Journal of Computer Science

Research Article

Development of a Machine Learning Based Mobile and Web Crop Recommendation System for Precision Farming in Wardha

Aishwarya Kadu and KTV Reddy

Department of Artificial Intelligence and Data Science, Faculty of Engineering and Technology, DMIHER DU: Datta Meghe Institute of Higher Education and Research, Wardha, India

Article history
Received: 27-04-2025
Revised: 19-07-2025
Accepted: 10-10-2025

Corresponding Author:
Aishwarya Kadu
Department of Artificial Intelligence and Data Science, Faculty of Engineering and Technology, DMIHER DU: Datta Meghe Institute of Higher Education and Research, Wardha, India
Email: aishwaryakadu17@gmail.com

Abstract: Agricultural productivity is crucial for economic growth of a nation, especially for countries like India which is highly dependent on agriculture. This study addresses the need for cutting-edge technology to enhance crop productivity by proposing a Mobile/Web-based precision farming system. Leveraging Machine Learning (ML) algorithms, the proposed system recommends optimal crops based on several factors such as environmental parameters (Nitrogen (N), Phosphorus (P), Potassium (K)), pH value, Temperature, Rainfall, Humidity and Soil Nutrients., The current study applied five ML algorithms Naïve Bayes, Random Forest, on a dataset comprising of 2200 records and 7 variables to create a predictive ML model. Among the five algorithms, the Naïve Bayes exhibits superior performance, achieving a precision rate of 99.50%. Despite challenges like data scarcity and the need for field validation, the research aims to provide Wardha farmers with a free and open-source precision farming platform. By facilitating informed crop management decisions, the platform seeks to augment agricultural productivity and foster sustainable economic growth in the region. The robustness of the proposed system is further validated by the identified algorithmic accuracy, with Random Forest emerging as the most efficient option for crop recommendation within the specified environmental context. This study under-scores the potential of precision farming to revolutionize agriculture and contribute to long-term economic development in emerging regions. The validation process involves a train-test split, cross-validation, and field validation to ensure the reliability and accuracy of the machine learning models. Additionally, the impact assessment highlights potential benefits such as improved crop yields, resource optimization, sustainable economic growth, and long-term environmental benefits for the Wardha region.

Keywords: Precision Farming, Machine Learning, Environmental Factors, Sustainable Economic Growth

Introduction

The agricultural world in India is a remarkable story, intricately woven with tradition, innovation, and immense importance. Its roots stretch back to the Neolithic period, and today, it stands as a powerhouse, ranking second globally in farm output. This sector serves as the lifeblood of India, particularly in rural areas, employing over half the work-force and contributing a significant 20.2% to the GDP (as per the 2020-21 economic survey). The land boasts incredible diversity, from being the world's second-largest producer of rice and wheat to leading in a vast array of crops like pulses, vegetables, fruits, sugar-cane, and spices. This abundance is a testament to India's agricultural prowess. However, the path forward is not without its challenges. Despite revolutions like the Green Revolution that ensured food security, issues like resource scarcity, soil degradation, and integrating modern practices require ongoing attention. In essence, Indian agriculture is a complex and fascinating story of tradition meeting innovation, striving for abundance while addressing significant challenges.

However, a new chapter is being written in this story. The agricultural world is witnessing a revolution in efficiency with the rise of precision farming. This innovative approach departs from the traditional one-size-fits-all methods, where farmers rely on intuition and experience to manage vast fields. Instead, precision farming embraces technology to cultivate smarter, not harder. It hinges on gathering and analyzing data to understand the specific needs of individual plants or even small field sections. With this granular knowledge, farmers can meticulously tailor inputs like water, fertilizer, and pesticides. This ensures optimal utilization of resources, minimizing waste and maximizing crop yields. This data-driven approach promises to transform Indian agriculture further, fostering sustainability and ushering in a new era of informed decision-making for a more productive future. By combining traditional knowledge with cutting-edge technology, precision farming holds immense potential to address the challenges faced by Indian agriculture and ensure its continued success in feeding a growing population.

In India, agriculture is vital as it provides employment for a large percentage of the population and makes up a sizeable piece of the nation's GDP. However, the industry needs help, including scarce water supplies, erratic weather patterns, and restricted access to vital data and resources. To improve agricultural output in the nation, there has been a noticeable increase in interest in developing a Crop Recommendation System (CRS) in response to these difficulties and the rising realization of the need for innovations (Abu et al., 2022).

CRS use analytical methods to forecast which crops would thrive in a given area and under particular environmental circumstances. These systems offer farmers tailored ad-vice using several data sources, including soil analysis, weather forecasts, satellite imaging, and past crop yields (Ahmed et al., 2021).

The current study aims to introduce a predictive ML model that uses a customized CRS to increase agricultural output in Wardha, Maharashtra. The approach entails a re-view of the body of research on CRS and analyzing the difficulties the Wardha agriculture industry faces. The report then describes the strategy used, including information on the dataset, prediction methods, and assessment criteria. The study's conclusions, which include forecast accuracy and CRS efficacy, are then explained (Arshaghi et al., 2023; Baltrušaitis et al., 2018; Bechar and Vigneault, 20216).

The introduction provides an overview of Artificial Intelligence (AI) in agricultural precision, emphasizing ML, an aspect of AI. The experimental assessment using several ML algorithms, Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), and Naïve Bayes (NB), is described in depth in the study’s methods section (Benos et al., 2021). The discussion section explores the study's implications for raising agricultural output in Wardha, Maharashtra, and suggests possible directions for further re-search.

Precision Farming

Precision Farming (PF) is a technology-driven method of managing farms that aims to maximize harvests and reduce waste by providing farmers with up-to-date data and information about their fields and animals. This technology helps in making well-informed decisions to maximize crop yields and minimize the losses (Cai et al., 2019). Data analysis from many sources is made possible by AI, which is essential for making well-informed judgments on crop management (Cao et al., 2020). PF has significantly benefited from ML that uses large datasets to identify patterns automatically and formulate forecasts built on those designs (Cedric et al., 2022).

IoT sensors are integrated into a system for PF to collect various indicators critical to farming, including environmental factors like pH values, Temperature (T), Humidity (H), and weather. The complete data set is further enhanced by indications relating to cattle, such as weight increase, feed consumption, and veterinarian well-being (Kadu and Reddy, 2023b). Then, using this data, many aspects of managing crops and animals are optimized, such as targeted fertilization, precise irrigation, disease avoidance, and efficient pest control (Kadu and Reddy, 2024a). Moreover, real-time decision-making is supported by the ongoing monitoring made possible by IoT sensors, which enables farmers to react quickly to situations that change. The sensors' data-driven insights support resource-efficient and sustainable agricultural methods, advancing precision farming programs' economic viability and environmental stewardship. Figure 1 shows essential elements of agricultural precision:

- IoT (Internet of Things): Collects real-time data from sensors (soil, weather, crops, etc.)
- AI (Artificial Intelligence): Analyzes this data to generate predictions, recommendations, and decisions
- Precision Farming (Overlap): When IoT data and AI analysis combine, farmers can apply the right inputs (water, fertilizer, pesticides) at the right time and place, leading to higher yield, reduced cost, and sustainable farming

After being sent to technical experts, the data collected by these sensors is interpreted, and valuable insights are extracted using various data processing techniques. Making timely and precise judgments in PF requires this knowledge. To improve comprehension, Figure 2 shows a summary of the overall process in Precision Agriculture:

- IoT Sensors in Farm Sensors collect real-time data from soil, crops, and weather conditions
- Communication Channel: The data is transmitted through networks like Wi-Fi, satellite, or mobile internet
- Application of AI: AI analyzes the data to give predictions, recommendations, and insights
- Information to Farmer The processed information is sent back to the farmer, helping in better decision-making (e.g., irrigation, fertilization, pest control)

