FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 8 No. 1, February, 2024, pp 61 - 68
DOI: https://doi.org/10.33003/fjs-2024-0801-2220

CROP YIELD PREDICTION USING SELECTED MACHINE LEARNING ALGORITHMS

***Nazifi Shuaibu, Obunadike, G. N. and Bashir Ahmad Jamilu**

Department of Computer Science, Federal University Dutsin-Ma. Katsina State, Nigeria

*Corresponding authors' email: nazifishuaibu1@gmail.com Phone: +2348067043264

ABSTRACT

Agriculture is paramount to global food security, and predicting crop yields is crucial for policy and planning. However, predicting these yields is challenging due to the myriad of influencing factors, from soil quality to climate conditions. While traditional methods relied on historical data and farmer experience, recent advancements have witnessed a shift towards machine learning (ML) for improved accuracy. This study explored the application of machine learning (ML) techniques in predicting crop yields using data from Nigeria. Previous efforts lacked transferability across crops and localities; this research aimed to devise modular and reusable workflows. Using data from the Agricultural Performance Survey of Nigeria, this study evaluated the performance of different machine learning algorithms, including Linear Regression, Support Vector Regressor, K-Nearest neighbor, and Decision Tree Regressor. Results revealed the Decision Tree Regressor as the superior model for crop yield prediction, achieving a prediction accuracy of 72%. The findings underscore the potential of integrating ML in agricultural planning in Nigeria where agriculture significantly impacts the economy. Further research is encouraged to refine these models for broader application across varying agroecological zones.

Keywords: Crop yield prediction, Decision Tree Regressor, K-Nearest neighbor, Linear Regression, Machine learning, Support Vector Regressor

INTRODUCTION

Crop yield can be described as the measurement of a farm product grown per unit area of land. The measurement unit of crops is usually by kilograms per hectare or bushels per acre. According to a report in (Factors That Influence Crop Yield - Omnia Nutriology®, 2017) shown that yield performance of many crops can be attributed to four most important factors including soil fertility, availability of water, climate, and diseases or pests. This is some of the most important information used by scientist to predict crop yield (Xu et al., 2019).

Predicting crop yield is critical to addressing one of the gaining problems in food security, particularly with the impact of global climate change (Ansarifar and Wang 2019). Predictions are vital but complex problems, which is needed for sustainable boosting and good use of natural resources (Phalan, Green, and Balmford 2014). Accurate crop yield prediction is very pertinent to global food production. This is because, the predictions not only aid farmers in making informed commercial and management decisions but also help in famine prevention activities (Ansarifar and Wang 2019).

Different approaches have been used to predict crop performance including field surveys, crop growth models, remote sensing, statistical models, and their combinations (Paudel et al., 2021). Each of these approaches addresses imperceptible different aspects of crop yield prediction independently. The field surveys approaches try to apprehend the ground truth while crop growth models simulate the crop growth and development, putting agronomic principles, environmental and management interactions into consideration (Chipanshi et al., 2015). Remote sensing depends on satellite instruments showing frequent, coarse resolution image time series for yield estimation (Atzberger et al., 2016).The statistical models rely on the use of weather variables and the output of field survey, crop growth models and remote sensing as predicators to develop linear relationships between the predicators and crop yield (Paudel et al., 2021). Some studies have combinedtwo or more of these approaches to predict crop yield. For example, in the studies of Zhao, Potgieter, Zhang, Wu and Hammer (2020) combined crop modelling and high-resolution remote sensing data to build statistical models to predict crop yield. Another study with similar approach conducted by Newlands et al. (2014) proposed a probabilistic yield prediction in Canada using crop modelling, remote sensing, Bayesian inference and statistical models.

Machine learning (ML) takes a data-driven or empirical modeling approach to learn useful patterns and relationships from input data (Willcock et al., 2018) and offers a promising opportunity for improving crop yield predictions (Paudel et al., 2021). Machine learning models have proven powerful performance in several data-driven applications including the crop yield prediction (Zhao, Potgieter, Zhang, Wu and Hammer 2020; Paudel et al., 2021). Many studies have employed machine learning approaches such as the multivariate regression, random forest, association rule mining, regression tree and artificial neural network for crop yield prediction (Khaki, Wang, and Archontoulis 2020). The machine learning models treat the output, crop yield as an inherent function of the input variables such as weather parameters and soil conditions, which might be a precise complex and nonlinear function (Khaki, Wang, and Archontoulis 2020). Just as in statistical models, machine learning algorithms can also use the output of other prediction approaches as features. Machine learning algorithms have some distinct benefits as can model non-linear relationships between multiple sources of data (Chlingaryan, Sukkarieh and Whelan 2018). The performance of Machine learning algorithms improves generally when more training is avail, where regularization techniques are employed to reduce variance and regularization error when the data is robust to noisy (Goodfellow, Bengio and Courville 2016). Therefore, machine learning could combine the benefits of other approaches, such as remote sensing, data-driven models, and crop growth modelling to make reliable

crop yield prediction (Paudel et al., 2021).

