International Journal Software Engineering and Computer Science (IJSECS)
6 (2), 2026, 532-547
Published Online August 2026 in IJSECS (http://www.journal.lembagakita.com/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v6i2.7343.

RESEARCH ARTICLE **Open Access**

WebGIS-Based K-Nearest Neighbor Classification for Agricultural Land Suitability Assessment Across Multiple Commodities in Curah Kalak Village, Situbondo Regency

Moh. Alfian Husni Mubarok [1*], **A. Hamdani** [2], **Adi Susanto** [3]

[1*,2,3] Department of Information Technology, Faculty of Science and Technology, Universitas Ibrahimy, Situbondo Regency, East Java Province, Indonesia.

**Corresponding author*: alvianhusni80@gmail.com.

Received: May 30, 2026; Accepted: July 12, 2026; Published: August 1, 2026.

Abstract: Agricultural land suitability assessment plays an important role in supporting agricultural planning and ensuring that land resources are used according to environmental characteristics and crop requirements. In Curah Kalak Village, Situbondo Regency, land suitability assessment is generally conducted manually, resulting in limited efficiency and difficulties in accessing spatial information for decision-making. This study aims to develop a WebGIS-based agricultural land suitability classification system by integrating the K-Nearest Neighbor (KNN) algorithm with Geographic Information System (GIS) technology. The classification process uses environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance. A total of 669 agricultural land records were used as the dataset, and land suitability classes were classified using KNN with K = 31 based on Euclidean distance. The developed system classified land suitability for sugarcane, chili, corn, and rice commodities. The classification results indicated that chili and sugarcane were categorized as Highly Suitable (S1), whereas corn and rice were categorized as Moderately Suitable (S2). Performance evaluation using an 80:20 train-test split showed that the KNN model achieved an accuracy of 73.00%, weighted precision of 68.00%, weighted recall of 73.00%, and weighted F1-score of 70.00%, indicating moderate classification performance. Furthermore, the WebGIS provides interactive digital maps for visualizing classification results and spatial information. Black-box testing confirmed that all implemented system functions operated according to the specified requirements. The novelty of this study lies in the integration of KNN-based multi-commodity land suitability classification with WebGIS visualization within a village-level platform for agricultural land assessment. The proposed system can support agricultural land management and spatially informed decision-making in Curah Kalak Village.

Keywords: Agricultural Land Suitability; Geographic Information System; K-Nearest Neighbor; WebGIS; Land Classification.

1. Introduction

Agriculture remains an important sector in supporting food security and economic growth in Indonesia. As an agrarian country, Indonesia has extensive agricultural land resources that can be managed to support productivity and sustainable agricultural development. According to the Ministry of Agriculture of Indonesia (2024), the agricultural sector continues to contribute to national food security and rural economic development. In Situbondo Regency, agriculture is also an important economic sector, highlighting the need for effective agricultural land management and planning to support regional development (BPS Situbondo, 2024). However, land use that does not adequately consider environmental and physical land characteristics may result in low productivity, inefficient resource management, and environmental degradation. Therefore, agricultural land suitability assessment is needed to determine whether land conditions are appropriate for particular crops and to support more appropriate land-use planning. Land suitability evaluation combined with spatial analysis can also support agricultural land management and planning (Artikanur *et al.*, 2023).

Curah Kalak Village, located in Jangkar District, Situbondo Regency, has agricultural potential for the cultivation of sugarcane, chili, corn, and rice. However, agricultural land in the village has varying environmental and physical characteristics, resulting in different suitability levels for each commodity. Based on the conditions addressed in this study, land suitability assessment is still conducted manually, making the evaluation process time-consuming and less efficient. In addition, land suitability information has not been integrated with GIS-based visualization, which limits access to spatial agricultural information for farmers and related institutions. These conditions can constrain agricultural planning and decision-making because land suitability information is not presented in an integrated spatial form. Therefore, a system that combines land suitability classification with GIS visualization is needed to provide spatially referenced information that can be accessed and interpreted by agricultural stakeholders.

Agricultural land suitability assessment is a complex process because it requires the evaluation of multiple environmental and spatial parameters. Important parameters include slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance, which are related to crop growth, water availability, nutrient availability, and agricultural productivity (Sarkar *et al.*, 2023). Land suitability assessment can also be used to identify appropriate crop commodities and reduce the risk associated with unsuitable land use (Zilvanhisna *et al.*, 2025). For example, steep slopes may increase erosion risk, while unsuitable soil pH may affect nutrient availability and soil fertility. Rainfall, altitude, and soil depth can also influence local environmental conditions, root development, and soil water retention. Because these factors must be considered together, an organized approach is required to process the available data and produce consistent land suitability classifications.

Recent developments in information technology have encouraged the use of digital systems to support agricultural land management. Geographic Information Systems (GIS) can manage, analyze, and visualize spatial information, while classification methods can be used to determine land suitability classes based on environmental and spatial characteristics (Artikanur *et al.*, 2023). One classification method that has been applied in various classification studies is the K-Nearest Neighbor (KNN) algorithm, which determines the class of testing data based on its similarity to training data (Uddin *et al.*, 2022). Previous research has also demonstrated the application of KNN and other machine learning classifiers to spatial and environmental classification problems (Ashiagbor *et al.*, 2023). In addition, machine learning and spatial analysis have been applied to agricultural land suitability assessment (Møller *et al.*, 2021). The integration of classification methods with GIS can provide classification results together with their spatial representation, which can support the interpretation of agricultural land information (Roy *et al.*, 2024). GIS-based land suitability systems have also been developed to support agricultural planning and land management through spatial visualization and suitability mapping (Wicaksono *et al.*, 2026).

Several previous studies have examined the use of GIS, machine learning, and WebGIS technologies in environmental and agricultural applications. Ashiagbor *et al.* (2023) evaluated several machine learning classifiers, including K-Nearest Neighbor (KNN), for land-use and land-cover classification in Ghana's cocoa landscape. The study demonstrated the application of machine learning for spatial classification, but its focus was not on integrating the classification results into a WebGIS-based agricultural land suitability platform. Radočaj and Jurišić (2022) examined machine learning approaches for cropland suitability prediction and discussed their application to agricultural land assessment. However, the study focused on land suitability prediction rather than the development of an integrated WebGIS platform for presenting suitability results. Meanwhile, Vinueza-Martinez *et al.* (2024) reviewed WebGIS architectures and their application for geospatial information visualization and decision support across different sectors. However, the study did not specifically address the integration of agricultural land suitability classification within a WebGIS environment. Based on the studies reviewed in this research, limited attention has been given to the integration of KNN-based land suitability classification and WebGIS visualization for multiple agricultural commodities at the village level. This gap motivates the development of a system that can classify land suitability and present the results through an interactive WebGIS platform.

