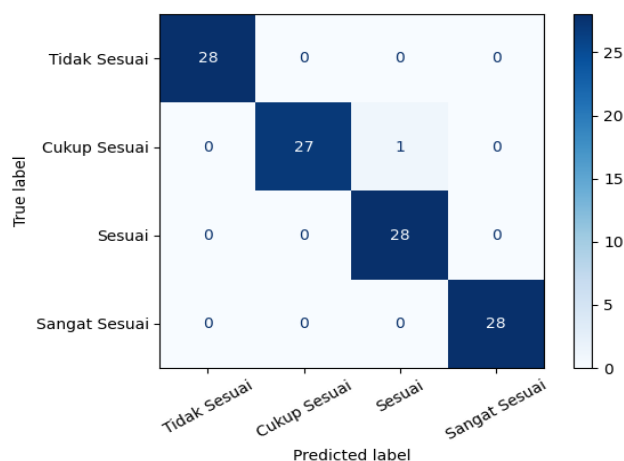
Tabel 13. *Classification Report* Model Terbaik

No.	Kelas Kesesuaian	Precision	Recall	F1- score	Support
1	Tidak Sesuai	1,00	1,00	1,00	28
2	Cukup Sesuai	1,00	0,96	0,98	28
3	Sesuai	0,97	1,00	0,98	28
4	Sangat Sesuai	1,00	1,00	1,00	28
	Accuracy			0,99	112
	Macro	0,99	0,99	0,99	112
	Average				
	Weighted	0,99	0,99	0,99	112
	Average				

Sumber: Hasil pengujian model penelitian, 2026.

Berdasarkan Tabel 13, nilai *macro average* dan *weighted average* masing-masing berada pada angka 0,99. Nilai ini menunjukkan bahwa model *Decision Tree* memiliki performa yang stabil pada seluruh kelas kesesuaian lahan. Karena setiap kelas memiliki jumlah data uji yang sama, yaitu 28 data, nilai *macro average* dapat digunakan sebagai indikator yang representatif untuk menilai keseimbangan performa model.

Selanjutnya, analisis *confusion matrix* dilakukan untuk mengetahui jumlah prediksi benar dan salah pada setiap kelas. Hasil *confusion matrix* menunjukkan bahwa model *Decision Tree* mampu mengklasifikasikan kelas Tidak Sesuai, Sesuai, dan Sangat Sesuai secara sempurna, masing-masing sebanyak 28 data. Pada kelas Cukup Sesuai, terdapat 27 data yang diklasifikasikan dengan benar dan 1 data yang diprediksi sebagai kelas Sesuai.



Gambar 5. *Confusion Matrix* Model Terbaik
(**Sumber:** Hasil pengujian model penelitian, 2026)

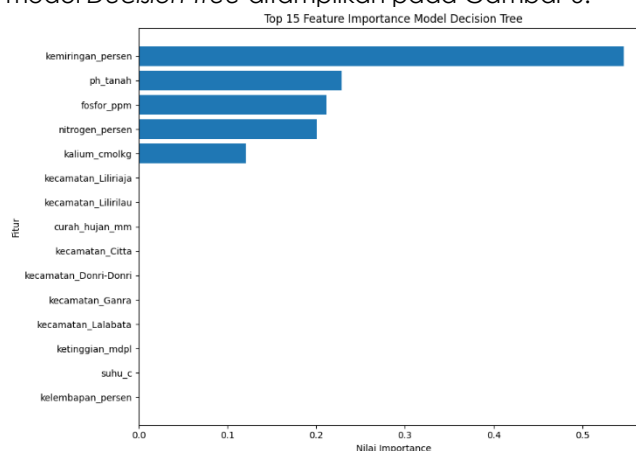
Berdasarkan hasil *confusion matrix*, hanya terdapat 1 kesalahan prediksi dari total 112 data uji. Kesalahan tersebut terjadi pada kelas Cukup Sesuai yang diprediksi sebagai kelas Sesuai. Hal ini dapat

terjadi karena karakteristik lahan pada kelas Cukup Sesuai dan Sesuai memiliki kedekatan nilai pada beberapa parameter, sehingga model mengalami kesulitan dalam membedakan satu sampel yang berada pada batas kedua kelas tersebut. Secara umum, hasil ini menunjukkan bahwa model *Decision Tree* memiliki kemampuan generalisasi yang sangat baik dalam membedakan kelas kesesuaian lahan jagung berbasis data tabular.

