Journal of Computer Science

Research Article

# An Integrated Machine Learning Framework for Citrus Grading and Yield Prediction Using Classifier-Informed Regression

**Lydia Dede Obeng[1] and Isaac O. Osunmakinde[2]**

[1]*Department of Engineering, Norfolk State University, Norfolk, Virginia 23504, United States*
[2]*Department of Computer Science, Norfolk State University, Norfolk, Virginia 23504, United States*

*Article history*
Received: 06-06-2025
Revised: 26-01-2026
Accepted: 28-02-2026

**Corresponding Author:**
Isaac O. Osunmakinde
Department of Computer Science, Norfolk State University, Norfolk, Virginia 23504, United States
Email: ioosunmakinde@nsu.edu

**Abstract:** Agriculture remains a cornerstone of human sustenance, yet accurate citrus fruit grading and yield predictions remain significant challenges due to visual similarities between varieties such as oranges and grapefruits. This includes the inefficiencies of manual grading. This study presents a new dual-function machine learning system that integrates classification and regression to simultaneously grade citrus fruits and forecast crop yields. This study utilizes a citrus classification dataset comprising 10,000 samples (5,000 oranges and 5,000 grapefruits) characterized by 5 features (diameter, weight, and RGB color values), and a crop yield dataset spanning 1990-2022 from 39 countries across Africa (17 countries, 1,122 records) and the Americas (22 countries, 1,452 records) totaling 2,574 records with environmental and nutrient features. An Artificial Neural Network (ANN) classifier first distinguishes visually similar citrus types, which achieves a grading accuracy of 98.5% with minimal variance. This output then informs a Random Forest (RF) regressor, which predicts yields with a high degree of precision ($R^2 = 0.905$). Compared to existing methods, such as CNN-SVM and XGBoost-based approaches that achieve lower accuracy and $R^2$ scores, such as 90.6% and 0.853, respectively, the proposed system demonstrates superior performance across both tasks. The integrated ANN-RF pipeline architecture uses classifier predictions as informative features for the regression model. This integrated design aligns with real-world agricultural practices that demand concurrent quality control and forecasting with improved predictive reliability. The system's modular architecture allows adaptation across different geographies, reinforcing its relevance to precision agriculture and food security.

**Keywords:** Citrus Grading, Integrated Systems, Machine Learning, Precision Farming, Yield Forecast

## Introduction

Agriculture serves as a cornerstone of human society, fulfilling our fundamental food needs. Over the past centuries, advancements in science and technology, environmental and climatic shifts, rapid population growth, and competitive corporate practices have significantly impacted this longstanding practice (Bender et al., 2020). The recent pandemic further highlighted the challenge of maintaining a stable agricultural supply chain, necessitating responsive actions from governments worldwide (Hossain, 2020).

Smart agriculture solutions, leveraging data collection, transmission, and processing with computer technologies, have emerged to address these evolving demands and enhance agricultural activity and productivity (Pandey et al., 2022). Machine Learning (ML) plays a crucial role in optimizing agrarian practices, as ML algorithms can identify patterns and relationships within datasets to address challenges such as disease and weed detection, crop quality assessment, and yield prediction. Fig. 1 shows the plot of the orange crop yield from 1990 to 2022 (FAOSTAT, 2025). The plot indicates that most crop yields do not always have a positive increment across the years.

SCIENCE Publications

© 2026 Lydia Dede Obeng and Isaac O. Osunmakinde. This open-access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

**DOI:** 10.3844/jcssp.2026.1344.1359

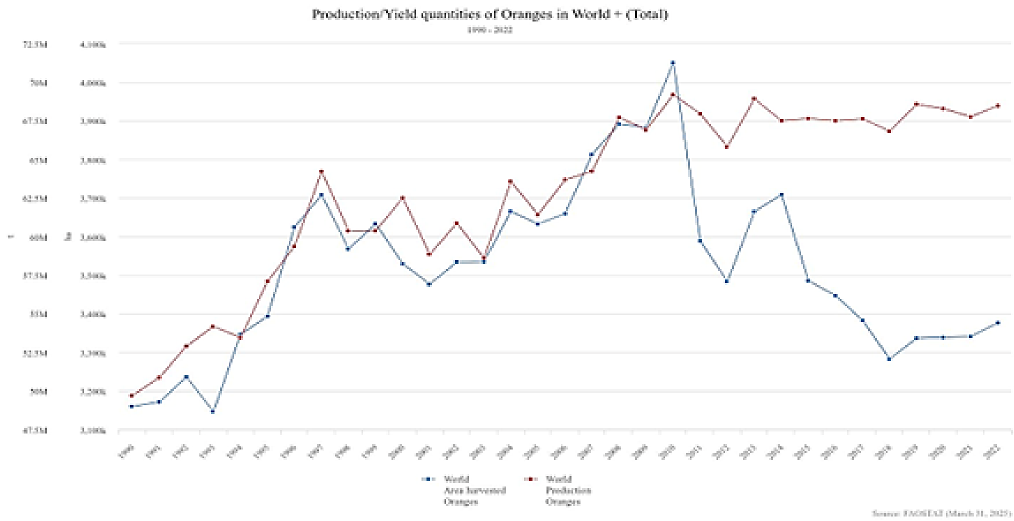



**Fig. 1:** Average Yield Plot of Oranges from 1990 to 2022 across the World (FAOSTAT, 2025)

Crop yields are influenced by several factors, including agriculture and environmental practices. This variability in yields highlights the need for crop yield predictions, mainly to help farmers make informed decisions about planting, harvesting, and potentially insurance.

Crop grading in farming involves sorting and classifying crops based on size, quality, ripeness, or variety. It directly influences market value, reducing waste and ensuring that produce meets consumer and industry standards. The complexity of crop grading arises from the subtle similarities between certain fruits and vegetables, which can confuse even experienced farmers. A notable example is the resemblance between oranges and grapefruits shown in Fig. 2, both citrus fruits that share similar shapes, colors, and textures, especially at certain stages of ripeness. Misidentification during manual sorting can lead to economic losses, as produce may fail to meet buyer expectations or disrupt processing lines. Such challenges highlight the limitations of traditional and labor-intensive methods that rely on human judgment.

Studies have explored the application of machine learning in agriculture, focusing on crop classification and yield prediction.

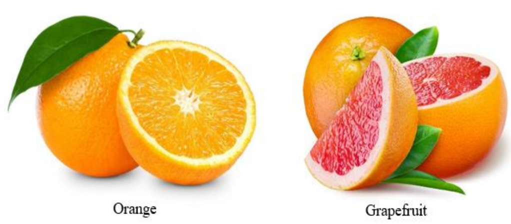



**Fig. 2:** Oranges and Grapefruit Similarity

Researchers have investigated crop yield prediction using various machine learning models, including Random Forest (RF), Support Vector Regression (SVR), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Auto Arima, demonstrating the potential of these techniques for yield forecasting (Özden and Karadoğan, 2024). Other studies have utilized machine learning regression models to predict crop yields based on meteorological parameters and pesticide data, highlighting the importance of considering environmental factors in yield prediction models (Burhan, 2022). These studies have laid the groundwork for this research, which advances crop intelligence systems by introducing a dual-function framework that integrates classification and regression. While previous approaches have contributed significantly to crop classification and yield prediction, they often treat these tasks independently, failing to leverage the relationship between crop grade and yield. Additionally, performance metrics, including accuracy or correlation scores in previous studies, indicate the need for improvement and exploration of alternative machine learning algorithms (Reddy et al., 2024; Sugumar et al., 2019). Addressing these limitations is crucial for advancing smart agriculture practices and ensuring food security.

Motivated by existing challenges in treating crop classification and yield prediction as isolated tasks, this research asks: How can a machine learning framework that integrates classification and regression better predict citrus yield across different countries? The main contributions of this study are:

- Design and implementation of an integrated machine learning framework where the output of a citrus classifier informs a yield prediction regressor, which enables coordinated grading and prediction

- High model performance and reliability, with the ANN classifier achieving 98.5% accuracy (±0.0033) for citrus grading, and the RF regressor delivering an $R^2$ of 0.905 (±0.025) for yield prediction
- Validation of generalizability through cross-country datasets ensures the model performs well across diverse agricultural conditions and supports broader applications
- Comparative and modular evaluation, demonstrating improved accuracy, efficiency, and interpretability over baseline models. It offers adaptability for other crop-specific contexts in advancing smart farming

## *Review of Literature*

The application of machine learning in agriculture, particularly for crop yield prediction and classification, has been a subject of interest to many researchers. Various studies have explored different methodologies and algorithms to address challenges in yield prediction and crop classification tasks. Table 1 summarizes key papers in this field, highlighting their focus, methodologies, and suggestions for future work.

## *Research Gaps*

Despite extensive research on crop classification and yield prediction using machine learning, existing studies typically treat these tasks as independent problems. This separation fails to exploit the potential interdependence between crop grade (a classification problem) and yield (a regression task), which limits model efficiency and insight generation. While numerous models, such as RF, SVR, CNN, and LSTM, have been applied, the accuracy and generalizability of these models remain suboptimal. This is essential when evaluated on real-world agricultural datasets. Prior work also lacks practical integration of multi-stage models for decision support in agricultural systems. Addressing these gaps, this research proposes a new dual-function ML system where a classifier guides the regressor for citrus grading and yield prediction. This aims to improve performance, interpretability, and agricultural decision-making support.