The contribution of this study consists of three main aspects. First, the study integrates KNN-based land suitability classification with WebGIS visualization within a single platform, allowing classification results to be presented together with their spatial distribution. Second, the system supports land suitability assessment for multiple commodities, namely sugarcane, chili, corn, and rice, within a unified classification framework. Third, the system is developed for village-level agricultural land assessment using environmental and spatial parameters collected from Curah Kalak Village, Situbondo Regency. To address the identified research gap, this study develops a WebGIS-based agricultural land suitability classification system using the K-Nearest Neighbor (KNN) algorithm and GIS technology. The system uses environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance, to classify land suitability for sugarcane, chili, corn, and rice. The dataset consists of 669 agricultural land records collected from Curah Kalak Village. The KNN classification model is evaluated using an 80:20 train-test split and Accuracy, Precision, Recall, and F1-score metrics to assess its classification performance. Based on the identified problems and research gap, the objectives of this study are: (1) to develop a WebGIS-based agricultural land suitability classification system using the K-Nearest Neighbor algorithm; (2) to classify land suitability for sugarcane, chili, corn, and rice based on environmental and spatial parameters; (3) to visualize the classification results through interactive GIS-based digital maps; and (4) to evaluate the performance and functionality of the developed system using quantitative classification metrics and black-box testing. By integrating KNN classification with GIS visualization, the proposed system is intended to provide accessible and spatially referenced land suitability information to support agricultural land management and planning in Curah Kalak Village.

2. Related Work

This section reviews the theoretical concepts and previous studies related to agricultural land suitability assessment, Geographic Information Systems (GIS), and the K-Nearest Neighbor (KNN) algorithm. The literature provides the theoretical basis for the classification method, spatial visualization, and development of the proposed system.

2.1 Agricultural Land Suitability

Web attack detection has attracted substantial research attention given the persistent and evolving threat Agricultural land suitability refers to the degree to which a particular area of land is appropriate for cultivating a specific agricultural commodity under existing environmental conditions. Land suitability assessment is conducted by evaluating land characteristics and comparing them with the requirements of particular crops to determine appropriate land use. Under the FAO land suitability framework, suitability is commonly classified into four categories: Highly Suitable (S1), Moderately Suitable (S2), Marginally Suitable (S3), and Not Suitable (N) (Cheng Han *et al.*, 2021). These classes represent different levels of compatibility between land characteristics and crop requirements, ranging from land with few or no significant limitations to land with severe limitations for the intended agricultural use. Land suitability evaluation generally considers environmental and spatial parameters such as slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation accessibility. These factors are associated with crop growth, nutrient availability, water retention, and agricultural productivity. Topographic, climatic, and soil physicochemical characteristics are important indicators in determining the suitability of land for agricultural production (Abate & Anteneh, 2024). Land suitability analysis is also applied in agricultural planning and land resource management because it provides information about the compatibility between land characteristics and crop requirements. The use of Geographic Information Systems (GIS) further supports land suitability assessment by enabling the spatial representation and analysis of land characteristics (AbdelRahman, 2025).

2.2 Geographic Information System (GIS)

A Geographic Information System (GIS) is a computer-based system used to capture, store, manage, analyze, and visualize geographic information. GIS combines spatial data, which represent geographic locations, with attribute data that describe the characteristics associated with those locations. This integration supports spatial operations such as mapping, spatial querying, overlay analysis, and environmental evaluation. These capabilities make GIS useful for analyzing land characteristics and supporting spatially informed decision-making. Abate and Anteneh (2024) reported that GIS-based analysis can support agricultural land suitability assessment by combining environmental, topographic, climatic, and soil-related information. The development of WebGIS has expanded access to GIS by allowing users to access and visualize spatial information through web-based platforms without requiring specialized GIS software. WebGIS supports information sharing and interactive digital mapping (Vinueza-Martinez *et al.*, 2024). In agriculture, GIS has been applied to land suitability assessment, crop monitoring, irrigation management, and agricultural planning.

The visual representation of spatial information can assist users in understanding the distribution and characteristics of agricultural land (Ersin Elbasi *et al.*, 2024). In this study, GIS is used to visualize agricultural land suitability classification results through an interactive WebGIS platform.

2.3 K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is a classification method that determines the class of new observations based on their similarity to previously labeled observations. The method identifies the nearest observations in the training dataset and assigns a class to the testing data based on the majority class among the selected neighbors. Because of its relatively simple classification mechanism, KNN has been applied to various classification problems involving multidimensional data (Elife Ozturk Kiyak, 2023). The KNN classification process begins by calculating the distance between testing data and observations in the training dataset. One commonly used distance measure is Euclidean Distance, which measures the distance between two observations in multidimensional space. The Euclidean Distance can be expressed as follows:

$$d\,(x,\,y) = \sqrt{\textstyle\sum_{i=1}^{n}(xi - yi)^2} \qquad (1)$$

Where:
$d(x, y) =$ distance between testing data and training data;
$x_i =$ value of the testing data attribute;
$y_i =$ value of the training data attribute;
$n =$ number of attributes used in the classification process.

After the distance values are calculated, the observations are sorted from the smallest to the largest distance. The KNN algorithm then selects the nearest observations according to a predefined value of K. The final class is determined using majority voting, in which the class occurring most frequently among the selected neighbors is assigned to the testing data (Rajib Kumar Halder *et al.*, 2024). Although KNN has a relatively simple classification mechanism, it has several limitations. The method can be sensitive to noisy observations and irrelevant attributes because classification decisions depend directly on the neighboring observations. KNN can also require greater computational effort as the number of observations increases because distances between testing data and training observations must be calculated. In addition, classification performance can be affected by the selection of the K value, as very small or very large values may influence the classification results. Several machine learning methods, including Support Vector Machine (SVM), Decision Tree, and Random Forest, have been applied to classification problems. KNN was selected in this study because its similarity-based classification mechanism is consistent with the use of environmental and spatial attributes for land suitability classification. The method also does not require a complex model-training process, which is suitable for the moderate-sized dataset used in this study.

2.4 Previous Studies

Previous studies have examined the application of GIS, spatial analysis, machine learning, and WebGIS in agricultural and environmental applications. These studies demonstrate the relevance of spatial information and classification methods for land evaluation and decision-making. However, the reviewed studies generally emphasize particular components, such as land suitability assessment, machine learning classification, spatial analysis, or WebGIS development. Based on the literature reviewed in this study, limited attention has been given to integrating KNN-based agricultural land suitability classification with WebGIS visualization for multiple agricultural commodities at the village level. A comparison between the previous studies and the proposed research is presented in Table 1.