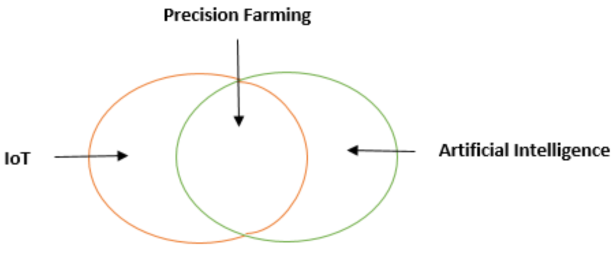

Fig. 1: Essential elements of agricultural precision

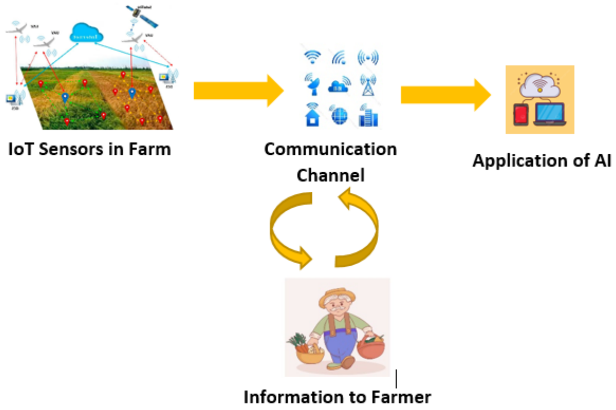

Fig. 2: Summary of Overall Process in Precision Agriculture

Furthermore, combining robots and data analytics in precision farming supports real-time decision-making and predictive analysis. With the help of this predictive ability, farmers may plan for possible obstacles, allocate resources as efficiently as possible, and take proactive steps to manage their animals and crops. The smooth data progression across these phases amplifies precision farming methods' overall efficacy and durability (Kadu and Reddy, 2023b; 2024b).

Precision farming benefits greatly from using ML, especially regarding disease identification and crop production prediction. Accurate crop production estimates are made possible by ML algorithms, which are skilled at evaluating data from various sensors, including weather stations, satellite imaging, and soil sensors (Kadu and Reddy, 2024c; Kadu et al., 2024). One such instance is a research (Kadu and Reddy, 2023a) that employed ML to predict China's winter wheat production using satellite images and meteorological data.

ML is also essential for the early identification of crop diseases in PF. When ML algorithms are used to analyze crop photos, they can detect symptoms of illness or stress and promptly notify farmers so they can take preventative measures (Kadu and Reddy, 2025) Research (Senoo et al., 2024) used ML in Tanzania to identify cassava illness from drone-captured photos. Interestingly, the machine learning model proved highly accurate at recognizing unhealthy plants even when the symptoms were invisible to the naked eye. These examples demonstrate how ML may improve PF techniques by improving yield prediction and disease detection.

Literature Review

The growing integration of ML in contemporary precision agriculture has led to numerous studies across diverse farming domains (Kadu and Reddy, 2024c; 2025; Senoo et al., 2024; Cao et al., 2020; Cedric et al., 2022; Chen et al., 2021). Several recent studies showcase the application of ML across various facets of modern precision agriculture. Table 1 provides a comprehensive summary of numerous research initiatives in this field. Fan et al. (2021) concentrates on predicting agricultural yields, leveraging ML approaches to enhance accuracy by integrating data from multiple sources. In disease detection, Fei et al. (2023) utilizes Convolutional Neural Networks (CNNs) and transfer learning to accurately identify leaf diseases in tomatoes and grapes. Agricultural management takes center stage for Folberth et al. (2016) utilizing ML to analyze aerial photos for precise forecasts regarding crop output and nutritional deficiencies. According to Forkuor et al. (2017) soil mapping employs sensor data and ML techniques to map soil features accurately. Foucquier et al. (2013) employs image analysis in pest detection, using deep convolutional neural networks to identify tomato pests precisely.

Fuentes et al. (2017) focuses on irrigation optimization, employing reinforcement learning to enhance precision irrigation for increased yields and water efficiency. Agricultural growth modelling is the focal point for Gasmi et al. (2022) utilizing deep learning with climatic and satellite data to simulate crop development trends precisely. Gebbers and Adamchuk (2010) explores wheat yield prediction, employing ML algorithms and UAV-based multi-sensor data fusion. Gondia et al., 2020) investigates the estimation of plant nitrogen status, demonstrating how hyperspectral imaging and machine learning yield accurate results.

He et al. (2019) employs deep learning to monitor crop health, specifically focusing on banana health, using aerial photo data. Huang et al. (2022) effectively applies deep learning and machine vision techniques for maize kernel seed selection. Javaid et al. (2023) discusses using UAV photos and machine learning algorithms for wheat yield estimation, while Jeong et al. (2016) concentrates on precisely identifying weeds in rice using Convolutional Neural Networks. In crop management, Jiang et al. (2022) integrates GIS. and random forests to identify nutrient deficits accurately based on aerial images.

Jones et al. (2017) employs Convolutional Neural Networks for precise identification in potato disease diagnosis, utilizing deep learning and image analysis. These studies underscore the myriad applications and significant outcomes of machine learning in precision

DOI: 10.3844/jcssp.2026.1045.1059

agriculture across various crop management and observation facets.

Table 1 showcases the work of many academics who have created AI results specifically for several agricultural zones, containing crop collection and farming methods. Using state-of-the-art methods, these solutions mainly comprise recommendation and prediction systems (Krasilnikov et al., 2022). Though these methods have potential, they fail to address the real-world requirement for accessible, field-tested solutions that enable farmers and decision-makers to maximize revenues and manage resources.

A small number of recent researches have examined this implementation difficulty in practice. Instead of practical answers, the majority of research often provides theoretical insights. Bridging the gap to making these cutting-edge solutions available to a larger audience especially farmers is the key to the dilemma. Our research approaches this problem methodically and offers a thorough road map for resolving these obstacles. It also presents instances of how these technologies might be made publicly accessible to democratize access and increase exposure for the good of the larger agricultural community (Li et al., 2022).

State of Art

Table 2 categorizes essential data types utilized in precision farming, highlighting their examples, importance, and collection methods. It emphasizes six key data types: Soil data, weather data, crop data, environmental data, economic data, and farmer management practices. Each category plays a vital role in facilitating informed decision-making within agricultural practices. For instance, soil data, which includes metrics such as pH levels and nutrient content, is crucial for assessing soil fertility and determining the suitability of land for specific crops. In contrast, weather data, comprised of temperature and precipitation readings, directly influences crop growth rates, irrigation needs, and the risks associated with pests and diseases.

Furthermore, crop data is indispensable for selecting appropriate crop varieties and predicting yields based on historical performance and pest occurrences.

Table 1: Precision Agriculture Research - Key Studies and Contributions Meteorological data

Author	Field Used	Algorithms/Techniques Used	Output/Conclusion
(Fan et al., 2021)	Crop Yield Prediction	ML Integration From Multiple Data Sources	Improved Accuracy
(Fei et al., 2023)	Disease Detection	CNNs, Transfer Learning	Identification of Leaf Diseases in Tomatoes and Grapes
(Folberth et al., 2016)	Crop Management	ML Analysis of Aerial Photos	Precise Forecasts on Crop Output and Nutritional Deficiencies
(Forkuor et al., 2017)	Soil Mapping	ML Techniques with Sensor Data	Accurate Mapping of Soil Features
(Foucquier et al., 2013)	Pest Detection	Deep CNNs	Precise Identification of Tomato Pests
(Fuentes et al., 2017)	Irrigation Optimization	Reinforcement Learning	Increased Yields and Water Efficiency in Precision Irrigation
(Gasmi et al., 2022)	Crop Growth Modeling	Deep Learning with Climatic and Satellite Data	Precise Simulation of Crop Development Trends
(Gebbers and Adamchuk, 2010)	Forecast for Grain Yield	ML Algorithms, Multi-sensor Data Fusion Based on UAVs	Precise Estimation of Grain Production
(Gondia et al., 2020)	Nitrogen Status Estimation	Hyperspectral Imaging, ML	Accurate Estimation of Plant Nitrogen Status
(He et al., 2019)	Crop Health Monitoring	Deep Learning with Aerial Photo Data	Focus on Banana Health
(Huang et al., 2020)	Seed Selection	Deep Learning, Machine Vision	Effective Maize Kernel Selection
(Javaid et al., 2023)	Wheat Yield Prediction	UAV Photos, ML Algorithms	Estimation of Wheat Yields
(Jeong et al., 2016)	Weed Detection	CNNs	Precise Identification of Weeds in Rice
(Jiang et al., 2022)	Crop Management	Random Forests, G.I.S.	Accurate Prediction of Nutrient Deficiency Based on Aerial Images
(Jiang et al., 2022)	Disease Detection	CNNs, Deep Learning, Image Analysis	Precise Identification in Potato Disease Diagnosis