Interpretasi Model Menggunakan *Explainable Machine Learning*

Interpretasi model dilakukan untuk mengetahui faktor-faktor yang paling berpengaruh terhadap hasil klasifikasi kesesuaian lahan jagung. Pada tahap ini, model terbaik yang telah diperoleh dari hasil evaluasi, yaitu *Decision Tree*, dianalisis menggunakan pendekatan *Explainable Machine Learning*. Pendekatan ini bertujuan agar hasil klasifikasi tidak hanya dilihat dari nilai performa model, tetapi juga dapat dipahami melalui kontribusi setiap variabel terhadap keputusan model.

Metode interpretasi pertama yang digunakan adalah *feature importance*. *Feature importance* digunakan untuk melihat tingkat pengaruh setiap fitur terhadap proses pengambilan keputusan model. Semakin tinggi nilai importance suatu fitur, semakin besar kontribusi fitur tersebut dalam menentukan kelas kesesuaian lahan jagung. Hasil *feature importance* model *Decision Tree* ditampilkan pada Gambar 6.

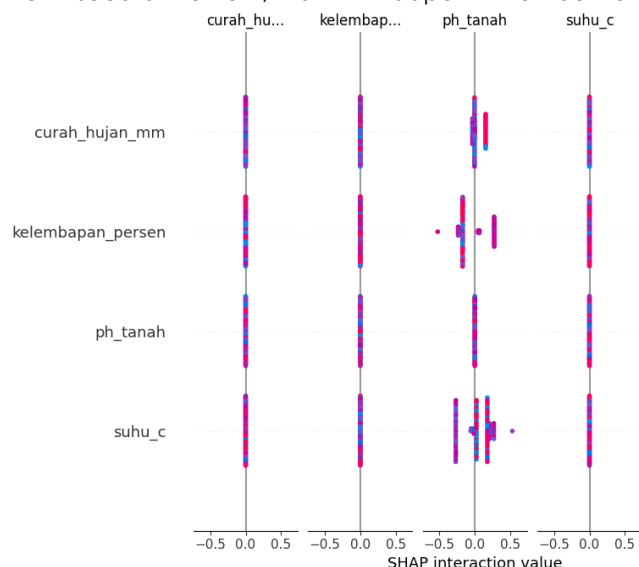


Gambar 6. *Feature Importance* Model *Decision Tree*
(**Sumber:** Hasil interpretasi model penelitian, 2026)

Berdasarkan Gambar 6, fitur yang memiliki pengaruh paling besar terhadap klasifikasi kesesuaian lahan jagung adalah kemiringan_persen. Hal ini menunjukkan bahwa kemiringan lahan menjadi faktor dominan dalam menentukan tingkat kesesuaian lahan pada dataset penelitian. Secara praktis, kemiringan lahan berhubungan dengan kemudahan pengolahan lahan, risiko erosi, pengelolaan air, dan efektivitas budidaya. Lahan dengan tingkat kemiringan yang tinggi cenderung memiliki risiko pengelolaan yang lebih besar, sehingga dapat memengaruhi kelas kesesuaian lahan untuk tanaman jagung. Fitur berikutnya yang juga memiliki pengaruh

tinggi adalah *ph_tanah*, *fosfor_ppm*, *nitrogen_persen*, dan *kalium_cmolk*. Temuan ini menunjukkan bahwa selain faktor topografi, kondisi kimia dan kesuburan tanah juga memiliki peran penting dalam menentukan kesesuaian lahan jagung. pH tanah berkaitan dengan ketersediaan unsur hara, sedangkan nitrogen, fosfor, dan kalium merupakan unsur hara utama yang mendukung pertumbuhan tanaman. Nitrogen berperan dalam pertumbuhan vegetatif, fosfor mendukung perkembangan akar dan pembentukan energi tanaman, sedangkan kalium berperan dalam ketahanan tanaman dan kualitas hasil. Dengan demikian, model menangkap bahwa kesesuaian lahan jagung sangat dipengaruhi oleh kombinasi antara kondisi fisik lahan dan kualitas unsur hara tanah.