Research	Method	Focus	Limitation	Gap
(Møller *et al.*, 2021)	Machine Learning for Land Suitability Assessment	Evaluated machine learning approaches for agricultural land suitability assessment	Focused on suitability prediction without GIS-based interactive visualization	Did not integrate spatial visualization and decision support systems.
(Ashiagbor *et al.*, 2023)	KNN, ANN, SVM, RF Classification	Compared machine learning classifiers for land-cover mapping	Focused on land-cover classification rather than	Did not classify agricultural land suitability for multiple commodities.

			agricultural land suitability	
(Sarkar *et al.*, 2023)	GIS, Remote Sensing, Bivariate Models	Agricultural land suitability for vegetable cultivation	Limited to GIS-based suitability mapping without machine learning classification	Did not implement predictive classification methods such as KNN.
(Artikanur *et al.*, 2023)	GIS and Analytic Network Process (ANP)	Sugarcane land suitability evaluation	Focused on a single commodity and ANP approach	Multi-commodity suitability classification was not addressed.
(Radočaj & Jurišić, 2022)	GIS-Based Machine Learning	Cropland suitability prediction using machine learning	Review-oriented study without WebGIS implementation	Did not provide an operational GIS-based decision support system
(Uddin *et al.*, 2022)	K-Nearest Neighbor Variants	Performance analysis of KNN algorithms	Conducted in disease prediction domain rather than agriculture	Did not investigate agricultural land suitability applications.
(Roy *et al.*, 2024)	Machine Learning and GIS Integration	Spatial analysis and decision-making using GIS and ML	General discussion of GIS-ML integration	No implementation for agricultural land suitability classification.
(Vinueza-Martinez *et al.*, 2024)	WebGIS Architecture	WebGIS development and geospatial information systems	Focused on WebGIS architecture trends	Did not integrate machine learning classification for agricultural land evaluation
(Zilvanhisna *et al.*, 2025)	Decision Support System and Fuzzy Mamdani	Plantation crop suitability recommendation	Used fuzzy logic and limited parameters	Did not apply KNN classification and GIS-based visualization
(Wicaksono Møller, 2026)	Rule-Based GIS and FAO Classification	Land suitability mapping for agricultural land	Rule-based approach without machine learning classification	Did not utilize KNN-based classification for multi-commodity suitability prediction.
	K-Nearest Neighbor (KNN) + GIS + WebGIS	Multi-commodity agricultural land suitability classification (rice, corn, chili, and sugarcane) in Curah Kalak Village	-	Integrates KNN classification, GIS spatial visualization, WebGIS accessibility, and multi-commodity land suitability evaluation in a single decision-support framework.

3. Methodology

This study developed a WebGIS-based agricultural land suitability classification system using the K-Nearest Neighbor (KNN) algorithm. The research workflow consisted of problem identification, data collection, data preparation, land suitability classification, GIS-based visualization, system implementation, and system evaluation. The overall research workflow is presented in Figure 1.

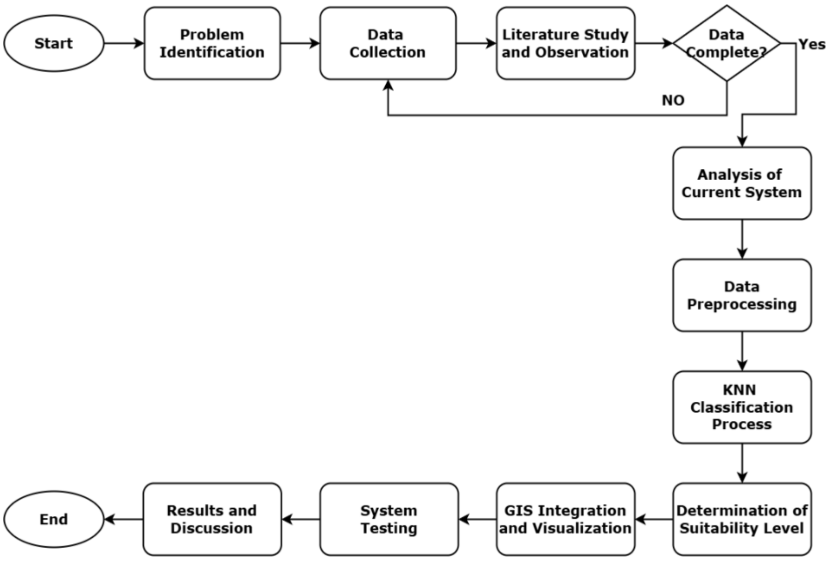

Figure 1. Research Flowchart

The research began by identifying the problems associated with agricultural land suitability assessment in Curah Kalak Village, Jangkar District, Situbondo Regency. Data were then collected from field observations, agricultural records, and relevant institutions, supported by a review of related literature. The collected data included slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance. These parameters were selected because they are related to land characteristics and crop growth requirements (Dornik *et al.*, 2024). The collected data were subsequently checked and prepared for classification. Land suitability classes were determined based on the suitability criteria for sugarcane, chili, corn, and rice and were used as class labels for the KNN classification process. The resulting classifications were then integrated with a Geographic Information System (GIS) to provide spatial visualization through a WebGIS platform. Finally, the classification model was evaluated using quantitative performance metrics, while black-box testing was used to verify the functionality of the developed system.

3.1 Research Workflow and Conceptual Framework

This study used a conceptual framework to describe the relationship between the research components and the sequence of activities involved in developing the WebGIS-based agricultural land suitability classification system. A conceptual framework can be used to describe relationships among research components and the sequence of processes involved in a study (Chukwuere, 2021). In this research, the framework describes the processing of environmental and spatial data, land suitability classification using KNN, and visualization of the classification results through GIS.

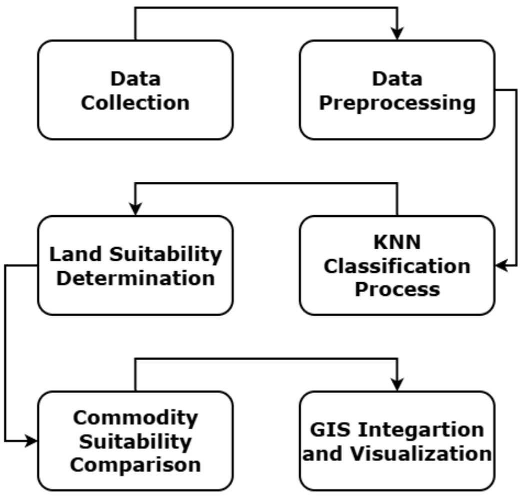

Figure 2. Conceptual Framework

Figure 2 illustrates that the research begins with the collection of environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance. The collected data are prepared and processed before the KNN classification is performed. The resulting suitability classes are used to represent the suitability levels of agricultural land for sugarcane, chili, corn, and rice. The classification results are subsequently integrated into a GIS-based WebGIS platform to provide interactive spatial visualization for agricultural land management and planning.

3.2 Data Collection

This study utilized environmental and spatial data obtained from field observations, agricultural records, and related institutions in Curah Kalak Village, Jangkar District, Situbondo Regency. These data were used as input variables in the K-Nearest Neighbor (KNN) classification process to determine agricultural land suitability levels for sugarcane, chili, corn, and rice commodities. The parameters used in this study consisted of slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance, as presented in Table 2.