Table 2: Key Data Sources and Their Importance in Precision Farming

Data Type	Examples	Importance	Collection Methods	Ref
Soil Data	pH levels, nutrient content (N, P, K), moisture content, soil temperature	Assesses soil fertility and suitability for crop growth	Soil sampling, laboratory analysis, remote sensing, sensors	(Maniruzzaman et al., 2020)
Weather Data	Temperature, precipitation, humidity, wind speed, solar radiation	Influences crop growth, irrigation needs, and pest/disease risks	Meteorological stations, satellite data, forecasting models	(Mendez et al., 2020)
Crop Data	Crop type, growth stages, historical yields, pest/disease occurrence	Essential for crop selection, yield prediction, and management	Farm management software, field surveys, remote sensing	(Ouzzani et al., 2016)
Environmental Data	Terrain, water availability, land use patterns, biodiversity	Provides context for sustainable land management	GIS, remote sensing, environmental monitoring systems	(Pande et al., 2021)
Economic Data	Market prices, input costs (seeds, fertilizers), labor costs	Informs decisions on crop selection, input management, and profitability	Market surveys, government reports, farm accounting records	(Patel et al., 2022)

Environmental data provides a broader context for sustainable land management by incorporating aspects such as terrain and water availability. Economic data empowers farmers to make financially sound decisions regarding crop selection and input management, while insights into farmer and management practices ensure that recommendations are practical and tailored to local conditions. The collection methods for these data types vary, encompassing a range of techniques from soil sampling and laboratory analysis to remote sensing and agricultural software. This comprehensive approach underscores the significance of data integration in optimizing agricultural productivity and sustainability.

Crop Recommendation System (CRS)

According to our analysis, the market for free and open-source precision agricultural solutions is dynamic and needs significant developments to rival proprietary options. Despite these obstacles, the involvement of a cooperative network of farmers and IT experts have been crucial in promoting the development of open-source substitutes (Bechar and Vigneault, 2016). This group effort guarantees a flexible environment sensitive to difficulties arising from integrating technology. From this angle, our primary goal was to launch a web- and mobile-based crop-recommendation system. This method uses machine learning algorithms to forecast the best crop to harvest while accounting for critical environmental factors, including soil nutrient levels, temperature, humidity, and rainfall.

The thorough performance analysis shows that Naive Bayes (N.B.) consistently outperformed our study's other machine learning models. As such, to utilize the robust features of Naive Bayes (N.B.) and Random Forest (R.F.) to forecast the ideal crop according to user-input data, such as pH levels, temperature, humidity, rainfall, and Potassium (K) and nitrogen (N). After determining which model performed the best, we serialized and saved the model using pickle, Python's built-in persistence mechanism. Our crop recommendation algorithm is built on top of this serialized model.

Decision Trees are tree-based models that partition data into branches based on feature values. They are characterized by their simplicity and the ease with which they can be interpreted and visualized. For instance, Decision Trees can effectively handle numerical and categorical data, making them a flexible choice for various datasets. However, they are prone to overfitting, particularly when dealing with noisy data, and even small changes in the data can lead to significantly different tree structures. In crop recommendation systems, Decision Trees are particularly useful for small to medium-sized datasets where clear decision rules exist. For example, a Decision Tree can be utilized to recommend crops by considering factors like soil type, local climate, and historical yield performance, allowing farmers to understand and apply the model's recommendations easily.

Random Forests are ensembles of multiple decision trees that average predictions to enhance accuracy and robustness. This method significantly reduces the risk of overfitting seen in single decision trees and performs well with large datasets, providing stability against noise. However, this model requires more computational resources and is less interpretable than individual trees. In agricultural applications, Random Forests are suited for large, complex datasets. They can combine the predictions from numerous decision trees to more accurately recommend suitable crops based on multifaceted input data, such as climate patterns, market trends, and soil analyses. Support Vector Machines (SVM) seek to find the optimal hyperplane to separate data points into distinct classes. They are particularly effective in high-dimensional spaces, making them suitable for binary and multiclass classification problems, including scenarios involving small datasets. The main drawback of SVM is its sensitivity to the choice of kernel and parameters, which can complicate model tuning, making SVM computationally demanding. In crop recommendation tasks, SVM can be used effectively for binary classification, such as determining whether to plant Crop A or Crop B, especially when differentiating factors are complex and nuanced.

The k-Nearest Neighbors algorithm classifies data based on the majority vote of the nearest k neighbors in the feature space. It is straightforward and intuitive, not requiring an extensive training process, making it particularly effective for non-linear data. Nevertheless, the model can become computationally expensive as the dataset grows larger and is sensitive to noise and irrelevant features. In crop recommendation systems, k-NN can suggest crops by examining the similarity of new data points, such as current soil and weather conditions, to previously successful crop scenarios. This method is most effective for smaller datasets in which clear patterns are easily observable.

Artificial Neural Networks comprise multi-layered networks that can learn intricate patterns from data. They excel in capturing non-linear relationships and are adaptable to various types of input data, allowing for complex interactions between features. However, ANN models require extensive training data, computational power, and careful tuning to avoid overfitting. When utilized in crop recommendation systems, ANNs can model the relationships between environmental factors, soil characteristics, and crop performance. This capability enables them to provide recommendations tailored to specific agricultural conditions and market dynamics.

Deep Learning encompasses advanced neural networks with multiple hidden layers, excelling in tasks such as image recognition and sequence prediction. These models handle large-scale data extremely well and learn complex patterns effectively. Nevertheless, they demand significant computational resources, can be challenging to interpret, and may be deemed excessive for simpler tasks. In agriculture, Deep Learning can be particularly beneficial in analyzing satellite images for crop health assessment or predicting future crop yields based on historical data patterns. This can provide farmers with critical insights and enhance decision-making processes.

Naive Bayes is a probabilistic model grounded in Bayes' theorem, making assumptions of feature independence. This model is known for its speed and efficiency, particularly with larger datasets, and minimal training data requirements. However, it may struggle with complex feature interactions due to its independence assumptions. Naive Bayes can facilitate quick and efficient crop recommendations in scenarios where datasets are small and factors like soil type, weather conditions, and crop requirements can be considered independently. This can be especially useful for preliminary analyses and fast decision-making.

Gradient Boosting Machines are a class of boosting methods that build models sequentially to correct the errors made by previous iterations. This approach often results in high accuracy, particularly in complex datasets, and it is robust against overfitting when properly tuned. However, the training process can be relatively slow, and careful parameter adjustments are necessary. In the context of crop recommendation systems, GBM can improve the accuracy of predictions by iteratively refining the model. This is especially valuable in highly variable agricultural environments, where nuanced crop recommendations can have significant impacts on yield and profitability.

The various ML models outlined provide a range of options for developing crop recommendation systems, each with unique advantages and considerations that can be tailored to the agricultural context and the specific needs of farmers.

Materials and Methods

Concept of the System

The first step in developing a specific dataset for crop recommendations was creating our Crop Recommendation System (CRS). We carefully preprocessed the dataset for thorough analysis using soil nutrient data (Cai et al., 2019; Cao et al., 2020; Cedric et al., 2022) and climatic data unique to Wardha (Maharashtra). The dataset seven variables and 2200 entries was thoroughly examined to learn more about its underlying properties. To complement this, we used Machine Learning (ML) methods such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naïve Bayes (NB), and Logistic Regression (LR). This multimodal method helped find the best characteristics and set the stage for building reliable machine learning systems Liu et al. (2020).

Our focus on refining the dataset and applying various machine learning techniques underscores the accuracy and sophistication of our Crop Recommendation System (CRS).

Figure 3 provides a detailed illustration of the steps that must be taken one after the other in order to formulate the design and develop the CRS. After the models were built, the next stage involved evaluating their performance. After the assessment, the emphasis turned to implementing the model that performed the best into an online application Makridakis et al., (2018). The core infrastructure of our Crop Recommendation System (CRS) is this Mobile/web-based application.

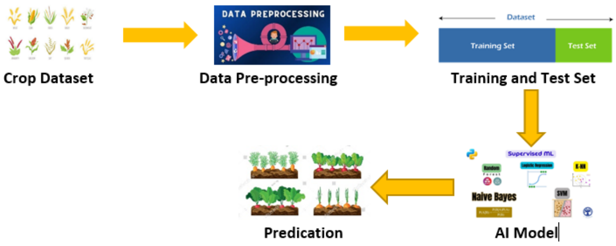

Fig. 3: Predicative model steps

Crop Dataset: The first dataset, as shown in Figure 3, contains information about various crops. Each dataset entry will likely include characteristics such as soil nutrients, climate, and other relevant elements (Maniruzzaman et al., 2020).