## *Machine Learning Models*

In agricultural classification and regression tasks, common supervised ML models used include Decision Trees, Random Forests, and Artificial Neural Networks. DTs handle both numerical and categorical data well, while RFs build upon decision trees, providing improved accuracy and robustness through ensemble learning. Artificial Neural Networks excel at capturing complex nonlinear relationships in data, making them suitable for complex agricultural systems.

## *Decision Trees*

DT is a supervised ML model used for classification and regression tasks. The model is trained on a dataset that includes the target variable, which can be either categorical for classification trees or continuous for regression trees. The structure of a DT consists of a root node, branches, internal decision nodes, and leaf nodes, representing the outcomes (Suthaharan, 2016). Building a decision tree involves splitting the data based on specific criteria to maximize information gain or minimize impurity, such as entropy or Gini impurity. Entropy measures the level of disorder or uncertainty in a dataset. It is used to quantify the purity of a dataset before and after a split. For a data set *S,* the entropy *H(S)* is given as:

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

Where $p_i$ is the probability of samples belonging to class $i$ and $n$ is the number of classes. Information gain, based on the concept of entropy, measures the reduction after splitting an attribute. The Information Gain (IG) for an attribute *A* is calculated as:

**Table 1:** Review of Related Literature

| Paper Type | Problem Addressed | Methodology | Future Work Suggested |
|---|---|---|---|
| Research paper (Dahlkar and Rode, 2014) | Develop various crop yield prediction models using Artificial Neural Networks (ANNs) to accurately estimate crop production. | ANN | Exploration of other ML methods for comparison and enhancing the efficiency of crop prediction models. |
| Review paper (Van-Klompenburg et al., 2020) | Exploration of machine learning algorithms used for crop yield prediction | RF, NN, CNN, DNN, Linear Regression, Gradient Boosting Tree (GBT), and LSTM | Translating models into practical decision support tools for farmers and agricultural stakeholders. |
| Research Paper (Camara-Guerra et al., 2024) | Classification of citrus crops (orange, mandarin, grapefruit) using satellite multispectral imagery. | Deep Neural Network (DNN) | Improving the resolution of satellite images and exploring additional vegetation indices for better accuracy. |
| Research Paper (Chakhar et al., 2024) | Assessing crop classification accuracy for citrus plots using multi-source data. | Support Vector Machine (SVM) with fused Sentinel-1 SAR and Sentinel-2 NDVI data. | Integrating higher-resolution imagery and advanced classifiers to reduce overestimation. |

$$IG(S, A) = H(S) - \sum_{v \epsilon Values(A)} \frac{|S_v|}{|S|} H(S_v) \quad (2)$$

Where $S_V$ is the subset of $S$ for which attribute $A$ has a value $v$ A higher information gain indicates a better split that reduces uncertainty.

Another impurity metric is the Gini impurity or Gini index, which measures how often a randomly chosen sample from the dataset would be incorrectly classified if it were labeled according to the distribution of labels (Nowozin et al., 2011). Gini impurity is zero for a perfectly pure node where all samples belong to one class. The formula for Gini impurity *G(S)* is given as:

$$G(S) = 1 - \sum_{i=1}^{n} p_i^2 \quad (3)$$

Decision trees are easy to interpret and require little data preparation, but they can be prone to overfitting and biases. A common metric that attempts to overcome DT's bias towards attributes with many values is the Gain Ratio, calculated as:

$$Gain\ Ratio\ (S, A) = \frac{IG(S,A)}{Split\ Info(S,A)} \quad (4)$$

Where *Spilt Info (S, A)* is the entropy of the dataset based on the values of attribute *A*.

## *Random Forests*

RF is an ensemble learning algorithm that creates a 'forest' of decision trees and combines their predictions to produce a more accurate and robust model. It operates on the principle of bagging, which involves training each tree on a random subset of the data with replacement, thus ensuring diversity among the trees. This means that some instances in the original dataset may appear multiple times in a sample, while others may be excluded. RF can handle classification and regression problems and is known for managing complex datasets without overfitting (Genuer and Poggi, 2020).

From a dataset $S = \{(x_i, y_i)\}_{i=1}^{N}$, $M$ S = {($x_i$ $y_i$)} new datasets $S_1, S_2, \ldots, S_M$ are created by randomly sampling N instances with replacement. A decision model T*m* is used to train each sample *Sm*. For regression tasks, the final prediction *Y* is the average prediction from *M* models:

$$\hat{y} = \frac{1}{M} \sum_{m=1}^{M} T_m(x) \quad (5)$$

For classification tasks, majority voting is used:

$$\hat{y} = mode(\{T_1(x), T_2(x), \ldots, T_M(x)\}) \quad (6)$$

RF ensures stability since the result is based on majority voting or averaging of DTs (Cutler et al., 2012).

## *Artificial Neural Networks*

ANNs are inspired by the biological neural networks of the human brain. They consist of interconnected nodes or neurons that process information using a weighted approach. The structure of an ANN typically includes an input layer, one or more hidden layers, and an output layer. The output of a neuron in layer l is calculated as:

$$a_j^l = f\left(\sum_i w_{ij}^l a_i^{l-1} + b_j^l\right) \quad (7)$$

Where:

$w_{ij}^l$ is the weight connecting neuron *i* in layer l-1 to neuron *j* in layer l.
$a_i^{l-1}$ is the output of neuron *i* in layer l-1
$b_j^l$ is the bias of neuron *j* in layer l
$f$ is the activation function

Each neuron applies an activation function to its input to introduce non-linearity. Common activation functions include Sigmoid, Rectified Linear Unit (ReLU), and Tanh (Krenker et al., 2011; Walczak, 2018). The learning process in ANNs involves adjusting the weights and biases to minimize the cost function. Weights and biases are updated using gradient descent, which is given by:

$$w_{ij}^l = w_{ij}^l - \alpha \frac{\partial J}{\partial w_{ij}^l} \quad (8)$$

$$b_j^l = b_j^l - \alpha \frac{\partial J}{\partial b_j^l} \quad (9)$$

Where *a* is the learning rate. ANNs can handle various tasks, including classification, regression, and unsupervised learning. They are particularly effective for processing large and complex datasets with numerous features (Yang and Wang, 2020). The selection of an ANN classifier over Convolutional Neural Networks (CNN) or traditional feature-engineered models was guided by the nature of the citrus dataset. The dataset comprises structured tabular data with numerical features rather than raw image data. For structured data with limited features, fully connected neural networks can effectively learn non-linear decision boundaries while maintaining computational efficiency.

## *Data Analysis and Model Development*

The scheme of implementation for the ML dual-function system for citrus classifier and crop yield prediction is illustrated in Fig. 3. The intercorrelation between the classifier and regressor is established via Classifier-Informed Integration. The classifier functions as a high-level feature extractor; its output (Fruit Class) is transformed into a binary encoded feature. This feature is then concatenated with the environmental input vectors for the regressor. This ensures that the regression model does not predict yield in a vacuum but rather optimizes its prediction parameters based on the specific physiological and growth requirements of the identified citrus variety.

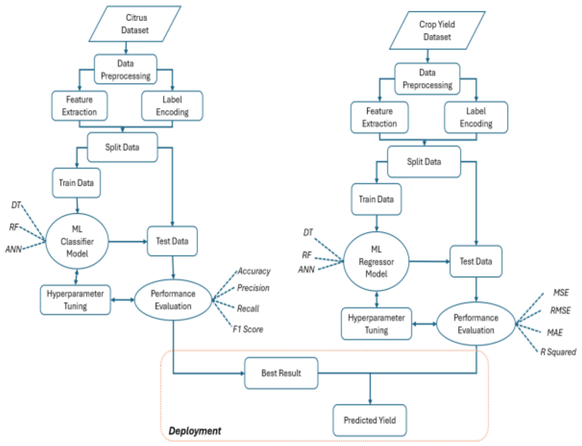



**Fig. 3:** Dual-Function Citrus Classifier and Yield Prediction

### *Data Acquisition and Preprocessing*

For the citrus classifier, two crops of similar features were selected: Oranges and grapefruits. These crops appear to have similar shapes, sizes, and colors at different stages of growth with little distinction in their features. Hence, it becomes a challenge when using traditional methods to distinguish oranges and grapefruits when harvested together. Grapefruit tends to be larger and heavier with pale yellow or pinkish peels, while oranges are smaller with orange-colored rinds (Grove, 2021). The dataset used for citrus grading was obtained from the Kaggle database (Kaggle, 2020), comprising 10,000 synthetic samples equally distributed between oranges (5,000) and grapefruits (5,000). The features used to classify oranges and grapefruit include weight, diameter, red, green, and Blue color values, and the target is the name of the crop. The synthetic nature of this dataset still captures realistic physical distinctions between citrus varieties. The features and targets for the citrus dataset are displayed in Fig. 4(a).