Table 2. Environmental and Spatial Parameters

No	Parameter	Description
1	Slope	Land slope level used to determine erosion risk
2	Soil pH	Soil acidity level affecting crop growth
3	Soil Type	Soil characteristics influencing land suitability
4	Rainfall	Annual rainfall intensity in the study area
5	Altitude	Elevation above sea level
6	Soil Depth	Effective soil depth for root development
7	Irrigation Distance	Distance from agricultural land to irrigation source

In addition to environmental and spatial parameter data, this study utilized agricultural land suitability criteria for sugarcane, chili, corn, and rice commodities. These criteria were used as reference values for determining land suitability classes based on environmental characteristics and crop requirements. Each parameter was categorized into four suitability classes: highly suitable (S1), moderately suitable (S2), marginally suitable (S3), and not suitable (N). The agricultural land suitability criteria applied in this study are presented in Tables 3–6.

Table 3. Agricultural Land Suitability Criteria for Sugarcane

Parameter	S1	S2	S3	N
Slope (%)	<5	6 - 10	11 – 15	> 16
Altitude (m asl)	0 – 700	701 – 1,200	1,201 – 1,500	> 1,501
Rainfall (mm)	650 - 800	550 - 649	450 - 549	< 450
Soil pH	6.0 – 7.4	5.5 – 5.9	5.0 - 5.4	<4.9 or >7.5
Soil Type	1	2	3	4
Soil Depth (cm)	> 61	41 - 60	30 - 40	< 29
Irrigation Distance (m)	< 300	299 - 600	601 - 800	> 801

Table 4. Agricultural Land Suitability Criteria for Corn

Parameter	S1	S2	S3	N
Slope (%)	< 8	9–15	16–25	> 26
Altitude (m asl)	0–800	801–1,200	1,201–1,500	> 1,501
Rainfall (mm)	600–800	500–599	400–499	< 399
Soil pH	5.8–7.0	5.5–5.7	5.0–5.4	< 4.9 or > 7.1
Soil Type	1	3	2	4
Soil Depth (cm)	> 51	41–50	30–40	< 29
Irrigation Distance (m)	< 200	201–500	501–800	> 801

Table 5. Agricultural Land Suitability Criteria for Chili

Parameter	S1	S2	S3	N
Slope (%)	< 5	6–10	11–15	> 16
Altitude (m asl)	0–899	900–1,299	1,300–1,699	> 1,700
Rainfall (mm)	650–900	550–649	450–549	< 450
Soil pH	6.1–6.7	5.6–6.0	5.0–5.5	< 4.9 or > 6.8
Soil Type	1	2	3	4
Soil Depth (cm)	> 61	41–60	29–40	< 30
Irrigation Distance (m)	< 200	201–400	401–700	> 701

Table 6. Agricultural Land Suitability Criteria for Rice

Parameter	S1	S2	S3	N
Slope (%)	< 3	4–9	10–15	> 16
Altitude (m asl)	0–500	501–800	801–1,200	> 1,201
Rainfall (mm)	700–900	600–699	500–599	< 499
Soil pH	5.5–7.5	5.0–5.4	4.5–4.9	< 4.4 or > 7.6
Soil Type	2	3	1	4
Soil Depth (cm)	> 51	41–50	31–40	< 30
Irrigation Distance (m)	< 100	101–300	301–600	> 601

Note: Soil Type 1 = Loam Soil, Soil Type 2 = Cubi Soil, Soil Type 3 = Clay Soil, and Soil Type 4 = Sandy Soil.

In this study, soil type was represented using numerical identifiers to enable its inclusion in the KNN classification process. The assigned values served as categorical labels rather than quantitative measurements and were applied consistently across the dataset. Although this approach allows categorical attributes to be processed within the Euclidean Distance framework, future studies may explore alternative encoding techniques, such as one-hot encoding, to minimize potential ordinal interpretation effects. The suitability criteria presented in Tables 3–6 were used to assign land suitability classes to agricultural land records. The dataset consisted of 669 records collected from Curah Kalak Village, including 296 sugarcane, 132 chili, 124 rice, and 117 corn observations. These classified records were used as training data, while testing data were evaluated using environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance, to determine land suitability classes.

Table 7. Distribution of Land Suitability Classes

Suitability Class	Number of Records
S1	245
S2	357
S3	67
N	0
Total	669

Table 7 shows that the dataset was not evenly distributed across suitability classes. The S2 class contained the largest number of observations (357 records), followed by S1 (245 records) and S3 (67 records), while no observations belonged to the N class. This class imbalance may influence KNN classification performance because classes with larger numbers of observations tend to dominate the majority voting process. The limited number of S3 observations may also affect the model's ability to distinguish marginal suitability classes accurately.

3.3 Data Preprocessing

Before the classification process, numerical attributes were normalized using the Min-Max Normalization technique to reduce differences in attribute scales. This step was necessary because the K-Nearest Neighbor algorithm calculates similarity using Euclidean Distance, which is sensitive to variations in numerical ranges among features. The normalization process transformed all numerical attributes into a range between 0 and 1 while preserving the relative distribution of the original data. The Min-Max Normalization formula is expressed as follows:

$$X^{1} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

Where:
X^1 = normalized value
X = original value
X_{min} = minimum value
X_{max} = maximum value

The normalized attributes included slope, soil pH, altitude, soil depth, and irrigation distance. Soil type was treated as a categorical attribute and encoded separately for classification purposes.

Table 8. Attribute Range Used for Min-Max Normalization

Parameter	Min	Max
Slope (%)	1	7
Altitude (m asl)	27	49
Soil pH	4.40	7.40
Soil Depth (cm)	27	75
Irrigation Distance (m)	45	990
Rainfall (mm/year)	652	652

The minimum and maximum values used in the Min-Max Normalization process were derived from the 669 agricultural land records contained in the training dataset. Therefore, the normalization ranges presented in Table 8 reflect the actual distribution of environmental and spatial attributes observed in the study area rather than the theoretical suitability thresholds presented in Tables 3–6. Rainfall was retained as an environmental parameter because it represents an important criterion in FAO-based agricultural land suitability assessment. Since all agricultural land records were collected from the same study area, Curah Kalak Village, the recorded annual rainfall value was identical across observations (652 mm/year). Consequently, rainfall exhibited zero variance and did not influence Euclidean Distance calculations during the KNN classification process.

3.4 KNN Classification Process

The K-Nearest Neighbor (KNN) method was applied to classify agricultural land suitability based on the similarity between testing data and training data. The classification process utilized environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance. These parameters were used to determine the suitability level of agricultural land for sugarcane, chili, corn, and rice commodities. The value of KK was initially determined using the square-root heuristic:

$$K \approx \sqrt{N} \quad (3)$$

Where N represents the total number of training records. Since N=669N = 669, the square-root heuristic produced a value of:

$$\sqrt{669} = 25.86 \quad (4)$$

Several odd-numbered K values around this range were evaluated during preliminary experiments, and (K=31) was selected because it provided more stable classification performance and reduced tie occurrences during majority voting. The classification process utilized 669 labeled agricultural land records and environmental parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance. Euclidean Distance was used to measure similarity between testing instances and training records. The 31 nearest neighbors were then identified, and the final land suitability class was determined through majority voting. The resulting classifications were grouped into four categories: Highly Suitable (S1), Moderately Suitable (S2), Marginally Suitable (S3), and Not Suitable (N).