Data Preprocessing: Figures 3 and 4 show that the dataset is preprocessed before being fed into a machine learning model. Managing missing data, scaling or standardizing features, encoding categorical variables, and ensuring data are presented correctly for training are some of the activities of this procedure (Mendez et al., 2020).

Training and Test Set: The standard division of the dataset into a training set and a test set is depicted in Figure 3. The machine learning model is trained on the training set, and its performance on unknown data is evaluated on the test set (Ouzzani et al., 2016). This section assesses the model's ability to generalize new and untested data.

AI Model Prediction: Figure 3 shows that the model may predict new or unknown data after training on the selected training set. This suggests that the A.I. model can predict the best crops in a crop recommendation system based on input parameters like soil nutrient levels and climate (Pande et al., 2021).

Predictive Model: The basis for producing crop forecasts is the predictive model graphically depicted in Figure 4. This model uses machine learning algorithms to forecast the best-cultivated crops by analyzing input parameters such as soil nutrient levels and climatic conditions (Patel et al., 2022).

Best Performing Model Selection: Various machine learning methods are thoroughly evaluated throughout choosing the best-performing model, as shown in Figure 4. After a thorough evaluation, the model that predicts crop outcomes with the best accuracy and effectiveness is selected for future development (Patel and Patel, 2020; Paymode and Malode, 2022).

Model creation as Mobile/Web App to Local Server: The progression from model creation to the deployment of the chosen model as a mobile/web application on a local server is shown in Figure 4. This phase entails turning the prediction model into an application that is easy to use and available on online and mobile platforms (Popescu et al., 2020; Ramcharan et al., 2017).

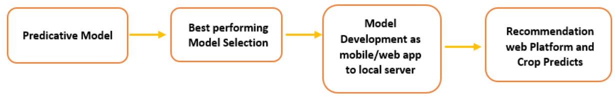

Fig. 4: Recommendation system (R.S.) building steps

Recommendation Crop Predictions: A recommendation system building last step, as shown in Figure 4. By integrating the selected prediction model, this platform lets customers obtain crop suggestions based on inputted or real-time environmental data. It provides an intuitive user interface for smooth system operation and accurate crop prediction (Paymode and Malode, 2022; Popescu et al., 2020; Ramcharan et al., 2017).

Setting Up the Dataset and Choosing the Features

The dataset selected for crop selection comprises crucial attributes such as soil pH, Temperature (T), Humidity (H), Rainfall (R), and nutrient levels of P, K, N, and pH values totalling 2,200 entries. The statistical information for each agricultural parameter, including Min, M.A.X., Mean, And Standard Derivation, is summarized in Table 2 and visually presented in Figure 6. Exploring the relationship between input variables and the target revealed an Average Correlation Value of 0.152, particularly notable in parameters like potassium, phosphorus, and nitrogen. This positive correlation signals promising connections, laying the groundwork for subsequent prediction processes. Conversely, negative correlation values in specific indicators underscore the dataset's diversity and insightful richness of information.

Figure 6 is a comprehensive visual representation of the Statistical Analysis of Agricultural Attributes, offering insights into the distribution and variability of Min, Max, Mean, and Standard Deviation for each agricultural parameter. This graphical summary aids in interpreting the dataset, providing valuable information for informed agricultural decision-making.

From a biological perspective, Nitrogen (N), Potassium (K), and Phosphorus (P) are vital micronutrients crucial for plant development. Nitrogen promotes leaf growth, potassium supports normal plant processes, and phosphorus aids fruit ripening. To solve soil shortages, fertilizers containing components like soil moisture and nutrients like P, K, N, and pH levels can be customized for particular crops. Understanding the ideal P.K.N. ratios is essential for optimizing fertilization operations and increasing crop yields (Raschka, 2018)

Data preprocessing involves examining box plots of features before and after scaling Figure 5. Before scaling, box plots are created for each feature, displaying quartiles, median, and possible outliers. Interpreting these plots helps understand the distribution and dispersion of values for each characteristic, assessing whether feature scales differ and might dominate the modelling process.