The European Commission's Joint Research Centre (JRC) and the National Agricultural Statistics Service (NASS) of US Department of Agriculture have a large-scale crop yield forecasting systems, such as the MARS Crop Yield Forecasting System (MCYFS) that relies on the infrastructure and historical data to build and assess crop prediction models for various crops in different localities (Paudel et al., 2021). The system utilizes statistical models from field survey results, crop growth model output, weather observations, remote sensing indicators and yield statistics (MARSWiki, 2020; USDA-NASS, 2012). However, performance evaluation of MCYFS from 1993 – 2015 shows no significant improvement in the performance from 2006 onwards (Van der Velde and Nisini, 2019). Machine learning could be the best model for such large-scale system.

Machine learning is a promising approach especially when a large amount of dataset is gathered and made public (Lokers, Knapen, Janssen, Randen, and Jansen 2016). For example, Jeong et al. (2016) employed multiple linear regression and random forest for yield prediction of potato, wheat and maize. The same machine learning algorithms were used by Shahhosseini, Martinez-Feria, Hu and Archontoulis (2019) to predict nitrate loss and corn yield. Awad (2019) proposed a mathematical optimization model and calculated biomass to predict potato yield. Several machine learning including decision tree and association rule mining for the classification of yield components of durum wheat and showed that association rule mining method best performance across all locations of the study (Romero, 2013). Ransom et al. (2019) evaluated machine learning approaches for corn nitrogen recommendation tool suing soil and weather information.

Related Works

various researchers explore the use of machine learning and data-driven approaches in optimizing agricultural practices and predicting crop yield. Chipanshi et al. (2015) focus on using an Extreme Learning Machine (ELM) model to accurately estimate coffee yield based on soil fertility properties, showing superior performance compared to traditional models. Goldstein et al. (2018) integrate data from various sources to predict irrigation recommendations for Jojoba crops, achieving high accuracy with regression and classification algorithms. Zhong, Li, Lobell, Ermon and Brandeau (2018) propose a hierarchical machine learning mechanism for seed variety selection, considering yield maximization and risk. Crane-Droesch (2018) introduce a deep neural network approach to model the relationship between weather and corn yield, outperforming traditional methods and showing less severe climate change impacts. Khanal, Fulton, Klopfenstein, Douridas and Shearer (2018) demonstrate the effectiveness of machine learning algorithms and remotely sensed data in predicting soil properties and corn yield. Taherei Ghazvinei et al. (2018) apply extreme learning machine to predict sugarcane growth, providing a swift and accurate model for the sugarcane industry. Ahmed et al. (2018) combines remote sensing and crop modeling to estimate maize yield, showcasing the potential of both techniques with high accuracy. These studies collectively highlight the value of machine learning and data-driven approaches in optimizing agricultural practices and yield prediction. Xu et al. (2019) developed an integrated climatic assessment indicator (ICAI) in Jiangsu Province, China, to evaluate the synthetic effects of meteorological factors on crop production. They used machine learning algorithms to construct the indicator, with Random Forest (RF) performing the best. The ICAI provided values for yield loss, normal conditions, and yield increment. The study assessed the past climatic suitability of winter wheat and predicted future suitability under global warming conditions. Filippi et al. (2019) explored the value of combining data from multiple fields and years for predicting crop yield. They used large farms in Western Australia as a case study and developed random forest models to predict crop yield. The models showed accurate predictions, improving as the season progressed and more within-season data became available. Ranjan and Parida (2019) focused on paddy acreage mapping and yield prediction in Sahibganj district, India, using Sentinel-based optical and SAR sensors data. They employed a Random Forest classification technique for mapping paddy acreage and developed a linear regression model for yield prediction. The study highlighted the usefulness of SAR data for accurate acreage mapping and the potential of timely information for decision-makers. Agarwal and Tarar (2021) addressed crop prediction in Indian agriculture using machine learning algorithms. They proposed an enhanced model, incorporating deep learning techniques such as Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Recurrent Neural Network (RNN). The model aimed to predict the most productive crop and provide information on soil ingredients and expenses. The study emphasized the use of climatic and soil conditions for accurate yield predictions and to assist farmers in decision-making processes. Paudel et al (2021) used Supervised regression and found that explainable features designed using principles of crop modeling can be used to predict crop yield at sub-national level. Ahmed, Adewumi, and Yemi-peters (2023) deployed Random Forest Algorithm to improve precision accuracy with minimal errors compared to manual process.