3.5 GIS Visualization Process

The land suitability classes generated by the K-Nearest Neighbor (KNN) classification process were integrated into a Geographic Information System (GIS) to produce spatial visualizations of agricultural land suitability. This integration enabled users to observe the distribution of land suitability classes interactively through a web-based mapping platform. The developed WebGIS utilized spatial and attribute data to represent agricultural land areas in Curah Kalak Village through digital map layers. Each land parcel was displayed according to its suitability level for sugarcane, chili, corn, and rice commodities. Through interactive map visualization, users could access land suitability information more easily, identify the spatial distribution of suitability classes, and support agricultural planning and land management activities.

3.6 System Architecture

The architecture of the developed system is presented in Figure 3. The system was designed as a WebGIS-based agricultural land suitability classification application that integrates the K-Nearest Neighbor (KNN) method with Geographic Information System (GIS) visualization.

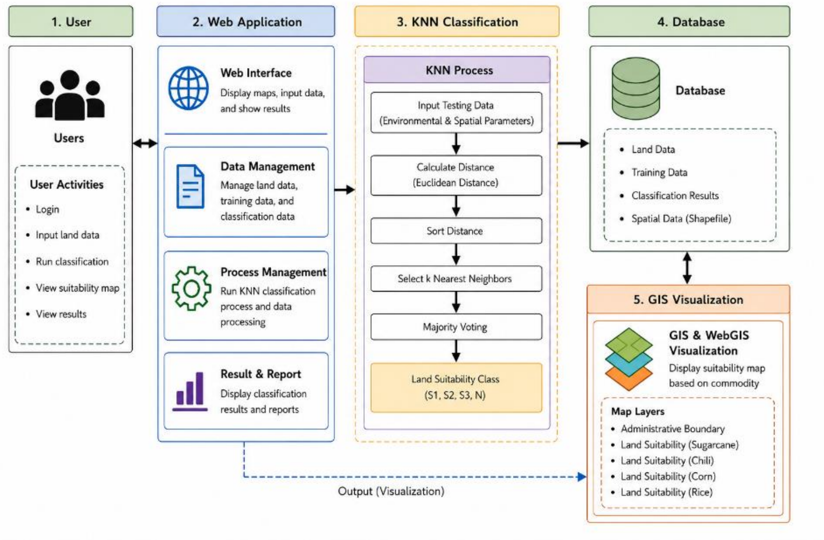

Figure 3. System Architecture

The system architecture consists of five main components: users, web application, KNN classification module, database, and GIS visualization. These components work together to support agricultural land suitability classification and spatial information visualization within a web-based environment. Users interact with the web application to manage land data, perform classifications, and view land suitability information. The KNN module processes environmental and spatial parameters by calculating Euclidean Distance, selecting the nearest neighbors, and determining land suitability classes through majority voting. The classification results are stored in the database and visualized through the WebGIS interface, enabling users to view the spatial distribution of land suitability classes for sugarcane, chili, corn, and rice commodities through interactive digital maps.

3.7 Tools and Technologies

Several software tools and technologies were utilized to support the development of the WebGIS-based agricultural land suitability classification system. PHP was used as the primary programming language, while MySQL served as the database management system for storing environmental, spatial, and classification data. The development process was carried out using XAMPP as the local server environment and Visual Studio Code (VS Code) as the source code editor. Microsoft Excel was used to organize and manage environmental and spatial parameter datasets before they were imported into the system. For spatial visualization, the system utilized Leaflet.js to display interactive digital maps and present agricultural land suitability information through a WebGIS interface.

3.8 Classification Performance Evaluation

In addition to black-box testing, the performance of the K-Nearest Neighbor (KNN) model was evaluated using quantitative classification metrics. An 80:20 train-test split was applied to the dataset, which consisted of 669 agricultural land records containing slope, altitude, rainfall, soil pH, soil type, soil depth, and irrigation distance attributes. The evaluation utilized numerical attributes normalized through the Min-Max Normalization procedure described in Section 3.3, transforming values into a common range of 0–1 to reduce scale-related bias in Euclidean Distance calculations. The normalized data were then classified using KNN with K=31K=31. Model performance was assessed using Accuracy, Precision, Recall, and F1-Score metrics, while a Confusion Matrix was used to analyze the distribution of correctly and incorrectly classified observations across suitability classes. The evaluation metrics used in this study were Accuracy, Precision, Recall, and F1-Score.

4. Result and Discussion

4.1 Results

The developed application was implemented as a WebGIS-based agricultural land suitability classification system that integrates the K-Nearest Neighbor (KNN) method with Geographic Information System (GIS) visualization. The system classifies agricultural land suitability using environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance. The implementation consists of two main modules: the KNN classification module and the GIS visualization module. The KNN module determines land suitability classes based on the similarity between testing data and training records, while the GIS module presents the classification results through interactive digital maps for sugarcane, chili, corn, and rice commodities. Figure 4 illustrates the KNN classification process, while Figure 5 presents the GIS visualization interface.

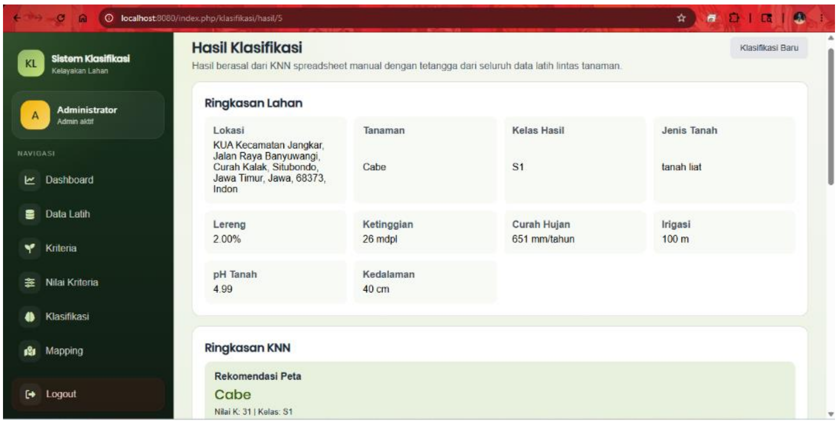

Figure 4. KNN Classification Process

Figure 4 presents an example of the agricultural land suitability classification generated by the system. Based on the environmental and spatial parameters entered into the system, the evaluated land was classified as Highly Suitable (S1) for chili cultivation using K = 31. The classification was performed using 669 labeled agricultural land records stored in the database. The result illustrates the implementation of the KNN method in generating land suitability predictions based on the similarity between the input data and the available training records.