In predicting the crop yield, the data selected were from citrus-producing countries in Africa and the Americas. From 36 American countries initially available in FAOSTAT, 22 were selected based on complete data availability across all features for the study period (1990–2022); the remaining 14 countries were excluded due to gaps in yield reporting for one or both crop types. Similarly, from African countries available in FAOSTAT, 17 were selected based on complete data availability across all features for the study period (1990-2022). The features of the dataset include the year, country, crop type (oranges or grapefruit), average rainfall per year in mm, average temperature in °C, and cropland mineral fertilizers consisting of Potassium (K), Nitrogen (N), and Phosphorus (P) per unit area measured in kg/ha. The target data was the annual citrus yield from 1990 to 2022. The year, crop type, country, and NPK per unit area values were obtained from the Food and Agriculture Organization (FAO) database (FAOSTAT, 2025). Where nutrient parameters (N, P, K) were recorded as zero values, country-specific mean values calculated from non-zero observations were used for imputation to avoid bias in model training. The average rainfall and temperature values per year for each country were taken from the World Bank Climate Knowledge Portal (World Bank, 2025), which provides quality-controlled climate data derived from the CRU TS dataset produced by the University of East Anglia. The features and target variables of the yield prediction dataset are displayed in Fig. 4(b). The final yield dataset comprises 2,574 records, 1,122 for Africa (17 countries × 33 years × 2 crops) and 1,452 for America (22 countries × 33 years × 2 crops).

### *Model Development for Machine Learning Frameworks*

The algorithms for DT, RF, and ANN integrated classifier and regressor models that were used to design the system framework are shown in Algorithms 1, 2, and 3.

```
     name  diameter  weight  red  green  blue
0  orange      2.96   86.76  172     85     2
1  orange      3.91   88.05  166     78     3
2  orange      4.42   95.17  156     81     2
3  orange      4.47   95.60  163     81     4
4  orange      4.48   95.76  161     72     9
(10000, 6)

Dataset Features:
----------------
diameter: float64
weight: float64
red: int64
green: int64
blue: int64

Target Variable:
---------------
name: object
Unique target values: ['orange' 'grapefruit']
```

(a)

```
   Year  Country     Crop  Avg_Rainfall  Avg_Temperature       P       N  \
0  1990  Algeria  Oranges         90.87            23.53  1.7988  8.2515
1  1991  Algeria  Oranges         87.43            22.85  1.5332  5.2524
2  1992  Algeria  Oranges         90.38            22.83  1.5465  5.7697
3  1993  Algeria  Oranges         66.67            23.07  1.6934  9.1518
4  1994  Algeria  Oranges         83.60            23.44  1.7988  7.2695

        K   Yield
0  3.1526  7568.1
1  2.1266  9405.3
2  1.8663  9749.9
3  2.2953  9055.5
4  2.5296  9432.2
(1122, 9)

Dataset Features:
----------------
Year: int64
Country: object
Crop: object
Avg_Rainfall: float64
Avg_Temperature: float64
P: float64
N: float64
K: float64

Target Variable:
---------------
Yield: float64
```

(b)

**Fig. 4:** Features and Target of (a) Citrus Dataset (b) Crop Yield Dataset

**Algorithm 1** Integrated Decision Tree Framework

**Inputs:**

***Classifier Inputs:***
***$F_c$***: Feature matrix for citrus grading including diameter, weight, red, blue, and green values.
***L***: Target vector for citrus grade (oranges and grapefruit)
***Regressor Inputs:***
***$F_T$***: Feature matrix for yield prediction includes year, country, crop, average rainfall, average temperature, phosphorous, nitrogen, potassium)
***Y***: Target vector for crop Yield

**Outputs :**
- Predicted citrus grade
- Predicted crop yield based on classified citrus grade

**Steps:**
Initialize the root node with the dataset ($F_c$, $L$, $F_T$, $Y$)
*In the Classifier Model*
1. For each node
   - Evaluate all possible split points for features in $F_c$
   - Calculate entropy for each split and select the one that maximizes information gain.
2. Partition data based on the selected split and create child nodes.
3. Recursively apply steps 1-2 until a stopping condition is met (i.e., all samples belong to one grade or maximum depth is reached).
4. At leaf nodes, assign the majority grade of samples.
5. Construct the decision tree classifier.

*In the Regressor Model*
1. For each node
   - Evaluate all possible split points for features in $F_T$
   - Calculate mean squared error (MSE) for each split and select the one that minimizes MSE.
2. Partition data based on the selected split and create child nodes.
3. Recursively apply steps 2-3 until a stopping condition is met (i.e., maximum depth or minimal MSE reduction).
4. At leaf nodes, assign the mean yield value of samples.
5. Construct the decision tree regressor.

**Prediction:**
- Input new citrus features into the trained classifier tree.
- Traverse the tree based on the feature values to predict its grade.
- Use the predicted citrus grade as part of $F_{T,new}$ along with other features.
- Input $F_{T,new}$ into the trained regressor tree
- Traverse the tree to predict crop yield

**Algorithm 2** Integrated Random Forest Framework

**Inputs:**

***Classifier Inputs:***
***$F_c$***: Feature matrix for citrus grading
***L***: Target vector for citrus grade
***Regressor Inputs:***
***$F_T$***: Feature matrix for yield prediction
***Y***: Target vector for crop Yield

**Outputs :**
- Predicted citrus grade
- Predicted crop yield based on classified citrus grade

**Steps:**
Initialize the root node with the dataset ($F_c$, $L$, $F_T$, $Y$)
*In the Classifier Model*
1. Initialize an empty RF model
2. For $i$ =1 to $n_{trees}$, create a bootstrap sample by randomly sampling with replacement from ($F_c$, $L$)
3. Grow a decision tree $T_i$ as done in the DT classifier algorithm
4. Construct the random forest classifier by aggregating all $T_i$ into the RF model.

*In the Regressor Model*
1. Initialize an empty RF regressor model
2. For $i$ =1 to $n_{trees}$, create a bootstrap sample by randomly sampling with replacement from ($F_T$, $Y$)
3. Grow a decision tree $T_i$ as done in the DT regressor algorithm
4. Construct the RF regressor by aggregating all $T_i$ into the RF model.

**Prediction:**
- Input new citrus features into the trained classifier tree.
- Traverse each tree $T_i$ based on the feature values to predict a class (orange/grapefruit).
- Determine the final citrus grade via majority voting across all trees
- Use the predicted grade as part of $F_{T,new}$ along with other features.
- Input $F_{T,new}$ into the trained RF regressor tree
- Traverse each tree $T_i$ to predict a yield value
- Calculate the crop yield as the average of all tree predictions

**Algorithm 3** Integrated Artificial Neural Network Framework

**Inputs:**

***Classifier Inputs:***
***$F_c$***: Citrus feature matrix
***L***: Target vector for citrus grade
***Regressor Inputs:***
***$F_T$***: Crop yield feature matrix
***Y***: Target vector for crop Yield

**Outputs :**
- Predicted citrus grade
- Predicted crop yield based on classified citrus grade

**Parameters:**
•Hidden layers: Architecture of hidden layers
•Activation: Activation function
•Solver: Optimization algorithm
•Max_iter: Maximum training iterations

**Steps:**
Create an MLP classifier and regressor with the specified parameters
*In the Classifier Model*
1. Split $F_c$ and $L$ into training and test sets
2. For $i$ =1 to *max_iter:*
   - Forward Propagation: Compute outputs across all layers.
   - Loss Calculation: Use cross-entropy loss.
   - Backward Propagation: Update weights via gradient descent.
3. Store the loss for each iteration.

*In the Regressor Model*
1. Split $F_T$ and $Y$ into training and test sets
2. For $i$ =1 to *max_iter:*
   - Forward Propagation: Compute outputs across all layers.
   - Loss Calculation: Use mean squared error (MSE)
   - Backward Propagation: Update weights via gradient descent.
3. Store the loss for each iteration.

**Prediction:**
- Input new citrus features into the trained ANN classifier.
- Perform a forward pass to predict citrus grade.
- Use the predicted citrus grade as part of $F_{T,new}$ along with other features.
- Input $F_{T,new}$ into the trained ANN regressor.
- Perform a forward pass to predict crop yield

## *Hyperparameter Tuning and Performance Evaluation*

In training the citrus classifier and crop yield model, the dataset is split into training and test data. DT, RF, and ANN are used to train the data. For hyperparameter tuning, the following parameters are tuned for each ML algorithm. In DT, the tuned hyperparameter is the test size or split size of the data. RF hyperparameters are the number of estimators, which represent the different decision trees that make up the forest, and the split size of the data (Cutler et al., 2012). ANN allows for tuning the number of hidden layers and neurons in each layer, learning rate, activation functions, number of iterations or epochs, and the test size (Zhang and Li, 2019). N-fold cross-validation is conducted on the training dataset for all models to obtain a more reliable estimate of model performance. In the citrus and yield datasets, up to 5 partitions were used for cross-validation. Evaluating the performance of the model allows users to select the best parameters for maximum response. Classifier ML models have standard evaluation metrics for all the algorithms. Similarly, all regressor models have standard performance metrics. The performance metrics are analyzed and compared among the selected ML models.