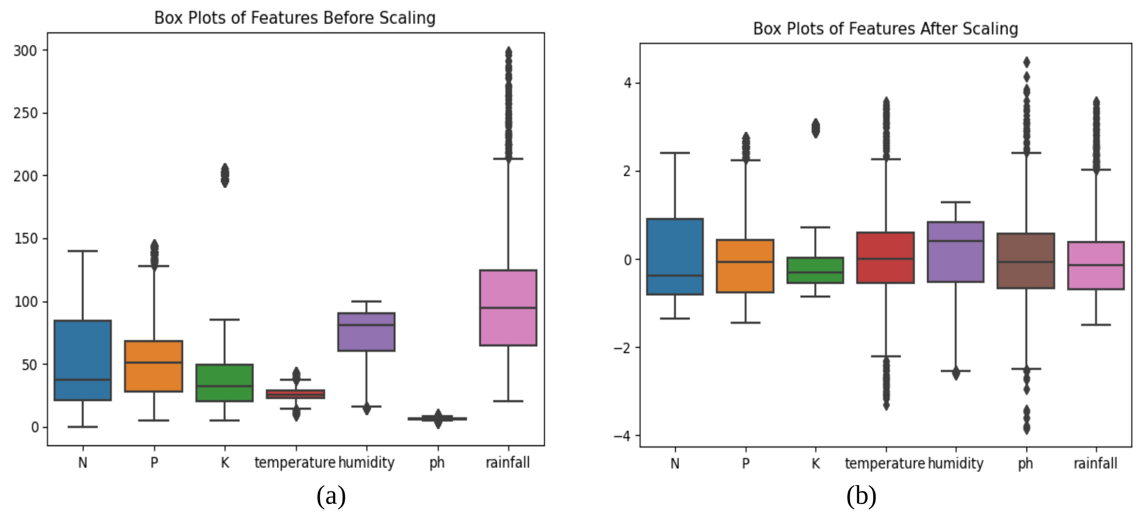

Fig. 5: Data Preprocessing Box Plots of Features Before and After Scaling

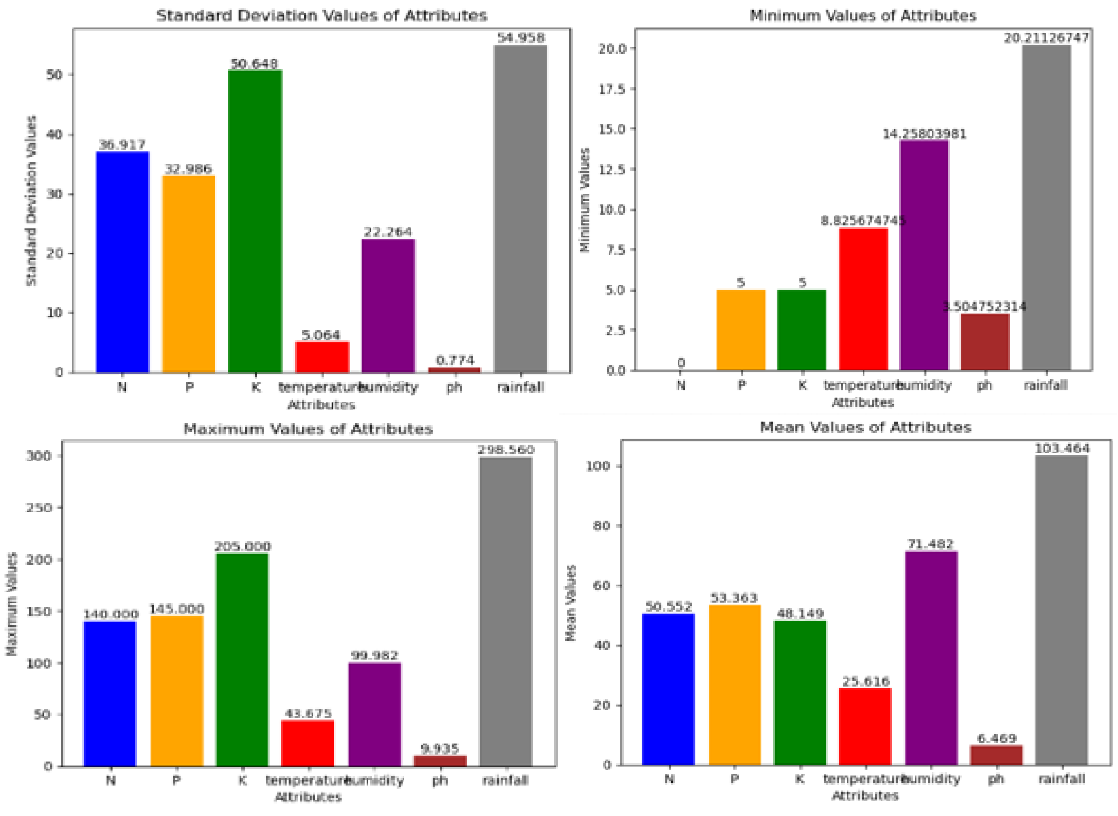

Fig. 6: Statistical Analysis of Agricultural Attributes

Scaling is crucial when features have varying scales, preventing a single characteristic from exerting undue influence on the model. Two standard scaling techniques are Min-Max Scaling, which limits values to a predetermined range (usually 0 to 1), and Standardization, which sets the standard deviation to one and the mean to zero.

After scaling, new box plots for each feature are created to observe median, quartiles, and outlier's changes. The interpretation involves examining the effects of scaling, ensuring features now have a comparable scale, and reducing the likelihood of one trait dominating others during modelling. This integrated approach, illustrated in Figure 5, enhances the

DOI: 10.3844/jcssp.2026.1045.1059

understanding and preparation of the dataset for subsequent predictive processes.

Other variables in our dataset, such as Rainfall (R), Temperature (T), Humidity (H), and soil pH, are essential in the crop selection process. While the amount of precipitation is critical to plant life, the pH value significantly impacts the presence of essential chemicals. The temperature of the air directly affects photosynthesis, and while too much or too little soil moisture can be harmful, it is necessary for plant development. In Figure 6, the crop data included in our dataset includes dates, apples, maize, grapes, oranges, peaches, potatoes, onions, tomatoes, olives, and watermelons. The characteristics above must be considered in order to anticipate these crop yields.

Development of Predictive ML Models

The Python programming language was used to create predictive machine learning models. Essential libraries, such as Matplotlib, NumPy, Pandas, and Scikit-learn, were loaded during the first round of data preparation. In order to prepare the dataset for model training, this stage included essential activities, including managing missing values and encoding categorical variables.

A 70:30 ratios were used to divide the dataset into training and testing sets. Then, five machine learning models were chosen for testing: Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), and Naive Bayes (NB) (Rehman et al., 2022)

The chosen training dataset was used to establish and train the models. After each model was trained, predictions were produced for each model on the test set. Several measures were used to evaluate each model's performance during the testing and assessment phase, including accuracy, precision, recall, and others.

A thorough examination of the outcomes involved contrasting each model's performance. Matplotlib and other visualization tools were used to provide the results in an understandable and straightforward format (Rokade et al., 2022). This procedure made it possible to fully comprehend how each model functioned and added to the predictive power within the parameters of the provided dataset.

Evaluation Metrics

When assessing the performance of regression models in machine learning, three commonly used metrics are the R^2-Score, Root Mean Squared Error (RMSE), and Mean Squared Error (MSE).

R^2-Score (Coefficient of Determination)

The R^2-Score indicates how well the independent variables explain the variability of the dependent variable. It ranges from 0 to 1, with 1 signifying a perfect fit and 0 indicating that the model fails to explain the variability in the data:

$$R^2 = \frac{\sum_1^n (yi - yi)^2}{\sum_1^n (yi - yi)^2} \tag{1}$$

Mean Squared Error (MSE)

The Mean Squared Error calculates the average of the squares of the differences between actual and predicted values. It provides a sense of the overall magnitude of prediction errors, with lower values indicating better model performance:

$$MSE = \frac{1}{n} \sum_{i=1}^n (yi - yi)^2 \tag{2}$$

Root Mean Squared Error (RMSE)

The Root Mean Squared Error is the square root of the Mean Squared Error. It represents the average magnitude of prediction errors in the same units as the target variable, making it more interpretable:

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (yi - yi)^2} \tag{3}$$

Results and Discussion

Prediction Results

Our study leveraged various machine learning (ML) algorithms to predict profitable crops based on input parameters. The experimental outcomes, summarized in Table 3, encompass scores for k-fold cross-validation, recall, accuracy, and precision diverse predictive capabilities of these ML algorithms (Sartore et al., 2022) In thoroughly examining Machine Learning (ML) algorithms for crop prediction that yields lucrative results, Naive Bayes (NB) is the clear winner in terms of all essential criteria. At 99.50%, NB's accuracy score is remarkable, outperforming rivals such as Random Forest (RF) with 99.41 Support Vector Machine (SVM) with 98.45, Decision Tree (DT) with 98.45, and Logistic Regression (LR) with 97.09 Moreover, NB surpasses RF, SVM, DT, and LR in all categories with outstanding precision (99.51), recall (99.55), and F1 Score (99.48%). Naive Bayes is the best algorithm for predicting the most lucrative crops since it consistently outperforms other algorithms in various criteria, demonstrating its dependability and firm performance in agricultural forecasting.

A visual depiction of the accuracy comparison of all machine learning models used in our investigation can be seen in Figure 8.

DOI: 10.3844/jcssp.2026.1045.1059

Table 3: Machine Learning (ML) Model for Crop Recommendation System

ML Model	Description	Strengths	Limitations	Best Use Cases	In Crop Recommendation
(DT)	Splits data by feature values	Easy, interpretable	Overfits, unstable	Small-medium datasets	Simple rules: Soil, climate, yield
(RF)	Ensemble of decision trees	Accurate, robust	Less interpretable, resource-heavy	Large/complex data	More accurate & stable predictions
SVM	Finds optimal boundary	Works in high dimensions	Needs kernel tuning, slow	Small–medium, complex classes	Binary crop choice (Crop A vs B)
kNN	Uses nearest neighbors	Simple, no training	Slow for big data, noise-sensitive	Small, clear patterns	Crop recommendation by similarity
ANN	Multi-layer networks	Captures non-linear relations	Needs big data, tuning	Complex tasks, large data	Models soil–weather–crop interactions
DL	CNNs, RNNs for complex tasks	Handles massive data	High resources, less interpretable	Images, time-series	Crop health (satellite), yield forecast
NB	Probabilistic, assumes independence	Fast, works with small data	Independence assumption unrealistic	Simple classification	Quick suggestions for small data
GBM	Sequential boosting models	High accuracy, robust	Slow, tuning needed	Large data, accuracy focus	Refined, accurate crop predictions

```
Accuracy: 0.99
Classification Report:
              precision    recall  f1-score   support

       apple       1.00      1.00      1.00        23
      banana       1.00      1.00      1.00        21
   blackgram       0.95      1.00      0.98        20
    chickpea       1.00      1.00      1.00        26
     coconut       0.96      1.00      0.98        27
      coffee       1.00      1.00      1.00        17
      cotton       1.00      1.00      1.00        17
      grapes       1.00      1.00      1.00        14
        jute       0.92      0.96      0.94        23
 kidneybeans       1.00      1.00      1.00        20
      lentil       0.92      1.00      0.96        11
       maize       1.00      0.95      0.98        21
       mango       1.00      1.00      1.00        19
   mothbeans       1.00      0.92      0.96        24
    mungbean       1.00      1.00      1.00        19
   muskmelon       1.00      1.00      1.00        17
      orange       1.00      1.00      1.00        14
      papaya       1.00      1.00      1.00        23
  pigeonpeas       1.00      1.00      1.00        23
 pomegranate       1.00      1.00      1.00        23
        rice       0.94      0.89      0.92        19
  watermelon       1.00      1.00      1.00        19

    accuracy                           0.99       440
   macro avg       0.99      0.99      0.99       440
weighted avg       0.99      0.99      0.99       440
```

Fig. 7: Crop Data Statistics

The Decision Tree (DT) method performed the worst out of all the machine learning models we tested. However, the Naive Bayes (NB) and Random Forest (RF) algorithms performed better than the other models. Regarding accuracy, precision, recall, and F1 score, these two algorithms NB and RF outperform others, demonstrating their efficacy in forecasting lucrative crops given the inputs at hand. Figure 10 shows the pie chart of the Accuracy of Different Classifiers.