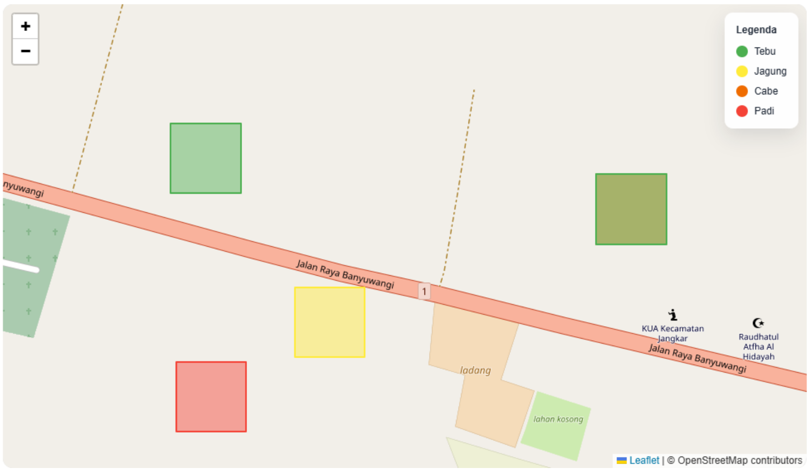

Figure 5. Spatial Distribution of Agricultural Commodities and Land Suitability Results

Figure 5 presents the spatial distribution of agricultural land parcels within the WebGIS platform. Each polygon represents an agricultural land parcel and its associated commodity information. The land suitability classification results are linked to the mapped parcels and can be accessed through the interactive interface. This visualization allows users to identify the geographical distribution of agricultural land and examine the corresponding suitability information within the study area.

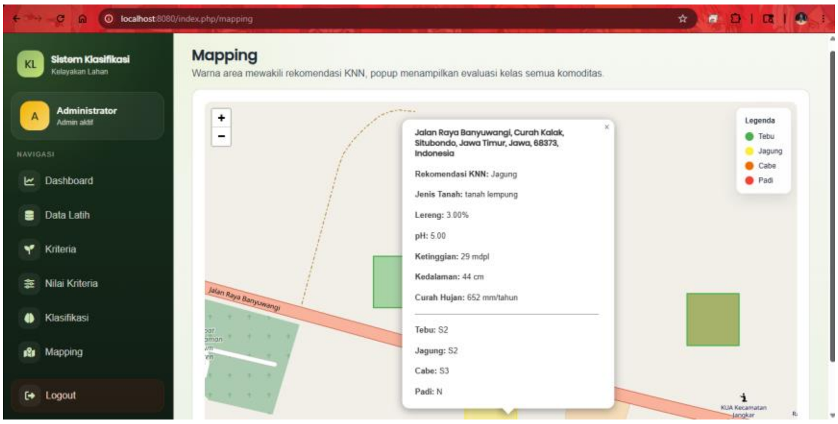

Figure 6. GIS Visualization

Figure 6 presents detailed information for individual land parcels, including commodity type, environmental parameters, and the resulting land suitability classification. Through the interactive map interface, users can access the classification results together with the environmental data used in the KNN process. This functionality supports the interpretation of land suitability information in relation to its spatial location.

Table 9. Examples of KNN Land Suitability Classification Results

Commodity	Slope (%)	Altitude (m asl)	Rainfall (mm/year)	Soil pH	Soil Type	Soil Depth (cm)	Irrigation (m)	k	Class
Chili	2.00	27	652	4.99	Clay Soil	40	100	31	S1
Sugarcane	2.00	27	652	5.00	Loam Soil	40	100	31	S1
Corn	3.00	29	652	5.00	Loam Soil	44	500	31	S2
Rice	3.00	28	652	4.99	Clay Soil	29	500	31	S2

Table 9 presents representative examples of classification outputs generated by the system. The examples include the environmental and spatial parameters used as input variables, the selected K value, and the resulting suitability class. The four examples represent the classification results for chili, sugarcane, corn, and rice. Although only four records are presented for illustration, the KNN classification process was applied to the agricultural land records used in the study. The overall distribution of suitability classes is presented in Table 7. The developed application was evaluated using black-box testing to verify the functionality of the implemented features. The testing covered the login page, training data management, suitability criteria management, KNN classification, and GIS visualization. The results are presented in Table 10. The testing results show that all tested functions operated according to the specified test scenarios. The login process was successfully executed, while the training data and criteria value modules successfully supported data addition, editing, deletion, display, and upload. The KNN module successfully processed testing data, calculated Euclidean distances, selected nearest neighbors, and generated classification results. The GIS module successfully displayed maps, land parcels, popup information, and classification results.

Table 10. Black-Box Testing Results

No	Tested Component	Testing Scenario	Result	Status
1	Login Page	Admin login process	Successful	Valid
2	Training Data Page	Add, edit, delete, display, and upload training data	Successful	Valid
3	Criteria Value Page	Add, edit, delete, display, and upload criteria values	Successful	Valid
4	KNN Classification Page	Input testing data, Euclidean Distance calculation, nearest neighbor selection, and classification results	Successful	Valid
5	GIS Mapping Page	Display GIS map, land area visualization, popup information, and classification results	Successful	Valid

The results indicate that the implemented functions met the expected functional requirements under the tested scenarios. However, the black-box testing results describe functional conformity based on the tested scenarios and should not be interpreted as evidence that the system is free from all possible errors In addition to functional testing, the KNN model was evaluated using an 80:20 train-test split. The evaluation used Accuracy, Precision, Recall, and F1-Score to measure the classification performance of the model. A Confusion Matrix was also used to examine the distribution of correct and incorrect predictions across the observed suitability classes. The dataset contained 669 agricultural land records. The 20% testing portion consisted of 133 observations, as reflected in the confusion matrix. The KNN model was configured with K = 31, and the numerical attributes were normalized using Min-Max Normalization before distance calculations.

Table 11. Classification Performance Results

Metric	Value
Accuracy	73.00%
Precision (Weighted Avg)	68.00%
Recall (Weighted Avg)	73.00%
F1-Score (Weighted Avg)	70.00%

The reported evaluation results indicate that the KNN model achieved an accuracy of 73.00%, precision of 68.00%, recall of 73.00%, and F1-score of 70.00%. The accuracy indicates that most testing observations were classified correctly, while the recall value indicates that approximately 73% of the actual class assignments were correctly identified. However, the precision and F1-score values should be verified against the confusion matrix in Table 12. The values reported in Table 11 should be recalculated using the same averaging method and testing observations represented in Table 12 to ensure consistency.