## *Classification Model Evaluation*

The evaluation metric used in this research is the confusion matrix. It is a table that summarizes the model's predictions compared to the actual target values. The matrix displays the counts of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). These counts are used to derive accuracy, precision, recall/sensitivity, and F1-score. The equations for these metrics are given by (Zhang and Li, 2019):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (10)$$

$$Precision = \frac{TP}{TP+FP} \times 100 \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (12)$$

$$F1\_Score = \frac{2TP}{2TP+FP+FN} \times 100 \quad (13)$$

### Regression Model Evaluation

The regression model evaluation metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). MSE measures the squared differences between the predicted and actual values. RSME is the square root of MSE, and MAE calculates the average absolute difference between predicted and actual values. The equations are given as (Zhang and Li, 2019):

$$MSE = \frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2 \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \quad (15)$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \quad (16)$$

Where $y_j$ is the actual value, $\hat{y}_j$ is the predicted value, and n is the number of instances. These metrics can be used to determine the accuracy by calculating the % Error given as:

$$\%\ Error = \frac{MSE\ or\ RMSE\ or\ MAE}{Average\ of\ Target} \times 100 \quad (17)$$

$$Accuracy = 100\% - Error \quad (18)$$

Another important metric in the regressor model is the coefficient of determination *($R^2$)*. It measures how much of the observed variation in the data is accounted for by the model's predictions. $R^2$ provides an indication of how well the model predicts an outcome. It is calculated as:

$$R^2 = 1 - \frac{\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}{\sum_{j=1}^{n}(y_j - \bar{y}_j)^2} \quad (19)$$

Where $\bar{y}_j$ represents the mean of the actual values. $R^2$ values closer to 1 indicate a better model prediction than those closer to 0.

### Experimental Design

The model setup, data processing, and deployment were conducted using Jupyter Notebooks and managed through Anaconda Navigator. Python version 3 was the programming language of choice, with key libraries including scikit-learn for machine learning tasks, pandas for data manipulation, matplotlib and seaborn for visualization, and NumPy for numerical computations.

### Data Preprocessing Stage

The dataset for the crop yield spanned from the year 1990 to 2022, which is the latest information available on the FAO and World Bank databases (FAOSTAT, 2025; Kaggle, 2020). The crop yield dataset comprises two regional subsets: Africa with 1,122 entries (17 countries) and America with 1,452 entries (22 countries), both spanning 1990-2022 with 8 features including year, country, crop type, average rainfall, average temperature, and NPK nutrients. A total of 10000 citrus and five features were used in training the ML classifier models, 5000 for oranges and 5000 for grapefruits. Having a balanced portion for the two targets is essential for model performance in classification tasks. When one class significantly outnumbers the other, the model may become biased towards the majority class, and class balance techniques such as SMOTE may be used. The features selected for classifying orange and grapefruit targets are diameter, weight, red, green, and blue values, and crop name.

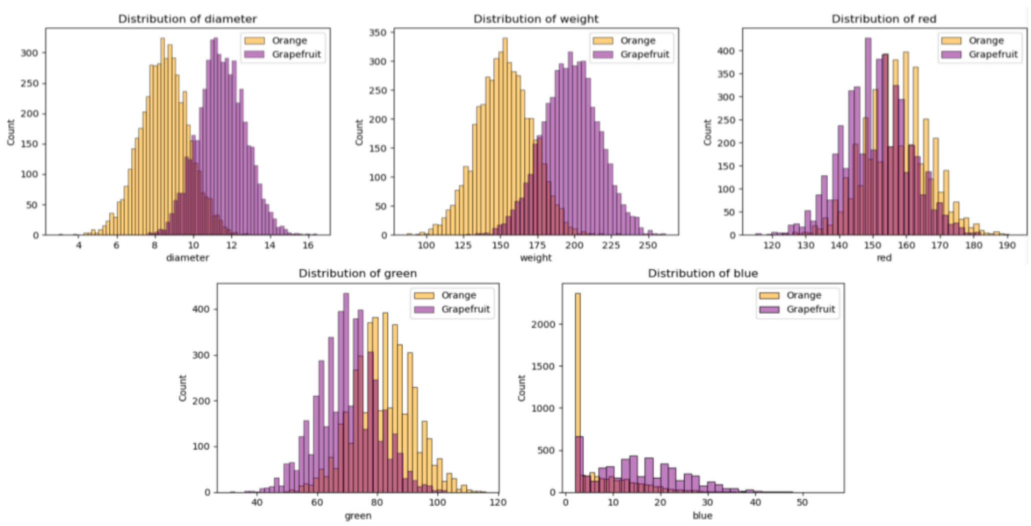



**Fig. 5:** Feature Distribution for Citrus Dataset

In the citrus dataset, all features but the crop name contained numerical values; the label encoder was used to assign integer values '0' and '1' for grapefruit and orange, respectively, as shown in Fig. 6(a). The crop type and country features in the yield dataset also contained categorical data, hence were encoded into numeric values. The encoded country data ranges from '0' to '16' as shown in Fig. 6(b).

The correlation between the features of the crop yield datasets is illustrated in Fig. 7(a). Heatmap values close to +1 or red indicate a positive correlation. Which means when one feature increases, the other also tends to increase. Negative correlation (close to -1 or blue) has the opposite effect; an increase in one tends to decrease the other. Yield has a positive correlation with P (0.35), N (0.24), and K (0.37). This means that higher levels of Phosphorus, Nitrogen, and Potassium are associated with higher crop yields, with K having the strongest positive effect among them. The year's feature indicates a slight upward trend in crop yield over time, as shown in Fig. 7(b), with a correlation value of 0.23. This is due to advances in farming techniques over the years. However, correlation does not imply causation; there might be other underlying factors at play.

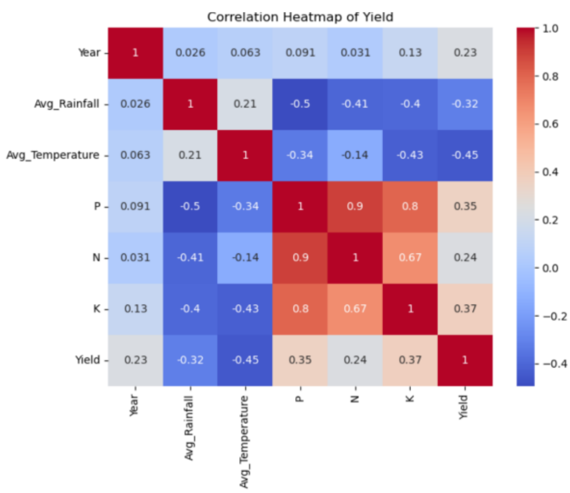
```
Encoded Crop Name Data:
            name  name_encoded
0         orange             1
1         orange             1
2         orange             1
3         orange             1
4         orange             1
...          ...           ...
9995  grapefruit             0
9996  grapefruit             0
9997  grapefruit             0
9998  grapefruit             0
9999  grapefruit             0
```

(a)

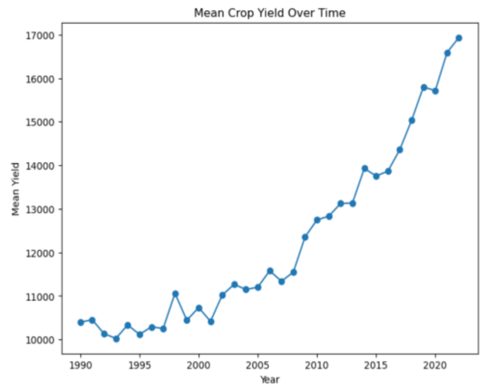
```
Encoded Country Data:
       Country  Country_enc
0      Algeria            0
1      Algeria            0
2      Algeria            0
3      Algeria            0
4      Algeria            0
...        ...          ...
1117  Zimbabwe           16
1118  Zimbabwe           16
1119  Zimbabwe           16
1120  Zimbabwe           16
1121  Zimbabwe           16
```

(b)

**Fig. 6:** (a) Encoded Crop Name for Citrus Classifier (b) Encoded Country data for Yield Prediction




(a)




(b)

**Fig. 7:** (a) Feature Correlation for Crop Yield Prediction (b) Mean Crop Yield from 1990 to 2022

## *Model Training and Setup*

For the citrus classification tasks, a dataset of 10000 entries was utilized, with 15% allocated as test data and 85% as training data. All features were included in the training process since they significantly contributed to the citrus target variable. The primary objective of this study was to test and evaluate the performance of different ML models in predicting orange or grapefruit based on diameter, weight, and RGB value. DT, RF, and ANN algorithms were employed for training and testing the datasets. In the RF model, aside from specifying a test size of 15%, the number of estimators was also varied. For the ANN model, additional hyperparameters tuned included the activation function, the number of hidden layers, the number of neurons in each layer, the learning rate, and the number of iterations. N-fold cross-validation was conducted on the training dataset to obtain a more reliable estimate of model performance. In the citrus dataset, up to five partitions were used for cross-validation, and the accuracy of each fold was determined and averaged.

For the crop yield datasets, 1122 entries were used, featuring variables such as year, country, crop type, average rainfall, average temperature, potassium, nitrogen, and crop yield as the target. The dataset was employed for training and testing the model. The aim of this regression task was to develop and evaluate the performance of ML models in predicting crop yields. The ML regressor algorithms used in this study were DT, RF, and ANN. The test size, number of estimators, number of hidden layers, activation function, and number of epochs were selected and varied to achieve maximum performance. The framework for crop yield was extended to countries in America, which have about 1452 entries in the datasets.