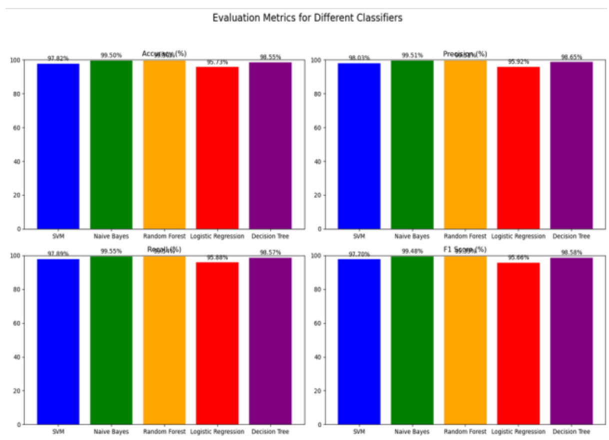

Fig. 8: Evaluation Matrix for Different Classifiers

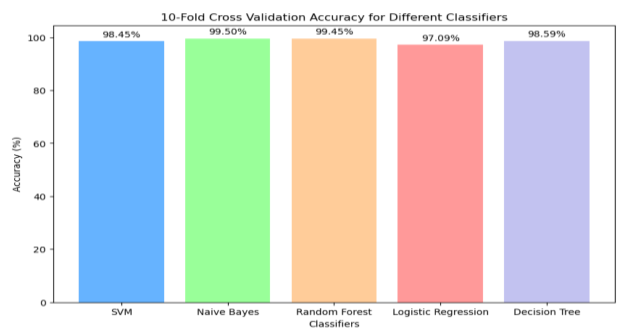

Fig. 9: Fold Cross Validation Accuracy for Different Classifiers

The Naive Bayes (NB) algorithm stands out as an exceptional performer, leading the pack with an outstanding F1 score of 99.48, emphasizing its ability to strike a harmonious balance between accuracy and recall.

DOI: 10.3844/jcssp.2026.1045.1059

Following closely is the Random Forest (RF) algorithm, securing an impressive F1 score of 99.34, reinforcing its efficacy in making accurate optimistic predictions while correctly identifying actual positives. Support Vector Machine (SVM) performs admirably, achieving an F1 score of 98.37 showcasing its strong performance in the evaluation metrics. Logistic Regression (LR) also contributes significantly, with a remarkable F1 score of 96.95, underscoring its commendable performance. The Decision Tree (DT) algorithm also showcases robust performance with an F1 score of 98.43%. Figure 8 and Table 3 visually represent the Evaluation Matrix for Different Classifiers, providing a comprehensive illustration of the models' performance across various metrics.

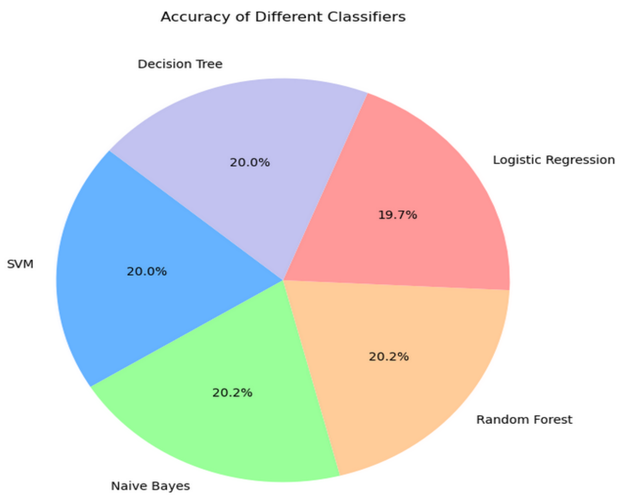

Fig. 10: Pie-chart of Accuracy of Different Classifier

While recognizing that accuracy alone is insufficient as a comprehensive performance indicator, a holistic evaluation of additional assessment metrics affirms that Naive Bayes (NB) outperforms the other machine learning methods, particularly in forecasting optimal harvests. Figure 8 and Table 3 illustrate the 10-fold cross-validation accuracy for different classifiers, visually representing the model's performance in this comprehensive evaluation.

Crop Recommendation (CR) Solution

According to our analysis, the market for free and open-source precision agricultural solutions is dynamic and needs significant developments to rival proprietary options. Despite these obstacles, the involvement of a cooperative network of farmers and IT experts have been crucial in promoting the development of open-source substitute (Sadeghi and Rahmati, 2017). This group effort guarantees a flexible environment sensitive to difficulties arising from integrating technology. From this angle, our primary goal was to launch a web- and mobile-based crop-recommendation system. This method uses machine learning algorithms to forecast the best crop to harvest while accounting for critical environmental factors, including soil nutrient levels, temperature, humidity, and rainfall.

The thorough performance analysis shows that Naïve Bayes consistently outperformed other tested machine learning models. As such, we want to utilize the robust features of Naive Bayes (NB) and Random Forest (RF) to forecast the ideal crop according to user-input data, such as pH levels, temperature, humidity, rainfall, and potassium (K) and nitrogen (N). After determining which model performed the best, we serialized and saved the model using pickle, Python's built-in persistence mechanism. Our crop recommendation algorithm is built on top of this serialized model. Figure 10 explains the complex processes for designing and implementing the mobile/web-based crop recommendation system.

Flask, a Python-based web framework designed for application development, was used in the system development process to facilitate the creation of Python Application Programming Interfaces (APIs). HTML (Hypertext Markup and Language) and CSS (Cascading Style Sheets) were used to create interactive websites for user interaction. The complete recommendation platform came to life using the Flask micro web framework. The recommendation platform is shown in Figure 11, a demonstration version currently hosted locally. Continuous attempts are being made to improve the accuracy of the platform.

The operational system and its corresponding development environment are accessible in an online directory (Sartore et al., 2022). This repository contains all the crucial code files needed for the frontend source code, prediction, training, and recommendation system application. Figure 12 illustrates the manual entry of field data acquired at the front end. It serves as a demo screenshot, providing an overview of the available features of the Implemented Crop Recommendation System (CRS). Notably, ongoing development efforts are directed towards fully implementing the web-based application platform for crop suggestions by farmers, as exemplified in the mentioned figure.

The analysis of Machine Learning (ML) model performance during training revealed that Naive Bayes (NB) consistently outperformed other models, including Random Forest (RF), across various metrics like recall, precision, and several performance parameters, despite the modest training dataset of 2200 entries. Consequently, NB was chosen as the preferred ML technique for constructing the recommendation system. Upon inputting the necessary parameters, the platform conducts verification checks to ensure entries comply with permissible criteria and searches for missing data.

DOI: 10.3844/jcssp.2026.1045.1059

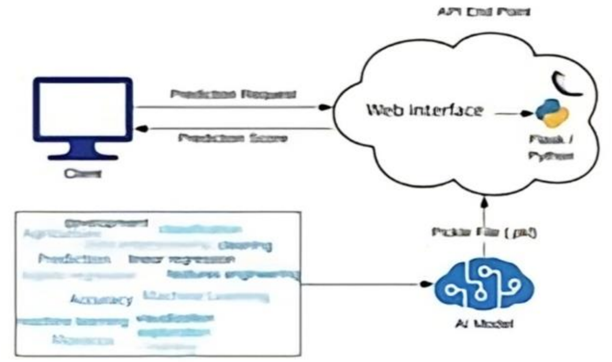

Fig 11: Steps of designing and implementing crop recommendation mobile/web-based system

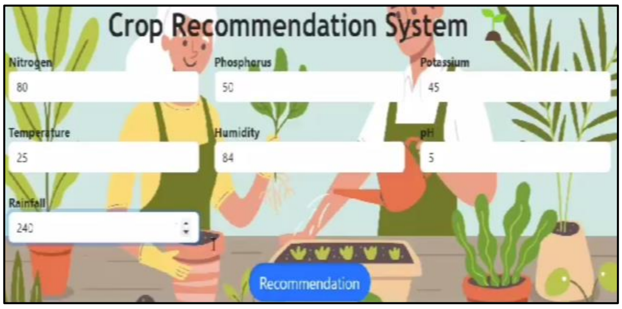

Fig. 12: Implemented crop recommendation system (CRS)