Machine learning models

Machine learning models are mathematical algorithms or computational systems that are designed to learn patterns and make predictions or decisions based on input data. These models are trained on large datasets to recognize and generalize patterns, enabling them to perform tasks such as classification, regression, clustering, or anomaly detection. Here are some machine learning models considered in this study.

Regression

In a machine learning regression model, the goal is to predict a continuous output value (y) given an input feature vector (x) (Xu et al., 2019). The predicted output value is represented as a function of the input features, which can be represented mathematically as shown in Equation 2.1:

$$y = f(x) + \varepsilon \quad (1)$$

where $f(x)$ is the predicted value of y given x, and ε is the error term.

In a linear regression model, the function $f(x)$ is a linear function of the input features (Equation 2.2):

$$f(x) = w1x1 + w2x2 + \cdots + wnxn + b \quad (2)$$

where $w1$, $w2$, ..., wn are the model coefficients (also known as weights) and b is the bias term. For example, if we have a single input feature x and a linear regression model with coefficient w and bias b, the predicted output value y is expressed in Equation 2.3:

$$y = wx + b \quad (3)$$

The model coefficients and bias are learned from the data during the training process. The goal is to find the values of the coefficients and bias that minimize the error between the predicted values and the true values of y in the training data. Once the model is trained, it can be used to make predictions on new, unseen data by plugging in the appropriate values for

x into the equation for $f(x)$ (Cravero, Pardo, Sepúlveda and Muñoz 2022).

Support Vector Regressor

Support Vector Regression (SVR) model is a powerful regression tool that predicts a continuous output value, represented as , based on a given input feature vector (Xu et al., 2019). The estimated output value is articulated as a function of the input features and can be expressed mathematically as follows:

where is the estimated output value given input feature vector, and ε signifies the error term, representing the error in prediction.

Distinct from traditional regression techniques, SVR does not aim to minimize the error. Instead, it aspires to fit the optimal hyperplane within a predefined error value ε, establishing an ε-insensitive tube (Agarwal and Tarar, 2021). The fundamental strategy of SVR is to identify a function which deviates from the actual response at most ε and at the same time is as flat as possible.

SVR operates by mapping the input space into a high-dimensional feature space via a kernel function, where the function in SVR is then formulated as a function of the input features.

The mathematical representation of the Support Vector Regression (SVR) model is a bit more complex due to the utilization of the kernel function for mapping the data to higher dimensions and the introduction of the ε-insensitive loss function (Cravero, Pardo, Sepúlveda and Muñoz 2022).

Formally, a linear Support Vector Regression function can be expressed in equation 2.4 as:

$f\ (x) = < w, x > \ + b$ (4)

where:

$f(x)$) is the regression estimate

$< w, x >$ denotes the dot product of the weight vector w and the input vector x

b is the bias term.

However, in most practical situations, the data is not linear. In such cases, SVR employs the kernel trick to map input data to a higher-dimensional feature space where the data can be linearly separated. This allows the use of linear methods (like SVR) to solve non-linear problems.

The kernelized version of the SVR function becomes (Equation 2.5):

$f(x)\ = \ \Sigma\ (ai - ai*)\ K(xi, x)\ + \ b$ (5)

where:

$f(x)$ is the regression estimate

$K(xi, x)$ is the kernel function that maps xi and x to a higher-dimensional space

ai and $ai*$ are Lagrange multipliers obtained from the solution of the dual problem.

b is the bias term

The optimization problem in SVR is to find the values of w and b that minimize the following:

$½\ ||w||\text{^}2\ +\ C\ \Sigma\ (\xi i\ +\ \xi i*)$

under the constraints:

$yi - < w, xi > \ - b\ <= \ \varepsilon\ +\ \xi i$

$< w, xi > \ + b - \ yi\ <= \ \varepsilon\ +\ \xi i*$

$\xi i, \xi i* >= \ 0\ for\ all\ i$

where:

$||w||\text{^}2$ is the square of the Euclidean norm of w

C is the regularization parameter.