Table 12. Confusion Matrix

Actual Class	Predicted S1	Predicted S2	Predicted S3	Pred N
S1	32	12	0	0
S2	14	65	0	0
S3	3	7	0	0
N	0	0	0	0

The confusion matrix contains 133 testing observations. Of these, 44 belonged to S1, 79 belonged to S2, and 10 belonged to S3. No testing observations belonged to the N class. The matrix shows that 32 of the 44 S1 observations were correctly classified, while 65 of the 79 S2 observations were correctly classified. None of the 10 S3 observations was predicted as S3. Instead, three S3 observations were classified as S1 and seven were classified as S2. This indicates that the model had difficulty distinguishing the S3 class from the neighboring suitability classes. The absence of predictions for S3 and N should be considered when interpreting the model performance. The S3 class had only 67 observations in the complete dataset, while the N class had no observations. The class distribution may therefore affect the majority-voting mechanism of KNN and reduce the model's ability to identify less-represented classes.

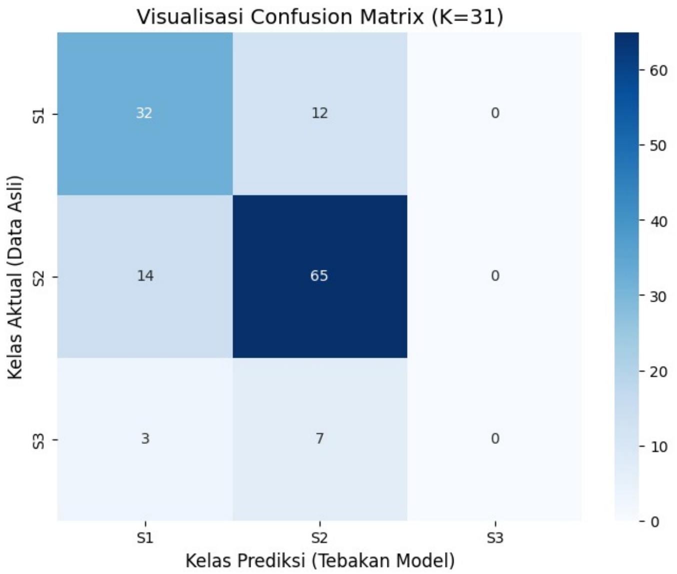

Figure 7. Confusion Matrix Visualization

Figure 7 presents the visual representation of the confusion matrix. The diagonal values indicate correctly classified observations, particularly for S1 and S2. Misclassifications mainly occurred between S1 and S2, while all S3 observations were classified as either S1 or S2. This pattern indicates that the model had difficulty separating the marginal suitability class from the more frequently observed classes.

4.2 Discussion

The developed system combines K-Nearest Neighbor (KNN)-based land suitability classification with Geographic Information System (GIS) visualization within a single WebGIS platform. Using environmental and spatial parameters from 669 agricultural land records, the system provides suitability classifications for sugarcane, chili, corn, and rice. The implementation shows that the classification results can be linked to the geographic location of agricultural land parcels, allowing suitability information to be accessed through an interactive spatial interface. This provides a practical connection between land classification and spatial information within the study area. The use of KNN is consistent with previous research that applied machine learning methods to land-use and environmental classification (Ashiagbor *et al.*, 2023). KNN was selected because its classification process is based on the similarity between an input observation and labeled training observations, which is appropriate for the environmental and spatial attributes used in this study. However, the model achieved a reported accuracy of 73.00%, indicating that its classification performance was moderate. Therefore, the KNN model should be considered a practical classification method for the developed WebGIS rather than a highly accurate predictive model.

The distribution of suitability classes is an important factor in interpreting the classification performance. The dataset consisted of 245 S1 records, 357 S2 records, and 67 S3 records, while no records belonged to the N class. The larger number of S2 observations may influence the majority-voting process of KNN, particularly for observations whose characteristics are located close to the boundaries between suitability classes. This pattern is reflected in the confusion matrix, where several S1 observations were classified as S2 and all S3 observations were classified as either S1 or S2. The limited number of S3 observations may therefore reduce the model's ability to distinguish this class from the more frequently represented classes. Rainfall also requires particular consideration in interpreting the KNN classification. All 669 agricultural land records had the same rainfall value of 652 mm/year. Consequently, rainfall had zero variance and did not contribute to differences in Euclidean Distance between observations. Although rainfall is included as an important parameter in the suitability criteria, its constant value in the dataset means that it could not provide discriminatory information for the KNN model. This condition should be clearly distinguished from the role of rainfall in the general land suitability framework. The parameter was retained because it forms part of the suitability criteria, but its contribution to the KNN distance calculation was effectively absent in the current dataset.

The representation of soil type also presents a methodological limitation. Soil type was converted into numerical identifiers to allow it to be included in the Euclidean Distance calculation. However, these values represent categorical classes rather than quantitative measurements. As a result, the Euclidean Distance calculation may treat the numerical difference between soil types as meaningful even though the categories do not have an inherent ordinal relationship. For example, a difference between codes 1 and 2 is treated as smaller than a difference between codes 1 and 4, although this numerical relationship does not necessarily represent the actual similarity between soil types. One-hot encoding or a distance method designed to handle mixed numerical and categorical variables could be considered in future studies. The absence of N-class observations is another limitation of the current dataset. Although the suitability framework defines four classes, namely S1, S2, S3, and N, no agricultural land records in the dataset were assigned to the N class. Consequently, the model's ability to identify unsuitable land cannot be evaluated from the current testing data. Additional observations representing N-class conditions would be required to assess the classification performance across all suitability categories.

The selection of K = 31 was based on the square-root heuristic followed by preliminary testing of several odd-numbered K values. The square-root heuristic provided an initial reference for determining the K value, while the final selection was based on the reported stability of the classification results. However, the manuscript does not present the comparative performance obtained from the different K values tested. The inclusion of these results would strengthen the justification for selecting K = 31 because readers could evaluate the performance of the selected K value against the other tested configurations. Overall, the findings indicate that the developed system can connect KNN-based classification with WebGIS visualization and provide spatially accessible land suitability information. Nevertheless, the moderate classification performance, imbalanced class distribution, constant rainfall variable, categorical soil-type encoding, and absence of N-class observations limit the interpretation of the model as a general predictive system. These limitations should be considered when the WebGIS is used to support agricultural land planning and land suitability assessment in Curah Kalak Village.

5. Conclusion and Recommendations

This study successfully developed a WebGIS-based agricultural land suitability classification system by integrating the K-Nearest Neighbor (KNN) method with Geographic Information System (GIS) technology. The system used environmental and spatial parameters, including slope, soil pH, soil type, rainfall, altitude, soil depth, and irrigation distance, to classify agricultural land suitability for sugarcane, chili, corn, and rice in Curah Kalak Village, Situbondo Regency. The classification process used 669 agricultural land records and K = 31. The representative classification results showed that chili and sugarcane were classified as Highly Suitable (S1), while corn and rice were classified as Moderately Suitable (S2). These results indicate that, for the evaluated examples, the land characteristics were more closely associated with the training patterns of the S1 class for chili and sugarcane and the S2 class for corn and rice. However, these classifications should be interpreted as KNN predictions based on similarity to the training data rather than as direct assessments in which every individual parameter satisfies the corresponding suitability criteria. The KNN model achieved a reported accuracy of 73.00%, precision of 68.00%, recall of 73.00%, and F1-score of 70.00%. These results indicate moderate classification performance. However, the evaluation results should be verified against the confusion matrix before publication because the reported precision and F1-score need to be consistent with the same testing data and averaging method used to obtain the performance metrics. In addition, the absence of N-class observations means that the model's ability to classify unsuitable land was not evaluated. The integration of GIS enabled the classification results to be associated with agricultural land parcels and displayed through interactive digital maps. Black-box testing also showed that the tested system functions operated according to the specified scenarios. Therefore, the developed WebGIS can serve as a functional platform for presenting land suitability information and supporting agricultural land planning in Curah Kalak Village. The moderate model performance and dataset limitations should, however, be considered when interpreting the classification results.