Subsequently, the algorithm employs these input characteristics to predict the most suitable crop for cultivation. This functionality empowers farmers to make informed decisions, reducing waste and optimizing returns on investment. In developing the crop recommendation system, the assumption was made that users would leverage existing meteorological data for parameter entry. The web-based nature of the platform allows users to access it conveniently from any location using any device. While acknowledging that the system was trained on Wardha (Maharashtra) data, with the understanding that unique circumstances in different locations may impact crop growth, there is a recognition of the need for more diverse data to extend the system's global reach. As the system evolves, it may aggregate data from various geographic regions, enabling users to input location-specific information for more accurate crop forecasts. This study lays the groundwork for future research in combining machine learning with robotic technology to create intelligent solutions for agriculture. Addressing the broader economic perspective, adequate agricultural investment involves prudent crop selection for optimal output. The proposed solution aims to assist farmers in making informed decisions tailored to their specific land and environmental conditions. With ongoing enhancements, including incorporating additional characteristics and critical indicators from diverse locations, the technology is envisioned to benefit a broader spectrum of farmers interested in adopting precision farming. In contrast to many companies and startups offering precision agricultural solutions with subscription-based models, the emphasis here is on the availability of free and open-source solutions for farmers. Farmers must know these alternatives, enabling them to embrace technology without significant upfront costs.

Conclusion and Future Scope

In conclusion, agriculture is essential in supporting the world's population, and the incorporation of technology is revolutionizing conventional farming practices to provide higher yields, better harvests, and less physical work. Data-based decision-making for farmers is made more accessible by the growing popularity of precision farming, which AI and ML drive. Our work presents a free and open-source alternative to for-profit solutions: A web-based machine learning-driven crop recommendation system. Our system's architecture demonstrates how Machine Learning (ML) and precision agriculture may be used to improve crop management. The model that performed the best was the Naïve Bayes algorithm, which had an astounding accuracy rate of 99.50%. Our research is to give Wardha (Maharashtra) farmers an open-source precision farming platform, encouraging increased agricultural output and sustained economic growth despite obstacles like the requirement for field validation and insufficient data. The determined algorithmic correctness highlights the resilience of our suggested system, with Random Forest emerging as the most effective choice for crop recommendation in the specified environmental context.

Further, real-time monitoring capabilities might be included in system developments by working with current data aggregation systems from sensors and the Internet of Things (IoT). This development would lead to a prediction framework that is more accurate and flexible, increasing precision farming's reach on a worldwide scale. In order to address the changing issues of feeding the world's expanding population, future versions may investigate the integration of Soil micronutrient data and additional climate characteristics for increased accuracy and improved crop yields.

Acknowledgment

We appreciate the anonymous reviewers thorough review and insightful suggestions, which greatly improved our work.

Funding Information

The authors did not receive any grants or funding for this study.

Author's Contributions

Aishwarya Kadu: Participated in all the experiments, including data collected and analysis. Written, reviewed, and edited.

KTV Reddy: Participated in data curation, investigation, and validation.

Ethics

The content presented here is the author's original research and has not been published elsewhere.

References

Abu, N. S., Bukhari, W. M., Ong, C. H., Kassim, A. M., Izzuddin, T. A., Sukhaimie, M. N., Norasikin, M. A., & Rasid, A. F. A. (2022). Internet of Things Applications in Precision Agriculture: A Review. *Journal of Robotics and Control (JRC)*, *3*(3), 338–347. https://doi.org/10.18196/jrc.v3i3.14159

Ahmed, U., Lin, J. C.-W., Srivastava, G., & Djenouri, Y. (2021). A nutrient recommendation system for soil fertilization based on evolutionary computation. *Computers and Electronics in Agriculture*, *189*, 106407. https://doi.org/10.1016/j.compag.2021.106407

Arshaghi, A., Ashourian, M., & Ghabeli, L. (2023). Potato diseases detection and classification using deep learning methods. *Multimedia Tools and Applications*, *82*(4), 5725–5742. https://doi.org/10.1007/s11042-022-13390-1

Baltrušaitis, T., Ahuja, C., & Morency, L. (2018). Multimodal machine learning: A survey and taxonomy. *Foundations and Trends in Machine Learning*, *41*(2), 423–443.

Bechar, A., & Vigneault, C. (2016). Agricultural robots for field operations: Concepts and components. *Biosystems Engineering*, *149*, 94–111. https://doi.org/10.1016/j.biosystemseng.2016.06.014

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, *21*(11), 3758. https://doi.org/10.3390/s21113758

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., & Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, *274*, 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010

Cao, J., Zhang, Z., Tao, F., Zhang, L., Luo, Y., Han, J., & Li, Z. (2020). Identifying the Contributions of Multi-Source Data for Winter Wheat Yield Prediction in China. *Remote Sensing*, *12*(5), 750. https://doi.org/10.3390/rs12050750

Cedric, L. S., Adoni, W. Y. H., Aworka, R., Zoueu, J. T., Mutombo, F. K., Krichen, M., & Kimpolo, C. L. M. (2022). Crops yield prediction based on machine learning models: Case of West African countries. *Smart Agricultural Technology*, *2*, 100049. https://doi.org/10.1016/j.atech.2022.100049

Chen, Y.-A., Hsieh, W.-H., Ko, Y.-S., & Huang, N.-F. (2021). An Ensemble Learning Model for Agricultural Irrigation Prediction. *Proceeding of the International Conference on Information Networking (ICOIN)*, 311–316. https://doi.org/10.1109/icoin50884.2021.9333852

Fan, C., Chen, M., Wang, X., Wang, J., & Huang, B. (2021). A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data. *Frontiers in Energy Research*, *9*, 652801. https://doi.org/10.3389/fenrg.2021.652801

Fei, S., Hassan, M. A., Xiao, Y., Su, X., Chen, Z., Cheng, Q., Duan, F., Chen, R., & Ma, Y. (2023). UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat. *Precision Agriculture*, *24*(1), 187–212. https://doi.org/10.1007/s11119-022-09938-8

Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L. B., Obersteiner, M., & van der Velde, M. (2016). Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. *Nature Communications*, *7*(1), 11872. https://doi.org/10.1038/ncomms11872

Forkuor, G., Hounkpatin, O. K. L., Welp, G., & Thiel, M. (2017). High Resolution Mapping of Soil Properties Using Remote Sensing Variables in South-Western Burkina Faso: A Comparison of Machine Learning and Multiple Linear Regression Models. *PLOS ONE*, *23*(1), e0170478. https://doi.org/10.1371/journal.pone.0170478

Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23*, 272–288. https://doi.org/10.1016/j.rser.2013.03.004

Fuentes, A., Yoon, S., Kim, S., & Park, D. (2017). A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition. *Sensors*, *17*(9), 2022. https://doi.org/10.3390/s17092022

Gasmi, A., Gomez, C., Chehbouni, A., Dhiba, D., & El Gharous, M. (2022). Using PRISMA Hyperspectral Satellite Imagery and GIS Approaches for Soil Fertility Mapping (FertiMap) in Northern Morocco. *Remote Sensing*, *14*(16), 4080. https://doi.org/10.3390/rs14164080

Gebbers, R., & Adamchuk, V. I. (2010). Precision Agriculture and Food Security. In *Science* (Vol. 327, Issue 5967, pp. 828–831). https://doi.org/10.1126/science.1183899

Gondia, A., Siam, A., El-Dakhakhni, W., & Nassar, A. H. (2020). Machine Learning Algorithms for Construction Projects Delay Risk Prediction. *Journal of Construction Engineering and Management*, *146*(1). https://doi.org/10.1061/(asce)co.1943-7862.0001736

He, Y., Zeng, H., Fan, Y., Ji, S., & Wu, J. (2019). Application of Deep Learning in Integrated Pest Management: A Real-Time System for Detection and Diagnosis of Oilseed Rape Pests. *Mobile Information Systems*, *2019*, 1–14. https://doi.org/10.1155/2019/4570808

Huang, H., Lan, Y., Yang, A., Zhang, Y., Wen, S., & Deng, J. (2020). Deep learning versus Object-based Image Analysis (OBIA) in weed mapping of UAV imagery. *International Journal of Remote Sensing*, *41*(9), 3446–3479. https://doi.org/10.1080/01431161.2019.1706112