Untuk memperkuat interpretasi model, penelitian ini juga menggunakan *SHAP Summary Plot*. *SHAP* digunakan untuk melihat kontribusi fitur secara lebih rinci, termasuk arah dan besarnya pengaruh fitur terhadap keputusan model. Berbeda dengan *feature importance* yang menunjukkan tingkat kepentingan fitur secara umum, *SHAP* dapat memberikan



Gambar 7. *SHAP Summary Plot Model Decision Tree* (Sumber: Hasil interpretasi model penelitian, 2026)

Berdasarkan Gambar 7, fitur *ph_tanah* menunjukkan sebaran nilai *SHAP* yang cukup bervariasi, terutama ketika berinteraksi dengan variabel seperti *curah_hujan_mm*, *kelembapan_persen*, dan *suhu_c*. Hal ini menunjukkan bahwa pH tanah tidak hanya berperan sebagai faktor tunggal, tetapi juga memiliki hubungan dengan kondisi lingkungan tumbuh lainnya dalam memengaruhi klasifikasi kesesuaian lahan jagung. Variasi nilai *SHAP* pada fitur pH tanah mengindikasikan bahwa perubahan tingkat keasaman tanah dapat memberikan pengaruh terhadap keputusan model, khususnya ketika dikombinasikan dengan kondisi iklim dan kelembapan.

Hasil interpretasi melalui *feature importance* dan *SHAP* menunjukkan bahwa faktor dominan dalam

klasifikasi kesesuaian lahan jagung meliputi kemiringan lahan, pH tanah, serta unsur hara seperti fosfor, nitrogen, dan kalium. Temuan ini selaras dengan konsep kesesuaian lahan yang menempatkan kondisi fisik lahan dan kualitas tanah sebagai faktor penting dalam mendukung pertumbuhan tanaman jagung.

Pembahasan

Hasil penelitian menunjukkan bahwa pendekatan *Machine Learning* mampu digunakan secara efektif untuk klasifikasi kesesuaian lahan jagung berbasis data tabular. Dari lima algoritma yang diuji, model *Decision Tree* memperoleh performa terbaik dengan nilai *accuracy* sebesar 0,9911, *precision* sebesar 0,9914, *recall* sebesar 0,9911, dan *F1-score* sebesar 0,9911. Hasil ini menunjukkan bahwa *Decision Tree* memiliki kemampuan yang sangat baik dalam mengenali pola hubungan antara karakteristik lahan dan kelas kesesuaian lahan jagung. Berdasarkan *confusion matrix*, model hanya menghasilkan satu kesalahan prediksi dari 112 data uji, yaitu pada kelas Cukup Sesuai yang diprediksi sebagai Sesuai. Kesalahan tersebut dapat terjadi karena kedua kelas memiliki karakteristik lahan yang relatif berdekatan, sehingga model mengalami sedikit kesulitan dalam membedakan batas antar kelas.

Hasil interpretasi menggunakan pendekatan *Explainable Machine Learning* menunjukkan bahwa fitur yang paling berpengaruh terhadap hasil klasifikasi adalah kemiringan_persen, *ph_tanah*, *fosfor_ppm*, *nitrogen_persen*, dan *kalium_cmolk*. Temuan ini sejalan dengan konsep kesesuaian lahan jagung yang menempatkan kondisi fisik lahan, tingkat keasaman tanah, dan ketersediaan unsur hara sebagai faktor penting dalam mendukung pertumbuhan tanaman. Dengan adanya *feature importance* dan *SHAP Summary Plot*, model tidak hanya menghasilkan prediksi kelas kesesuaian lahan, tetapi juga mampu menjelaskan alasan di balik keputusan tersebut. Hal ini menunjukkan bahwa pendekatan *Explainable Machine Learning* dapat meningkatkan transparansi model dan mendukung pengambilan keputusan pertanian berbasis data yang lebih objektif, informatif, dan mudah dipahami oleh petani, penyuluh, maupun pengambil kebijakan.

KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, pendekatan *Explainable Machine Learning* dapat digunakan secara efektif untuk klasifikasi kesesuaian lahan jagung berbasis data tabular. Dataset yang digunakan terdiri atas 560 data dengan 10 variabel input utama yang merepresentasikan karakteristik tanah, iklim, topografi, dan unsur hara. Setelah melalui tahapan *preprocessing*, data dibagi menjadi 448 data latih dan 112 data uji dengan distribusi kelas yang seimbang. Hasil pengujian menunjukkan bahwa model *Decision Tree* memperoleh performa terbaik dibandingkan *Random Forest*, *Support Vector Machine*, *K-Nearest Neighbor*, dan *XGBoost*, dengan nilai *accuracy* sebesar 0,9911, *precision* sebesar 0,9914, *recall* sebesar

0,9911, dan *F1-score* sebesar 0,9911. Hasil *confusion matrix* juga menunjukkan bahwa model hanya menghasilkan satu kesalahan prediksi dari total 112 data uji, sehingga model memiliki kemampuan generalisasi yang sangat baik dalam membedakan kelas Tidak Sesuai, Cukup Sesuai, Sesuai, dan Sangat Sesuai.