In the classification tasks, accuracy, precision, recall, and F1 score were calculated for each parameter combination for the model. The best response was selected and used to classify new datasets. For the crop yield prediction, MAE, MSE, and RMSE were used to evaluate the performance of the model. These parameters were used to calculate the model's accuracy and $R^2$ scores.

## Results and Discussion

### *Citrus Grading*

In citrus grading of oranges and grapefruit, the hyperparameters that produced optimal performances for different ML classification models are summarized in Table 2.

After tuning the hyperparameters, the accuracy, precision, F1-score, and recall values were obtained. Fig. 8 indicates the average of each performance metric obtained for DT, RF, and ANN models. From Fig. 8, the ANN achieved the highest accuracy of 98.5% compared to the RF and DT models, which had the same accuracy of 95.7%. The ANN model also performed better in other metrics compared to DT and RF. Hence, ANN is recommended for deployment since it has much higher reliability and classification success.

The classifier models were also evaluated through the n-fold cross-validation, with the outcome being the mean of the accuracy and the standard deviation. Table 3 shows the cross-validation values of the accuracy for 5 folds, the mean accuracy, and the standard deviation values. DT obtained a mean accuracy of 0.9498 with a standard deviation of 0.0086, followed by the RF model, which achieved a slightly higher mean accuracy of 0.9514 +/0.0028. ANN had the highest average value of 0.98 +/-0.0033 of all 5 folds.

**Table 2:** Tuning Parameters for ML Classifier Models

| ML Models | Hyperparameters | |
|---|---|---|
| DT, RF, ANN | Test Size | 15% |
| | N-fold Cross Validation | N=5 |
| RF | Estimators | 30 |
| ANN | Activation Function | ReLu |
| | Hidden Layer Number | 3 |
| | Maximum Iteration | 200 |

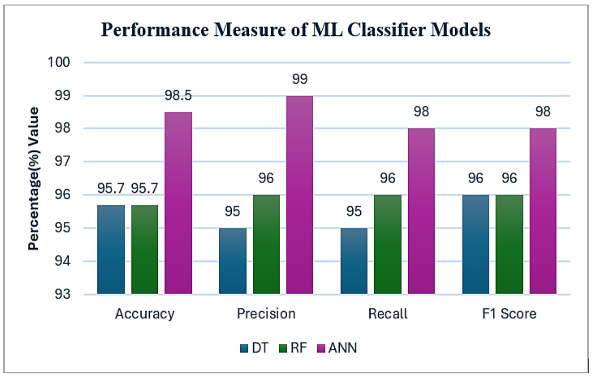



**Fig. 8:** Performance Measure of ML Classifier in Percentage

**Table 3:** N-fold Cross-Validation Values for Classifier Models

| Models | Accuracy Scores for 5 Folds | Mean Accuracy | Standard Deviation (+/-) |
|---|---|---|---|
| DT | [0.9415, 0.955, 0.9615, 0.9525, 0.9385] | 0.9498 | 0.0086 |
| RF | [0.9495, 0.954, 0.9475, 0.951, 0.955 ] | 0.9514 | 0.0028 |
| ANN | [0.981, 0.974, 0.98, 0.984, 0.981] | 0.98 | 0.0033 |

All models showed consistent performance across all folds, with the most deviation observed in the DT model.

The confusion matrix plot was obtained and evaluated for the different models to illustrate their actual and predicted values for orange and grapefruit targets. Table 4 shows the confusion matrix for all four ML models, with ANN achieving the highest correct prediction for both orange and grapefruit. Fig. 9 illustrates the confusion matrix plot for the ANN model. For a 15% test size, it means 1500 out of the 10000 samples are tested on the model, 743 grapefruits and 757 oranges.

**Table 4:** Confusion Matrix Plot of ML Models

| DT, RF, ANN | | Predicted | |
|---|---|---|---|
| | | Grapefruit | Orange |
| Actual | Grapefruit | 716 | 27 |
| | | 721 | 22 |
| | | 738 | 5 |
| | Orange | 37 | 720 |
| | | 42 | 715 |
| | | 18 | 739 |

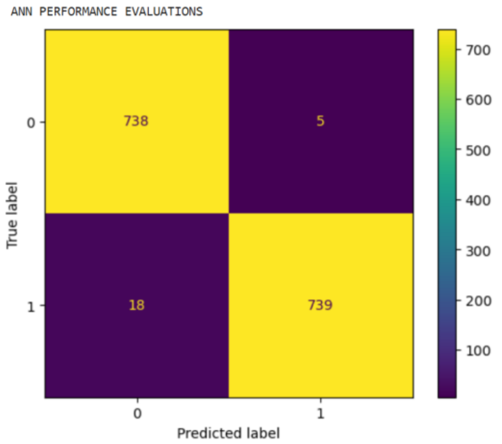


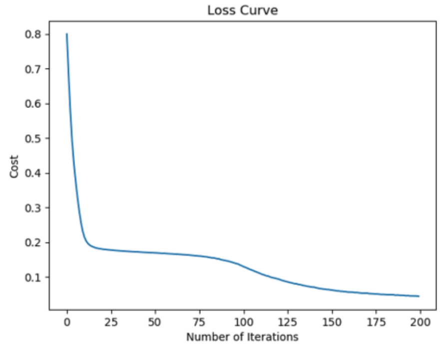

**Fig. 9:** Confusion Matrix for ANN Model

Fig. 9 shows that out of 743 grapefruit samples, the ANN model was able to predict 738 grapefruit (correct prediction) and 5 as orange (incorrect prediction). And for the 757 orange samples, it predicted 739 as orange and only 18 as grapefruit.

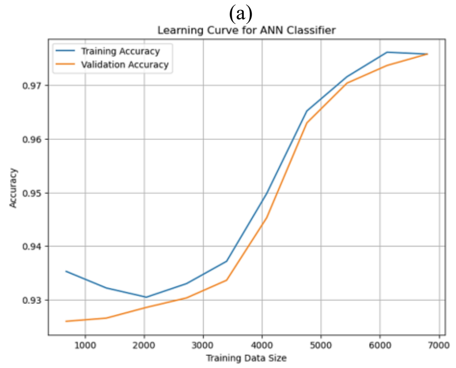
Also in the ANN model, the loss curve in Fig. 10(a) indicates a steady decrease in cost as the number of iterations increases. This means the neural network is learning and improving its predictions as the training progresses. Fig. 10(b) also shows the learning curve of the ANN classifier model; the close alignment in the training and validation accuracies suggests that the model is not overfitting. Additionally, both accuracies converge to a high value, indicating the ANN model is learning effectively and generalizing well to unseen data.

### *Crop Yield Prediction*

In performing regression tasks on the crop yield datasets, the hyperparameters that resulted in maximum performance are in Table 5. It illustrates the different tuning parameters for the DT, RF, and ANN models.




(a)




(b)

**Fig. 10:** (a) Loss Curve for ANN Classifier Model (b) Learning Curve for ANN Classifier Model

**Table 5:** Hyperparameters for ML Regressor Models

| ML Models | Hyperparameters | |
|---|---|---|
| DT, RF, ANN | Test Size | 10% |
| | N-fold Cross Validation | N=5 |
| RF | Estimators | 21 |
| ANN | Activation Function | ReLu |
| | Hidden Layer Number | 5 |
| | Maximum Iteration | 800 |

The error metrics, including RMSE, MAE, and MSE, were calculated for the three models. Tables 6 and 7 show the comparison of the error metrics for the different ML regressor models as a percentage of the mean yield. The mean yield for the data was obtained as 12240.1 for Africa and 14439.6 for America. From Tables 6 and 7, Random Forest demonstrates the best performance with the lowest MAE (12.33, 14.74%) and MSE (7.33, 5.16%), indicating that its predictions are closer to actual values on average and have fewer large deviations. The Neural Network performs poorly in both continents, with extremely high MAE (45.91, 38.60%) and MSE (32.52, 35.63%). This indicates that its predictions are far from accurate and suffer from significant errors due to overfitting or underfitting of the training data by the model. The RMSE for the ANN is relatively higher compared to the other models.

The coefficient of determination ($R^2$) indicates how well a model fits the observed data. In Fig. 11, the RF model achieved an $R^2$ value of 0.856 for Africa and 0.905 for America. These high $R^2$ values suggest that the RF models can accurately predict the variance in the data. Conversely, the ANN model, with a value of 0.062 and 0.333, performed poorly, indicating that it is not a good fit for the crop yield data for both continents. $R^2$ values range from 0 to 1, with higher values indicating a better fit.