ξi and $\xi i*$ are slack variables introduced to cope with the infeasible constraints of the optimization problem.

ε is the width of the insensitive tube.

In simple terms, the SVR algorithm tries to find a function $f(x)$ that has at most ε deviation from the actually obtained target yi for all the training data, and at the same time, is as flat as possible (Agarwal and Tarar, 2021). This is achieved by minimizing $||w||$, which gives the flatness. In the case where this is not possible, the function is allowed to deviate more than ε, but these deviations are penalized in the objective function of the optimization problem.

K-Nearest Neighbour

K-Nearest Neighbors (K-NN) is a simple, yet effective supervised learning algorithm used for both classification and regression (Xu et al., 2019). It works based on the assumption that similar inputs have similar outputs, and the algorithm's output is determined by the properties of its neighboring data points.

The K-NN algorithm operates by identifying 'K' instances that are nearest to the test instance and classifies the input based on the most common class in the neighborhood.

In the case of a regression problem, it takes the mean (or median, depending on the use case) of the values of its nearest neighbors.

The distance between two instances can be measured in many ways, such as Euclidean distance, Manhattan distance, Minkowski distance, etc. The choice of distance measure depends on the problem at hand (Cravero, Pardo, Sepúlveda and Muñoz 2022).

To explain it mathematically, let's denote x as the input vector to be classified or used for prediction, and D as the dataset.

The 'K' nearest neighbors are identified by the function as shown in equation 2.6:

$NNk(x)\ =\ argmin\ (d(x, xi))\ for\ all\ xi\ in\ D$ (6)

Here, $d(x, xi)$ is a distance metric like the Euclidean distance, which for two points $x\ =\ (x1, x2, \ldots, xn)\ and\ y\ =\ (y1, y2, \ldots, yn)$ in an n-dimensional space can be computed as (Equation 2.7):

$d(x, y)\ =\ sqrt[(x1 - y1)\text{^}2\ +\ (x2 - y2)\text{^}2\ +\ \ldots\ +\ (xn - yn)\text{^}2]$ (7)

This formula is used to calculate the distance between the new instance and all the instances in the training data.

For classification, once the K neighbors are identified, the algorithm assigns the class that is most common among the neighbors:

$y\ =\ mode(yi)\ for\ all\ xi\ in\ NNk(x)$

Here, mode(yi) is the most common output (class) among the K nearest neighbors.

For regression, the predicted output is typically the mean or median of the K nearest neighbors:

$y\ =\ mean(yi)\ for\ all\ xi\ in\ NNk(x)$

or

$y\ =\ median(yi)\ for\ all\ xi\ in\ NNk(x)$

K-NN is a non-parametric, lazy learning algorithm meaning it doesn't learn a discriminative function from the training set but 'memorizes' the training dataset instead (Cravero, Pardo, Sepúlveda and Muñoz 2022). The parameter K is crucial in this algorithm and choosing the right K is a complex task. A smaller K value will have a more flexible fit which will have low bias but high variance, whereas a larger K will have a smoother decision boundary (less variance) but increased bias.

Decision Tree

A decision tree is a machine learning model used for classification and regression tasks (Xu et al., 2019). It is a tree-like model that makes decisions based on the value of an input feature and splits the data into different branches based on the decision. The final decisions at the leaf nodes of the tree determine the output class or value for the input data (Cravero, Pardo, Sepúlveda and Muñoz 2022).

In a decision tree model, the goal is to predict a class label (in the case of classification) or a continuous output value (in the case of regression) based on a set of input features.
For example, in a classification tree, the Gini impurity at a node t is calculated as Equation 2.8):

$$Gini(t) = 1 - \sum p(i|t) \wedge 2 \quad (8)$$

where $p(i|t)$is the proportion of the samples at node t that belong to class i.
The decision tree is constructed by recursively splitting the data at each node until the tree is fully grown. The final tree can then be used to make predictions on new, unseen data by following the decisions made at each node and reaching a leaf node, at which the output class or value is determined.