Hasil interpretasi menggunakan *feature importance* dan *SHAP Summary Plot* menunjukkan bahwa variabel yang paling berpengaruh terhadap klasifikasi kesesuaian lahan jagung adalah kemiringan lahan, pH tanah, fosfor, nitrogen, dan kalium. Temuan ini menunjukkan bahwa model tidak hanya mampu menghasilkan prediksi kelas kesesuaian lahan, tetapi juga dapat menjelaskan faktor dominan yang memengaruhi keputusan model.

Saran untuk penelitian selanjutnya adalah mengembangkan model dengan menggunakan data lapangan yang lebih luas dan beragam agar kemampuan generalisasi model dapat diuji pada wilayah yang berbeda. Selain itu, penelitian berikutnya dapat mengintegrasikan data tabular dengan data spasial, citra satelit, atau data temporal seperti musim tanam dan perubahan iklim agar analisis kesesuaian lahan menjadi lebih komprehensif. Pengembangan sistem pendukung keputusan berbasis web atau mobile juga disarankan agar hasil klasifikasi dan interpretasi model dapat dimanfaatkan secara langsung oleh petani, penyuluh pertanian, pemerintah daerah, dan pemangku kepentingan lainnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Dinas Tanaman Pangan, Hortikultura, Perkebunan dan Ketahanan Pangan Kabupaten Soppeng atas dukungan dan kontribusi yang diberikan dalam pelaksanaan penelitian ini. Dukungan tersebut sangat membantu penulis dalam memperoleh informasi terkait kondisi pertanian, karakteristik wilayah, serta data pendukung yang relevan dengan analisis kesesuaian lahan jagung. Ucapan terima kasih juga disampaikan kepada seluruh pihak yang telah membantu proses pengumpulan data, validasi informasi, dan penyusunan penelitian ini sehingga artikel dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