**Table 6:** Performance Measure of ML Regressor Models as % of Mean Yield_ Africa

| Regressor Models | MAE | MSE | RMSE |
|---|---|---|---|
| DT | 12.80% | 8.78% | 29.62% |
| RF | 12.33% | 7.33% | 27.08% |
| ANN | 45.91% | 32.52% | 57.03% |

**Table 7:** Performance Measure of ML Regressor Models as % of Mean Yield_ America

| Regressor Models | MAE | MSE | RMSE |
|---|---|---|---|
| DT | 15.35% | 8.00% | 28.29% |
| RF | 14.74% | 5.16% | 22.71% |
| ANN | 38.60% | 35.63% | 59.69% |

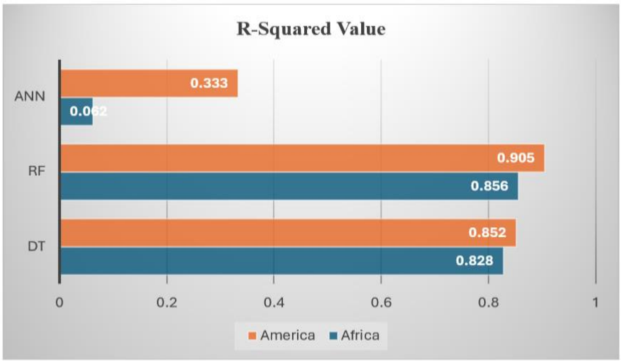



**Fig. 11:** $R^2$ Results for ML Regression Models

The ANN regressor achieved substantially lower performance ($R^2$ = 0.062 for Africa, 0.333 for America) compared to RF. Several factors explain this gap, including the yield dataset containing only five input features, which is suboptimal for neural networks that excel with high-dimensional data. Additionally, the sample-to-parameter ratio is unfavorable, potentially causing overfitting, and agricultural yield data exhibits threshold-based relationships that RF captures naturally through axis-aligned splits, while ANN's smooth activation functions struggle with such discontinuities.

A 5-fold cross-validation of the $R^2$ was performed on the regressor models, as shown in Table 8, including the mean $R^2$ and standard deviation values. The RF model achieved the highest mean $R^2$ of 0.8837 +/0.025 for yield in Africa. Similarly, results were seen for America with values 0.7480 +/- 0.0718 for the RF model, as shown in Table 9. ANN had the lowest average value of all 5 folds for both continents.

Figures 12 and 13 show that the predictions made for crop yield for DT and RF models are close to the actual yield. This means both RF and DT can be used to forecast future yields, with the RF model having a slightly higher accuracy than DT. However, ANN had the least closeness in actual and predicted yields as indicated in Fig. 14. Hence, deploying the ANN model for future predictions is not recommended.

**Table 8:** Cross-Validation Values for Regressor Models - Africa

| Models | $R^2$ Scores for 5 Folds | Mean $R^2$ | Standard Deviation (+/-) |
|---|---|---|---|
| DT | [0.6515, 0.8496, 0.6862, 0.7983, 0.8684] | 0.7708 | 0.087 |
| RF | [0.8556, 0.8559, 0.8951, 0.8914, 0.9206] | 0.8837 | 0.025 |
| ANN | [0.1082, 0.1853, 0.1031, 0.1525, 0.1213] | 0.1341 | 0.0308 |

**Table 9:** Cross Validation Values for Regressor Models - America

| Models | $R^2$ Scores for 5 Folds | Mean $R^2$ | Standard Deviation (+/-) |
|---|---|---|---|
| DT | [0.7711, 0.4121, 0.6243, 0.5795, 0.5929] | 0.5960 | 0.1145 |
| RF | [0.8382, 0.6401, 0.8135, 0.7411, 0.7072] | 0.7480 | 0.0718 |
| ANN | [0.1122, 0.2406, 0.1789, 0.1262, 0.3031] | 0.1922 | 0.0715 |

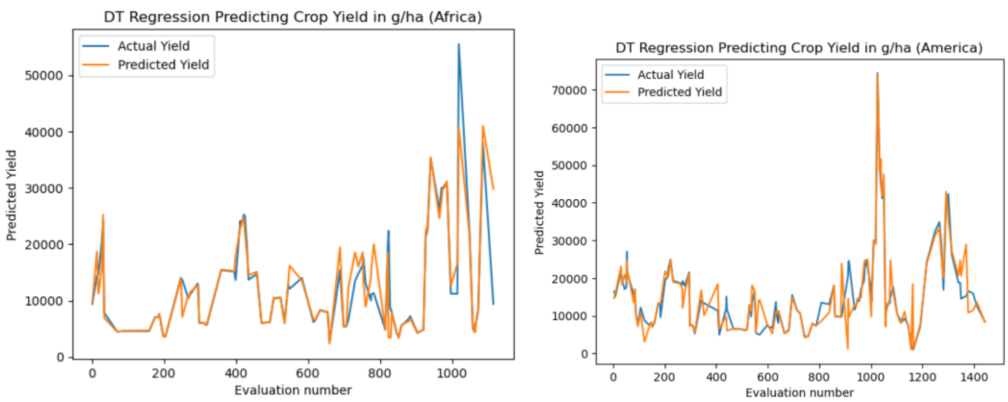



**Fig. 12:** Actual and Predicted Yield for DT Regressor

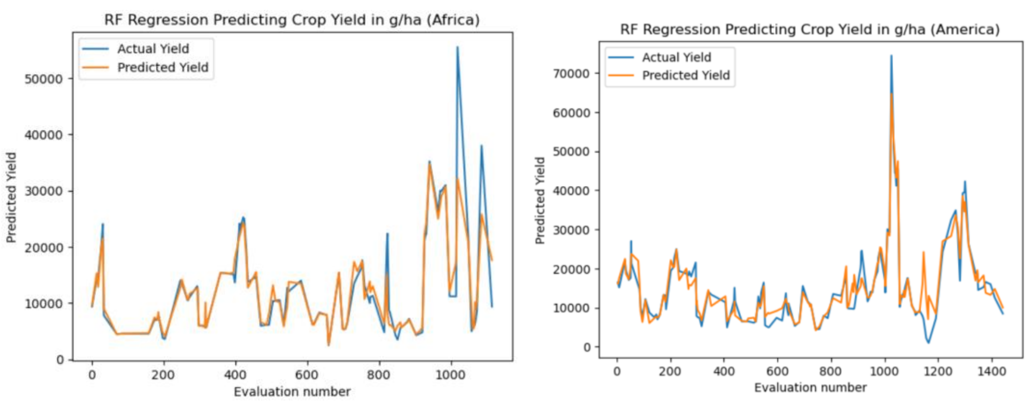



**Fig. 13:** Actual and Predicted Yield for RF Regressor

DOI: 10.3844/jcssp.2026.1344.1359

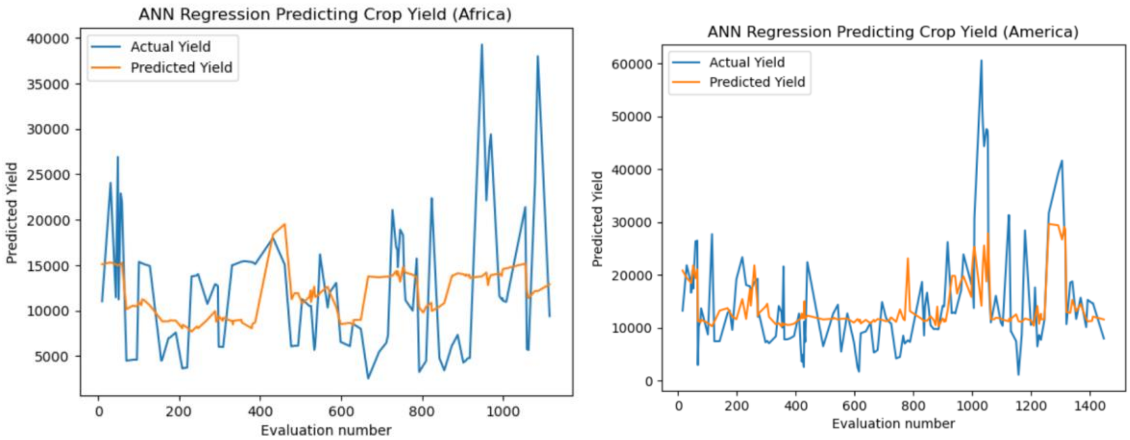



**Fig. 14:** Actual and Predicted Yield for ANN Regressor

### *Feature Importance and Error Distribution Analysis*

To understand the drivers of model performance, feature importance scores were extracted from the Random Forest models (Fig. 15). For citrus classification, physical size features dominated, diameter (42.3%) and weight (33.9%) together accounted for over 76% of the predictive signal, with the green color channel as the most discriminative color feature. This ranking aligns with botanical characteristics, as grapefruits are typically larger with more yellow-green coloration compared to oranges. For yield predictions across both Africa and America, environmental factors showed the highest importance, temperature (24.2%) and rainfall (23.0%) collectively explained nearly half of the yield variance, consistent with agronomic principles where thermal conditions and water availability are primary determinants of citrus productivity. Among nutrients, phosphorus (22.3%) showed the highest importance, followed by nitrogen (17.3%) and potassium (13.3%).

Fig. 16 illustrates the distribution of prediction errors for both regression models across the combined Africa and America datasets. The RF regressor exhibits a near-symmetric error distribution centered close to zero, indicating unbiased predictions with balanced over- and under-estimation. In contrast, the ANN regressor shows a positive bias with higher variance, suggesting systematic under-prediction of yields. RF's error distribution further supports its selection as the preferred regression model for this application.