MATERIALS AND METHODS

In the proposed framework, four (4) machine learning algorithms including Linear Regression, Support Vector Regressor, K-Nearest Neighbor, and Decision Tree Regressor were executed to predict best crop yield predictions. Multiple but most common cash crops based on atmosphere, locations, and climatic parameters were taken into consideration for selections. In this model, data extracted from multiple sources with a variety of parameters was loaded, followed by the loading of useful libraries and packages for data pre-processing. Feature selection was performed to extract the most important features in the dataset for the best performance. The dataset was then divided into training and testing ratios which were later used for both training and testing by employing the Machine Learning algorithms. The testing dataset was then used for various performance metric evaluations as in figure 1

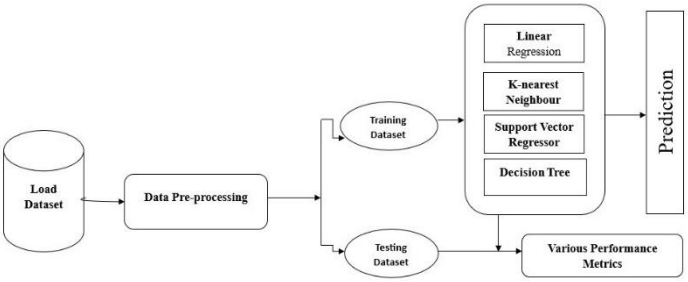

Figure 1 Architecture of Proposed Model for Crop Yield Prediction

Data Collection and Description

The crop yield/performance datasets were generated from the Agricultural Performance Survey of Nigeria by the National Agricultural Extension and Research Liaison Services (NAERLS) and Federal Department of Agricultural Extension (FDAE) for a five (5) year period. There are several crops taken in this dataset like wheat, rice, maize, millet, yam, cocoyam, and sorghum. Climatic data were collected from the Nigerian Meteorological Agency (Nimet) for the same 5 years period. The prediction parameters in this dataset include temperature, rainfall, relative humidity, soil moisture, soil surface, and area. Several values are available for each prediction parameter for a single crop. For instance, when taking a crop such as wheat, any value can be given to the prediction parameters among a set of values available in the dataset, for wheat. It is the same for the entire crops available in the dataset.
A sample of the dataset is captured below in table 1.

Table 1: Dataset

ID	Year	State	Crop	land_area	yield	humidity	soil_surface	max_temp	min_temp	soil_moisture	root_zone	max_speed	min_speed	radiation	rainfall
1	2015	Abia	Maize	79.62	1.16	85.12	0.69	32.78	13.13	0.85	0.68	6.15	0.07	35.77	1439.65
2	2015	Abia	Rice	14.59	2.4	85.12	0.69	32.78	13.13	0.85	0.68	6.15	0.07	35.77	1439.65
4	2015	Abia	yam	175.2	3.96	85.12	0.69	32.78	13.13	0.85	0.68	6.15	0.07	35.77	1439.65
5	2015	Abia	groundnut	7.2	1.344	85.12	0.69	32.78	13.13	0.85	0.68	6.15	0.07	35.77	1439.65
7	2015	Abia	cassava	196.86	9.52	85.12	0.69	32.78	13.13	0.85	0.68	6.15	0.07	35.77	1439.65
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2467	2019	Zamfara	benniseed	7.45	1.57	49.44	0.38	41.55	10.62	0.42	0.4	8.27	0.01	35.13	669.73
2468	2019	Zamfara	cotton	17.61	0.41	49.44	0.38	41.55	10.62	0.42	0.4	8.27	0.01	35.13	669.73
2469	2019	Zamfara	cassava	130.17	1.79	49.44	0.38	41.55	10.62	0.42	0.4	8.27	0.01	35.13	669.73
2470	2019	Zamfara	tomatoe	39.61	5.14	49.44	0.38	41.55	10.62	0.42	0.4	8.27	0.01	35.13	669.73
2471	2019	Zamfara	Onion	27.97	4.52	49.44	0.38	41.55	10.62	0.42	0.4	8.27	0.01	35.13	669.7

Data preprocessing

The raw extracted data was cleaned, transformed, and organized. Exploratory Data analysis was performed to identify outliers, missing values, feature scaling, and data transformation where it is necessary. All the features were evaluated and only the best candidate was selected for the machine learning prediction. The dataset was then divided into training and testing datasets at a 0.2 ratio. This means that 80% would be used for training while the remaining 20% for testing and subsequent performance metrics evaluation.

Feature Extraction

After data cleaning, a feature selection process was undertaken. Utilizing an algorithm based on a tree-structured model such as a Random Forest or Gradient Boosting, importance scores were attributed to each feature. The results highlighted land_area, crop, and humidity as having the highest importance scores, hence being the most influential variables in the dataset.