For future research, the dataset should be expanded to include a more balanced representation of suitability classes, particularly the N class and the relatively small S3 class. Additional data should also be collected for environmental parameters that vary among land parcels. In the current dataset, rainfall had the same value for all observations and therefore did not contribute to the Euclidean Distance calculation. Soil type should also be represented using an encoding method that is more appropriate for categorical variables, such as one-hot encoding, or through a distance measure designed for mixed numerical and categorical data. Further studies should report the comparative results of different K values to provide stronger justification for selecting K = 31 and should compare KNN with other classification methods, such as Decision Tree, Random Forest, and Support Vector Machine (SVM). These improvements may provide a more reliable basis for evaluating and applying agricultural land suitability classification within the WebGIS platform.

References

Abate, S. G., & Anteneh, M. B. (2024). Assessment of agricultural land suitability for cereal crops based on the analysis of soil physico-chemical characteristics. *Environmental Systems Research, 13*, Article 6. https://doi.org/10.1186/s40068-024-00333-y

AbdelRahman, M. A. E., Yossif, T. M. H., & Metwaly, M. M. (2025). Enhancing land suitability assessment through integration of AHP and GIS-based for efficient agricultural planning in arid regions. *Scientific Reports, 15*, Article 31370. https://doi.org/10.1038/s41598-025-14051-7

Artikanur, S. D., Widiatmaka, Setiawan, Y., & Marimin. (2023). An evaluation of possible sugarcane plantations expansion areas in Lamongan, East Java, Indonesia. *Sustainability, 15*(6), Article 5390. https://doi.org/10.3390/su15065390

Ashiagbor, G., Asare-Ansah, A. O., Amoah, E. B., Asante, W. A., & Mensah, Y. A. (2023). Assessment of machine learning classifiers in mapping the cocoa-forest mosaic landscape of Ghana. *Scientific African, 20*, Article e01718. https://doi.org/10.1016/j.sciaf.2023.e01718

Cheng, H., 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), Article 370. https://doi.org/10.3390/agriculture11040370

Chukwuere, J. E. (2021). Theoretical and conceptual framework: A critical part of information systems research process and writing. *Review of International Geographical Education (RIGEO), 11*(9), 2678–2683. https://doi.org/10.48047/rigeo.11.09.234

Dornik, A., Chetan, M. A., Crisan, T. E., Heciko, R., Gora, A., Dragut, L., & P. P. [Author information should be verified against the final published record]. (2024). Geospatial evaluation of the agricultural suitability and land use compatibility in Europe's temperate continental climate region. *International Soil and Water Conservation Research, 12*(4), 908–919. https://doi.org/10.1016/j.iswcr.2024.01.002

Elbasi, E., Mostafa, N., Zaki, C., AlArnaout, Z., [remaining authors should be verified against the final published record]. (2024). Optimizing agricultural data analysis techniques through AI-powered decision-making processes. *Applied Sciences, 14*(17), Article 8018. https://doi.org/10.3390/app14178018

Kawung, Y., Tooy, D., & Pakasi, S. (2023). Design of a web-based geographic information to show spatial information of land used for horticulture. *Agro Bali: Agricultural Journal, 6*(3), 581–594. https://doi.org/10.37637/ab.v6i3.1373

Kiyak, E. O., Ghasemkhani, B., & Birant, D. (2023). High-level K-nearest neighbors (HLKNN): A supervised machine learning model for classification analysis. *Electronics, 12*(18), Article 3828. https://doi.org/10.3390/electronics12183828

Møller, A. B., Mulder, V. L., Heuvelink, G. B. M., Jacobsen, N. M., & Greve, M. H. (2021). Can we use machine learning for agricultural land suitability assessment? *Agronomy, 11*(4), Article 703. https://doi.org/10.3390/agronomy11040703

Patel, S., Kaur, B., Verma, S., Sood, A., Litoria, P. K., & Pateriya, B. (2023). Web GIS based decision support system for agriculture monitoring and management. *Geoinformatics & Geostatistics: An Overview, 11*(1), Article 1000133. https://doi.org/10.4172/2327-4581.1000133

Radočaj, D., & Jurišić, M. (2022). GIS-based cropland suitability prediction using machine learning: A novel approach to sustainable agricultural production. *Agronomy, 12*(9), Article 2210. https://doi.org/10.3390/agronomy12092210

Roy, P. P., Abdullah, M. S., & Siddique, I. M. (2024). Machine learning empowered geographic information systems: Advancing spatial analysis and decision making. *World Journal of Advanced Research and Reviews, 22*(1), 1387–1397. https://doi.org/10.30574/wjarr.2024.22.1.1200

Sarkar, D., Saha, S., & Mondal, P. (2023). Modelling agricultural land suitability for vegetable crops farming using RS and GIS in conjunction with bivariate techniques in the Uttar Dinajpur district of Eastern India. *Green Technologies and Sustainability, 1*(2), Article 100022. https://doi.org/10.1016/j.grets.2023.100022

Uddin, S., Haque, I., Lu, H., Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports, 12*, Article 10358. https://doi.org/10.1038/s41598-022-10358-x

Vinueza-Martinez, J., Correa-Peralta, M., Ramirez-Anormaliza, R., Arias, O. F., & Vera Paredes, D. (2024). Geographic information systems (GISs) based on WebGIS architecture: Bibliometric analysis of the current status and research trends. *Sustainability, 16*(15), Article 6439. https://doi.org/10.3390/su16156439

Wicaksono, A., Manalu, D. S. T., Sebayang, V. B., Pratama, A. J., Pamungkas, M. A., & Puspa, A. S. (2026). Geographic information system for land suitability mapping of partner farmers at Okiagaru Indonesia Agricoop using rule-based system and prototype methodology. *Jurnal Teknik Informatika (JUTIF), 7*(2), 1455–1467. https://doi.org/10.52436/1.jutif.2026.7.2.5471

Zilvanhisna, E. F., Irawan, A. D., Madjid, A., & Imron, A. M. N. (2025). Implementation of decision support system for analyzing the suitability of plantation crops. *International Journal of New Media Technology, 12*(1), 8–13.