### *Model Deployment for Test Cases*

Using the ANN model from the citrus grade and the RF model from the crop yield prediction, the models were integrated and deployed. Figures 17 and 18 show the deployment where the system uses the outcome of the citrus grade to predict the yield for countries in Africa and America, respectively.

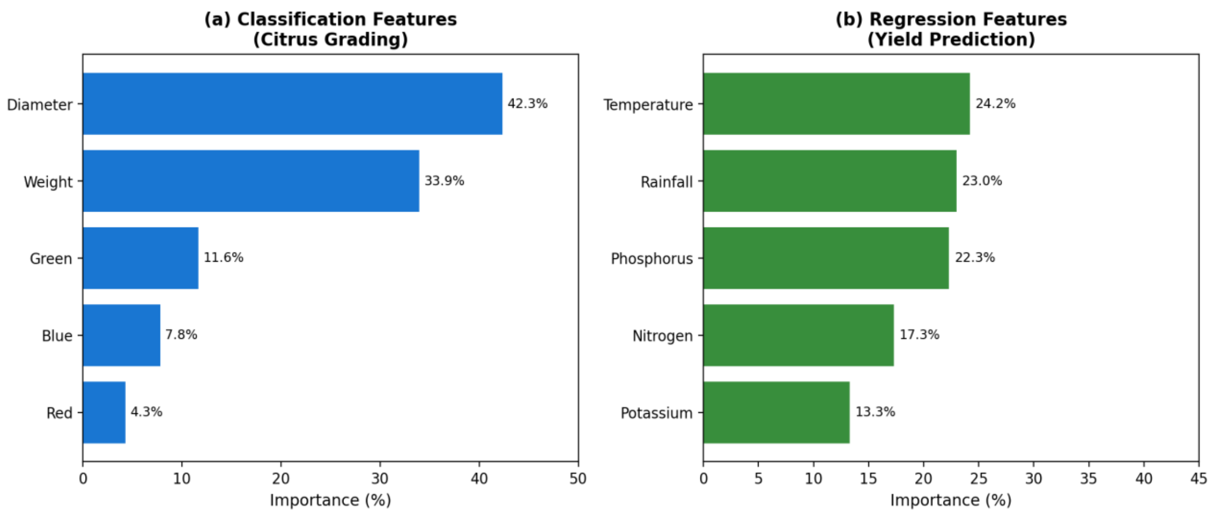



**Fig. 15:** Feature Importance Rankings from Random Forest Models

**DOI:** 10.3844/jcssp.2026.1344.1359

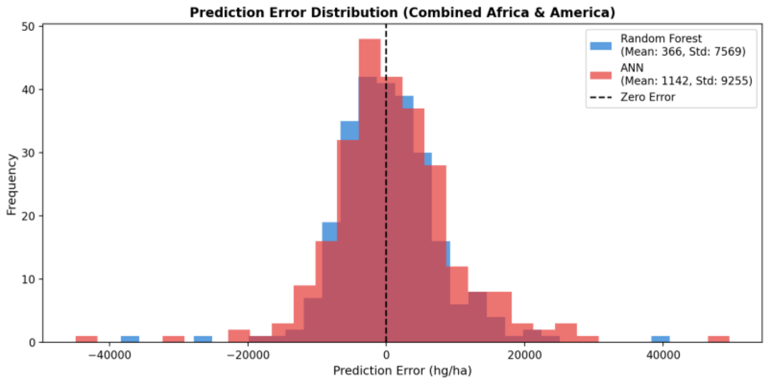



**Fig. 16:** Prediction Error Distribution for RF and ANN Regressors

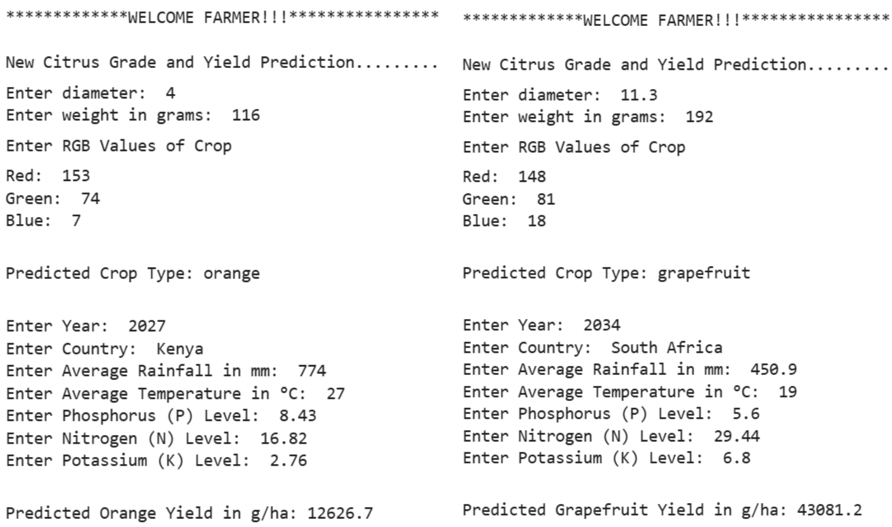



**Fig. 17:** Deployment of the Integrated Model for Countries in Africa

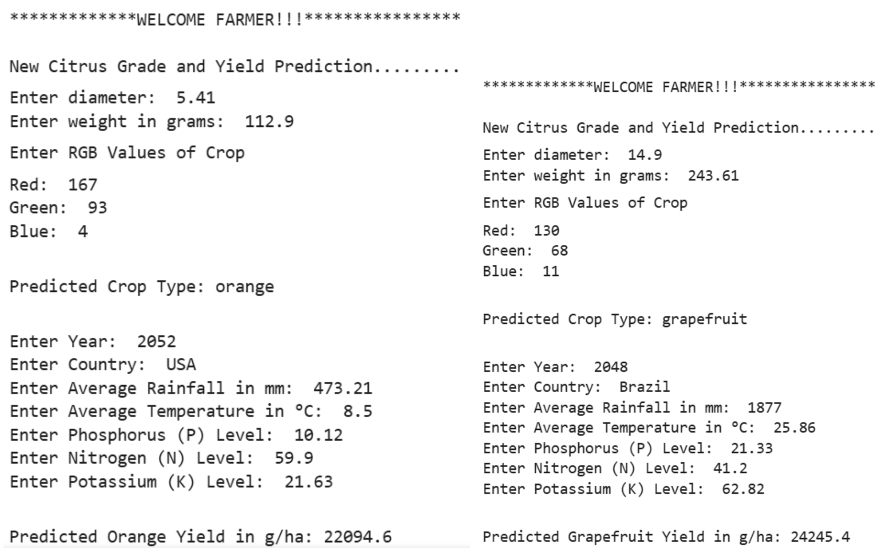



**Fig. 18:** Deployment of the Integrated Model for Countries in America

**DOI:** 10.3844/jcssp.2026.1344.1359

*Comparison with Existing Literature*

Table 10 illustrates the superior performance of the proposed integrated machine learning system in both citrus classification and yield prediction tasks, compared to existing approaches. A key innovation of this research lies in the dual-function framework, where the citrus grading classifier directly informs the yield regressor. This design not only enhances predictive accuracy but also aligns with real-world agricultural workflows that require simultaneous quality assessment and yield forecasting. In the classification task, the system achieves a maximum accuracy of 98.5% using ANN, outperforming prior methods such as SVM and CNN reported in related studies (Ismail and Malik, 2022). For yield prediction, the Random Forest regressor achieves an $R^2$ score of 0.905, which reflects high model reliability and improved generalization over previous models. This integrated approach streamlines both quality control and agricultural planning, which makes it particularly valuable for precision farming applications where rapid and accurate decision-making is essential.

**Table 10:** Comparison of Proposed System with Existing Literature

| Publication | Research Focus | Methods | Data Source | Performance Score |
|---|---|---|---|---|
| (Ismail and Malik, 2022) | Fruit grading with computer vision and deep learning | ResNet, DenseNet, MobileNetV2, NASNet, and EfficientNet | (Li et al., 2009; Mazen and Nashat,, 2019) | Accuracy: 96.7% apples and 93.8% bananas |
| (Cai et al., 2024) | Decay detection in citrus fruits | Hybrid CNN-SVM | Captured Citrus reflectance images | Accuracy: 90.6% |
| (Xu et al., 2025) | Citrus Yield Prediction | RF, SVM, XGBOOST | Manual Selection of Citrus Trees in the fruit maturity stage. | Max $R^2$ Score: 0.853 XGBoost |
| (Morales and Villalobos, 2023) | Predicting wheat and sunflower yields | RF, ANN, Standard Linear Regression | (Ritchie et al., 1985; Villalobos et al., 1996) | Max $R^2$ Score: 0.75 RF Model |
| Proposed System | Integrated Citrus grading and yield prediction | ANN-RF Classifier-Regressor, DT | (Kaggle, 2020; FAOSTAT, 2025; World Bank, 2025) | Accuracy: 98.5% (citrus grade), Max $R^2$=0.905 (yield) |

## Conclusion

Agriculture remains the cornerstone of human society, providing essential sustenance and economic stability globally. In response to growing challenges such as climate change, resource constraints, and rising food demand, this study introduces a new dual-function machine learning system that integrates citrus grading and yield prediction. Using an Artificial Neural Network (ANN) classifier with 98.5% accuracy (±0.0033 SD) to distinguish between visually similar citrus varieties, followed by a Random Forest (RF) regressor achieving an $R^2$ of 0.905 (±0.025 SD), the proposed framework demonstrates the value of classifier-informed regression. Comparative analysis shows performance improvements of 2.5–8.5% in classification and 10–15% in yield prediction over traditional models like SVM, SVR, and DNN.