Javaid, M., Haleem, A., Khan, I. H., & Suman, R. (2023). Understanding the potential applications of Artificial Intelligence in Agriculture Sector. *Advanced Agrochem*, *2*(1), 15–30. https://doi.org/10.1016/j.aac.2022.10.001

Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K.-M., Gerber, J. S., Reddy, V. R., & Kim, S.-H. (2016). Random Forests for Global and Regional Crop Yield Predictions. *PLOS ONE*, *11*(6), e0156571. https://doi.org/10.1371/journal.pone.0156571

Jiang, J., Atkinson, P. M., Zhang, J., Lu, R., Zhou, Y., Cao, Q., Tian, Y., Zhu, Y., Cao, W., & Liu, X. (2022). Combining fixed-wing UAV multispectral imagery and machine learning to diagnose winter wheat nitrogen status at the farm scale. *European Journal of Agronomy*, *138*, 126537. https://doi.org/10.1016/j.eja.2022.126537

Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., Godfray, H. C. J., Herrero, M., Howitt, R. E., Janssen, S., Keating, B. A., Munoz-Carpena, R., Porter, C. H., Rosenzweig, C., & Wheeler, T. R. (2017). Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. *Agricultural Systems*, *155*, 269–288. https://doi.org/10.1016/j.agsy.2016.09.021

Kadu, A. V., & Reddy, K. T. V. (2023a). Enhancing Crop Yield Prediction and Monitoring: Current Developments, Limitations, and Future Directions–A Review. *Journal of Technical Education*.

Kadu, A., & Reddy, K. T. V. (2023b). Enhancing Crop Prediction Accuracy Via Soil and Environmental Feature Selection Techniques in Machine Learning. *2023 1st DMIHER International Conference on Artificial Intelligence in Education and Industry 4.0 (IDICAIEI)*, 1–6. https://doi.org/10.1109/idicaiei58380.2023.10406703

Kadu, A. V., & Reddy, K. T. V. (2024a). Agriculture Yield Forecasting via Regression and Deep Learning with Machine Learning Techniques. *Communication and Intelligent Systems*, 219–233. https://doi.org/10.1007/978-981-97-2079-8_18

Kadu, A. V., & Reddy, K. T. V. (2024b). Empowering Agriculture: Smart Crop Recommendation with Machine Learning Challenges and Future Prospects. *Recent Advances in Science, Engineering & Technology*, 176–182. https://doi.org/10.1201/9781003527442-20

Kadu, A. V., & Reddy, K. T. V. (2025). A Comprehensive Survey of Machine Learning Applications in Agricultural Soil Nutrient Analysis, Fertility Prediction, and Soil Mapping in Wardha District, India. *ICT Analysis and Applications*, 203–215. https://doi.org/10.1007/978-981-97-9529-1_19

Kadu, A., & Reddy, K. T. V. (2024c). Comprehensive analysis of deep and machine learning approaches for predicting crop yields. *AIP Conference Proceedings*, *3188*(1), 080002. https://doi.org/10.1063/5.0240204

Kadu, A., Reddy, K. T. V., & Gawande, U. (2024). Innovative AgTech: A Predictive Machine Learning-Driven Precision Farming Solution for Enhancing Agricultural Productivity in Wardha. *Proceeding of the DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI)*, 1–6. https://doi.org/10.1109/idicaiei61867.2024.10842778

Krasilnikov, P., Taboada, M. A., & Amanullah. (2022). Fertilizer Use, Soil Health and Agricultural Sustainability. *Agriculture*, *12*(4), 462. https://doi.org/10.3390/agriculture12040462

Li, D., Miao, Y., Ransom, C. J., Bean, G. M., Kitchen, N. R., Fernández, F. G., Sawyer, J. E., Camberato, J. J., Carter, P. R., Ferguson, R. B., Franzen, D. W., Laboski, C. A. M., Nafziger, E. D., & Shanahan, J. F. (2022). Corn Nitrogen Nutrition Index Prediction Improved by Integrating Genetic, Environmental, and Management Factors with Active Canopy Sensing Using Machine Learning. *Remote Sensing*, *14*(2), 394. https://doi.org/10.3390/rs14020394

Liu, Y., Ma, X., Shu, L., Hancke, G. P., & Abu-Mahfouz, A. M. (2020). From Industry 4.0 to Agriculture 4.0: Current status, enabling technologies, and research challenges. *IEEE Transactions on Industrial Informatics*, *17*(6), 4322–4334. https://doi.org/10.1109/TII.2020.3003910

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, *13*(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

Maniruzzaman, Md., Rahman, Md. J., Ahammed, B., & Abedin, Md. M. (2020). Classification and prediction of diabetes disease using machine learning paradigm. *Health Information Science and Systems*, *8*(1), 1–7. https://doi.org/10.1007/s13755-019-0095-z

Mendez, K. M., Broadhurst, D. I., & Reinke, S. N. (2020). Migrating from partial least squares discriminant analysis to artificial neural networks: a comparison of functionally equivalent visualisation and feature contribution tools using jupyter notebooks. *Metabolomics*, *16*(2), 1–17. https://doi.org/10.1007/s11306-020-1640-0

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Pande, S. M., Ramesh, P. K., Anmol, A., Aishwarya, B. R., Rohilla, K., & Shaurya, K. (2021). Crop Recommender System Using Machine Learning Approach. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 1066–1071. https://doi.org/10.1109/iccmc51019.2021.9418351

Patel, J., Amipara, C., Ahanger, T. A., Ladhva, K., Gupta, R. K., Alsaab, H. O., Althobaiti, Y. S., & Ratna, R. (2022). A Machine Learning-Based Water Potability Prediction Model by Using Synthetic Minority Oversampling Technique and Explainable AI. *Computational Intelligence and Neuroscience*, *2022*, 1–15. https://doi.org/10.1155/2022/9283293

Patel, K., & Patel, H. B. (2020). A state-of-the-art survey on recommendation system and prospective extensions. *Computers and Electronics in Agriculture*, *178*, 105779. https://doi.org/10.1016/j.compag.2020.105779

Paymode, A. S., & Malode, V. B. (2022). Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG. *Artificial Intelligence in Agriculture*, *6*, 23–33. https://doi.org/10.1016/j.aiia.2021.12.002

Popescu, D., Stoican, F., Stamatescu, G., Ichim, L., & Dragana, C. (2020). Advanced UAV–WSN System for Intelligent Monitoring in Precision Agriculture. *Sensors*, *20*(3), 817. https://doi.org/10.3390/s20030817

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., & Hughes, D. P. (2017). Deep Learning for Image-Based Cassava Disease Detection. *Frontiers in Plant Science*, *8*, 1852. https://doi.org/10.3389/fpls.2017.01852

Raschka, S. (2018). Model Selection, and Algorithm Selection in Machine Learning. *J. Open Source Software*, *3*(26), 969. https://doi.org/arXiv

Rehman, A., Saba, T., Kashif, M., Fati, S. M., Bahaj, S. A., & Chaudhry, H. (2022). A revisit of internet of things technologies for monitoring and control strategies in smart agriculture. *Agronomy*, *12*(1), 127. https://doi.org/10.3390/agronomy12010127

Rokade, A., Singh, M., Malik, P. K., Singh, R., & Alsuwian, T. (2022). Intelligent Data Analytics Framework for Precision Farming Using IoT and Regressor Machine Learning Algorithms. *Applied Sciences*, *12*(19), 9992. https://doi.org/10.3390/app12199992

Sadeghi, K., & Rahmati, T. (2017). Integrating assessment as, for, and of learning in a large-scale exam preparation course. *Assessing Writing*, *34*, 50–61. https://doi.org/10.1016/j.asw.2017.09.003

Sartore, L., Rosales, A. N., Johnson, D. M., & Spiegelman, C. H. (2022). Assessing machine leaning algorithms on crop yield forecasts using functional covariates derived from remotely sensed data. *Computers and Electronics in Agriculture*, *194*, 106704. https://doi.org/10.1016/j.compag.2022.106704

Senoo, E. E. K., Anggraini, L., Kumi, J. A., Karolina, L. B., Akansah, E., Sulyman, H. A., Mendonça, I., & Aritsugi, M. (2024). IoT Solutions with Artificial Intelligence Technologies for Precision Agriculture: Definitions, Applications, Challenges, and Opportunities. *Electronics*, *13*(10), 1894. https://doi.org/10.3390/electronics13101894