- [1] H. Afzal *et al.*, "Incorporating soil information with machine learning for crop recommendation to improve agricultural output," *Sci. Rep.*, vol. 15, no. 1, p. 8560, 2025, doi: 10.1038/s41598-025-88676-z.
- [2] A. Balasundaram, A. B. Abdul Aziz, A. Gupta, A. Shaik, and M. S. Kavitha, "A fusion approach using GIS, green area detection, weather API and GPT for satellite image based fertile land discovery and crop suitability," *Sci. Rep.*, vol. 14, no. 1, p. 16241, 2024, doi: 10.1038/s41598-024-67070-1.
- [3] A. Bhullar, K. Nadeem, and R. A. Ali, "Simultaneous multi-crop land suitability prediction from remote sensing data using semi-supervised learning," *Sci. Rep.*, vol. 13, no. 1, p. 6823, 2023, doi: 10.1038/s41598-023-33840-6.
- [4] M. Bouni, B. Hssina, K. Douzi, and S. Douzi, "Interpretable Machine Learning Techniques for an Advanced Crop Recommendation Model," *J. Electr. Comput. Eng.*, vol. 2024, no. 1, p. 7405217, Jan. 2024, doi: <https://doi.org/10.1155/2024/7405217>.
- [5] A. Fitzgibbon, D. Pisut, and D. Fleisher, "Evaluation of Maximum Entropy (Maxent) Machine Learning Model to Assess Relationships between Climate and Corn Suitability," 2022, doi: 10.3390/land11091382.
- [6] B. A. Ganati and T. M. Sitote, "Predicting land suitability for wheat and barley crops using machine learning techniques," *Sci. Rep.*, vol. 15, no. 1, p. 15879, 2025, doi: 10.1038/s41598-025-99070-0.
- [7] M. Ismaili *et al.*, "Assessment of Soil Suitability Using Machine Learning in Arid and Semi-Arid Regions," 2023, doi: 10.3390/agronomy13010165.
- [8] L. Li, L. Liu, Y. Peng, Y. Su, Y. Hu, and R. Zou, "Integration of multimodal data for large-scale rapid agricultural land evaluation using machine learning and deep learning approaches," *Geoderma*, vol. 439, p. 116696, 2023, doi: <https://doi.org/10.1016/j.geoderma.2023.116696>.
- [9] A. Li *et al.*, "Spatial suitability evaluation based on multisource data and random forest algorithm: a case study of Yulin, China," *Front. Environ. Sci.*, vol. Volume 12-2024, 2024, [Online]. Available: <https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2024.1338931>
- [10] A. B. Møller, V. L. Mulder, G. B. M. Heuvelink, N. M. Jacobsen, and M. H. Greve, "Can We Use Machine Learning for Agricultural Land Suitability Assessment?," 2021, doi: 10.3390/agronomy11040703.
- [11] A. Y. Ozalp and H. Akinci, "Evaluation of Land Suitability for Olive (*Olea europaea* L.) Cultivation Using the Random Forest Algorithm," 2023, doi: 10.3390/agriculture13061208.
- [12] D. Radočaj, M. Jurišić, M. Gašparović, I. Plaščak, and O. Antonić, "Cropland Suitability Assessment Using Satellite-Based Biophysical Vegetation Properties and Machine Learning," 2021, doi: 10.3390/agronomy11081620.
- [13] D. Radočaj and M. Jurišić, "GIS-Based Cropland Suitability Prediction Using Machine Learning: A Novel Approach to Sustainable Agricultural Production," 2022, doi: 10.3390/agronomy12092210.
- [14] D. Radočaj, A. Šiljeg, I. Plaščak, I. Marić, and M. Jurišić, "A Micro-Scale Approach for Cropland Suitability Assessment of Permanent Crops Using Machine Learning and a Low-Cost UAV," 2023, doi: 10.3390/agronomy13020362.
- [15] D. Radočaj, M. Gašparović, and M. Jurišić, "Cropland Suitability Prediction Method Based on Biophysical Variables from Copernicus Data and Machine Learning," 2025, doi: 10.3390/app15010372.
- [16] A. Rajbongshi *et al.*, "Leveraging explainable AI for sustainable agriculture: a comprehensive review of recent advances," *Artif. Intell. Rev.*, vol. 59, no. 3, p. 105, 2026, doi: 10.1007/s10462-025-11459-5.
- [17] S. Roopashree, J. Anitha, S. Challa, T. R. Mahesh, V. K. Venkatesan, and S. Guluwadi, "Mapping of soil suitability for medicinal plants using machine learning methods," *Sci. Rep.*, vol. 14, no. 1, p. 3741, 2024, doi: 10.1038/s41598-024-54465-3.
- [18] M. K. Senapaty, A. Ray, and N. Padhy, "A Decision Support System for Crop Recommendation Using Machine Learning Classification Algorithms," 2024, doi: 10.3390/agriculture14081256.
- [19] M. Y. Shams, S. A. Gamel, and F. M. Talaat, "Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making," *Neural Comput. Appl.*, vol. 36, no. 11, pp. 5695–5714, 2024, doi: 10.1007/s00521-023-09391-2.
- [20] S. Shastri, S. Kumar, V. Mansotra, and R. Salgotra, "Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence," *Sci. Rep.*, vol. 15, no. 1, p. 25498, 2025, doi: 10.1038/s41598-025-07003-8.

- [21] A. J. D. I Made Candra Girinata, Budi Styawan, Arwin Wahyu Saputra, M Aidil Arif, "Prediksi Pergerakan Harga Ethereum Menggunakan Machine Learning Dengan Algoritma Random Forest Dan Xgboost," *JUMISTIK J. Manaj. Inform. Sist. Inf. dan Teknol. Inform.*, vol. 4, no. 2, pp. 624–631, 2025, doi: 10.70247/jumistik.v4i2.222.
- [22] S. Ismail, "Data Mining Klasifikasi Penduduk Miskin Menggunakan Metode Support Vektor Machine," *J. Ilm. Sist. Inf. dan Tek. Inform.*, vol. 8, no. April, pp. 142–152, 2025, doi: <https://doi.org/10.57093/jisti.v8i1.283>.
- [23] I. Hermin, "Prediksi Produktivitas Jagung Berbasis Explainable Machine Learning dan Seleksi Fitur Adaptif," *J. Ilm. Sist. Inf. dan Tek. Inform.*, vol. 9, no. April, pp. 55–67, 2026.
- [24] I. Ismail and A. Hidayah, "Implementasi Machine Learning Dengan Metode Regresi Linear Untuk Prediksi Gaji Karyawan Berdasarkan Masa Kerja," *J. RISTER Ris. Sist. Cerdas*, vol. 2, no. 1, pp. 1–7, 2025, doi: 10.25126/Rister.