The sequential integration of grading and yield prediction aligns with real-world agricultural workflows, enabling more reliable and data-driven quality control and forecasting.

This study has several limitations that warrant consideration. First, the citrus classification dataset is synthetic rather than derived from real agricultural measurements, which may limit generalizability despite capturing realistic feature distributions. Second, the yield prediction model relies on historical data from 1990 to 2022, which may not fully account for recent climate change impacts or evolving agricultural practices.

Future work will focus on several directions to enhance the framework's utility and robustness. First, validation with larger, more diverse citrus datasets, including real field measurements, will establish practical applicability. Second, incorporating real-time environmental data through IoT sensors (temperature, humidity, soil moisture) could improve yield prediction accuracy. Third, the development of a mobile application for field use would enable farmers to access predictions directly during growing seasons. Fourth, exploring ensemble methods that combine multiple classifiers may further improve grading accuracy. Finally, extending the framework to other crop types beyond citrus will test generalizability across agricultural domains.

## Acknowledgment

The authors gratefully acknowledge the resources made available by Norfolk State University, Virginia, USA.

## Funding Information

The authors have not received any financial support or funding to report.

## Authors' Contributions

The authors equally contributed to this research.

## Ethics

This research did not involve human participants, animal subjects, or the use of personally identifiable data. All datasets utilized are publicly available and were used in accordance with their respective terms of use. No ethical concerns were identified in the conduct of this study.

## References

Bender, F. E., Douglass, L. W., & Kramer, A. (2020). *Statistical Methods for Food and Agriculture*. https://doi.org/10.1201/9781003073116

Burhan, H. A. (2022). Crop Yield Prediction by Integrating Meteorological and Pesticides Use Data with Machine Learning Methods: An Application for Major Crops in Turkey. *Ekonomi, Politika; Finans Araştırmaları Dergisi*, *7*(IERFM Özel Sayısı), 1–18. https://doi.org/10.30784/epfad.1148948

Cai, Z., Sun, C., Zhang, H., Zhang, Y., & Li, J. (2024). Developing universal classification models for the detection of early decayed citrus by structured-illumination reflectance imaging coupling with deep learning methods. *Postharvest Biology and Technology*, *210*, 112788. https://doi.org/10.1016/j.postharvbio.2024.112788

Camara-Guerra, A., Artyounian-Vieyra, C., Gonzalez-Cuellar, E., Trevino-Escamilla, A., Salazar-Garibay, A., & Hernandez-Gutierrez, A. (2024). Classification of Citrus Crops Using Satellite Multispectral Imagery and Deep Neural Network. *Proceeding of the 2024 16th International Conference on Computer and Automation Engineering (ICCAE)*, 351–356. https://doi.org/10.1109/iccae59995.2024.10569452

Chakhar, A., Zitouna-Chebbi, R., Hernández-López, D., Ballesteros, R., Mahjoub, I., & Moreno, M. A. (2024). Assessing the Accuracy of Multiple Classification Algorithms Combining Sentinel-1 and Sentinel-2 for the Citrus Crop Classification and spatialization of the Actual Evapotranspiration Obtained from Flux Tower Eddy Covariance: Case Study of Cap Bon, Tunisia. *Proceedings of IAHS*, *385*, 443–448. https://doi.org/10.5194/piahs-385-443-2024

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random Forests. *Ensemble Machine Learning*, 157–175. https://doi.org/10.1007/978-1-4419-9326-7_5

Dahlkar, S. S., & Rode, S. V. (2014). Agricultural crop yield prediction using artificial neural network approach. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, *2*(1), 683–686.

FAOSTAT. (2025). Crops and livestock products. *Food and Agriculture Organization*. https://www.fao.org/faostat/en/?utm_source=copilot.com#data/QCL/visualize

Genuer, R., & Poggi, J.-M. (2020). Random forests. *Random Forests with R*, 33–55. https://doi.org/10.1007/978-3-030-56485-8_3

Grove, H. (2021). Grapefruit vs. oranges. *Hale Groves Blog*. https://www.halegroves.com/blog/grapefruit-vs-oranges/

Hossain, S. T. (2020). Impacts of COVID-19 on the Agri-food Sector: Food Security Policies of Asian Productivity Organization Members. *Journal of Agricultural Sciences – Sri Lanka*, *15*(2), 116–132. https://doi.org/10.4038/jas.v15i2.8794

Ismail, N., & Malik, O. A. (2022). Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. In *Information Processing in Agricultur* (Vol. 9, Issue 1, pp. 24–37). https://doi.org/10.1016/j.inpa.2021.01.005

Kaggle. (2020). Oranges vs. grapefruit. *Kaggle (Dataset Repository)*. https://www.kaggle.com/datasets/joshmcadams/oranges-vs-grapefruit

Krenker, A., Bester, J., & Kos, A. (2011). Introduction to the Artificial Neural Networks. *Artificial Neural Networks - Methodological Advances and Biomedical Applications*, 1–18. https://doi.org/10.5772/15751

Li, G., Holguin, G., Park, J., Lehman, B., & Hull, L. (2009). The CASC IFW database. *Purdue University*.

Mazen, F. M. A., & Nashat, A. A. (2019). Ripeness Classification of Bananas Using an Artificial Neural Network. *Arabian Journal for Science and Engineering*, *44*(8), 6901–6910. https://doi.org/10.1007/s13369-018-03695-5

Morales, A., & Villalobos, F. J. (2023). Using machine learning for crop yield prediction in the past or the future. *Frontiers in Plant Science*, *14*, 1–20. https://doi.org/10.3389/fpls.2023.1128388

Nowozin, S., Rother, C., Bagon, S., Sharp, T., Yao, B., & Kohli, P. (2011). Decision tree fields. *Proceeding of the 2011 International Conference on Computer Vision*, 1668–1675. https://doi.org/10.1109/iccv.2011.6126429

Özden, C., & Karadoğan, N. (2024). Wheat yield prediction for Turkey using statistical machine learning and deep learning methods. *Pakistan Journal of Agricultural Sciences*, *61*(2), 429–435. https://doi.org/10.21162/PAKJAS/24.182

Pandey, C., Sethy, P. K., Behera, S. K., Vishwakarma, J., & Tande, V. (2022). Smart agriculture: Technological advancements on agriculture—A systematical review. *Deep Learning for Sustainable Agriculture*, 1–56. https://doi.org/10.1016/b978-0-323-85214-2.00002-1

Reddy, N. V., Kumar, A., Reddy, P., Manoj, M., & Kumar, M. (2024). Crop yield prediction and remedies recommendation using feature selection techniques in machine learning. *IRIMETS*, *6*(3), 203–210.

Ritchie, J. T., Godwin, D. C., & Otter-Nacke, S. (1985). *CERES-Wheat: A Simulation Model of Wheat Growth and Development*.

Sugumar, D., Harshavarthan, V., Kavisri, S., Vallavi, M., & Vanathi, P. (2019). Citrus classification and grading using machine learning algorithms. *International Journal of Innovative Technology and Exploring Engineering*, *8*(10), 2616–2621. https://doi.org/10.35940/ijitee.J9349.0881019

Suthaharan, S. (2016). Decision Tree Learning. *Machine Learning Models and Algorithms for Big Data Classification*, *36*, 237–269. https://doi.org/10.1007/978-1-4899-7641-3_10

Villalobos, F. J., Hall, A. J., Ritchie, J. T., & Orgaz, F. (1996). OILCROP-SUN: A Development, Growth, and Yield Model of the Sunflower Crop. *Agronomy Journal*, *88*(3), 403–415. https://doi.org/10.2134/agronj1996.00021962008800030008x

van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, *177*, 105709. https://doi.org/10.1016/j.compag.2020.105709

World Bank, (2025). Climate Knowledge Portal. *Climate Change Knowledge Portal*. https://climateknowledgeportal.worldbank.org/country#country-map

Walczak, S. (2018). *Artificial neural networks*. https://doi.org/10.4018/978-1-5225-7368-5.ch004

Xu, W., Liu, X., Dong, J., Tan, J., Wang, X., Wang, X., & Wu, L. (2025). Improvement of Citrus Yield Prediction Using UAV Multispectral Images and the CPSO Algorithm. *Agronomy*, *15*(1), 171. https://doi.org/10.3390/agronomy15010171

Yang, G. R., & Wang, X.-J. (2020). Artificial Neural Networks for Neuroscientists: A Primer. *Neuron*, *107*(6), 1048–1070. https://doi.org/10.1016/j.neuron.2020.09.005

Zhang, X. & Li, Y. (2019). A Review of Scikit-learn Package in Python Programming Language. *Journal of Educational and Behavioral Statistics*, *44*(3), 348–361. https://doi.org/10.3102/1076998619832248