Following the feature selection process, several models were trained, and their performances were evaluated via two metrics - Mean Squared Error (MSE) and R-Squared (R2 Score). The model utilizing all features ("Full features") did not yield optimal results, with a low R2 score of only 0.17. The Random Forest model displayed a negative R2 score, indicative of its unsuitability for this specific dataset or a possibility of overfitting.

However, the model employing Recursive Feature Elimination (RFE) for feature selection displayed the best performance, with an R2 score of 0.47. This underscores the need for prudent feature selection, as several features may not substantially contribute to the predictive capacity of the model and could therefore be removed.

RESULTS AND DISCUSSION

Crop Yield Prediction Scores of the Various Machine Learning Algorithms

The crop yield prediction scores are expressed in percentage values. The higher the score, the better the algorithm's performance in predicting crop yield as presented in Table 3 below. For the Full-Features dataset, the Linear Regression algorithm achieved a prediction score of 0.92%, the Support Vector Regressor obtained 5.94%, the Decision Tree Regressor achieved 42.11%, and the K-Nearest Neighbor algorithm achieved a score of 25.98%.

When using the RFE-Features dataset, the Linear Regression algorithm obtained a slightly higher prediction score of 1.15%, the Support Vector Regressor achieved 5.89%, the Decision Tree Regressor showed a significant improvement with a score of 71.59%, and the K-Nearest Neighbor algorithm obtained a score of 25.93%.

These results indicate that the performance of the algorithms varies depending on the dataset used. In the case of the Decision Tree Regressor, the algorithm performed significantly better with the RFE-Features dataset compared to the Full-Features dataset as shown in Table 2

Table 2: Crop Yield Prediction Scores of the Various Machine Learning Algorithms

Dataset	Linear Regression	Support Vector Regressor	Decision Tree Regressor	K-Nearest Neighbor
Full-Features	0.92%	5.94%	42.11%	25.98%
RFE-Features	1.15%	5.89%	71.59%	25.93%

Discussion

The varying prediction scores achieved by different machine learning algorithms indicate the importance of dataset selection. The Decision Tree Regressor performed significantly better with the RFE-Features dataset compared to the Full-Features dataset. This finding suggests that the use of feature selection techniques, such as RFE, can improve the performance of prediction models for crop yield. These findings are consistent with studies by Gopal and Bhargavi (2019), which demonstrated the impact of dataset quality and feature selection on crop yield prediction accuracy.

The superior performance of the Decision Tree Regressor on both the Full-Feature and RFE-Feature datasets, as indicated by lower MSE and MAE values and higher R-Squared, implies its effectiveness in predicting crop yield. These results align with the findings of previous studies by Kuradusenge *et al*. (2023) and Javadinejad, Eslamian and Ostad-Ali-Askari (2021), which highlighted the superiority of decision tree-based algorithms in agricultural forecasting. The implication is that employing the Decision Tree Regressor, particularly with the RFE-Feature dataset, can lead to more accurate crop yield predictions.

Our study advances the field by providing a comprehensive understanding of the connection between algorithmic choice, feature selection, and prediction accuracy—and by highlighting the advantages of the Decision Tree Regressor with the RFE-Feature dataset. This thorough study contributes to the repository of existing knowledge and provides practitioners with valuable data to assist them improve crop yield prediction precision.

CONCLUSION

Lastly, considering Nigeria's agricultural climate, our research validates the revolutionary impacts of integrating Machine Learning (ML) into crop yield prediction models. Our contribution is the meticulous evaluation of the Decision Tree Regressor, even though our findings are in line with previous research on the efficacy of machine learning.

With enhanced measurements and a prediction score of 72%, the Decision Tree Regressor demonstrates its robustness in crop yield prediction by regularly outperforming competing algorithms. Beyond simply validating past research, our work offers specific insights for practitioners to optimize crop yield predictions in the Nigerian agriculture setting.

In summary, our research offers a significant contribution to the field by examining the distinct use of the Decision Tree Regressor and providing practitioners and policymakers with useful suggestions. These results enrich existing knowledge and provide strategic direction for the growth of sustainable agriculture in Nigeria and other comparable economies.

REFERENCES

Agarwal, S., and Tarar, S. (2021). A hybrid approach for crop yield prediction using machine learning and deep learning algorithms. *In Journal of Physics: Conference Series* (Vol. 1714, No. 1, p. 012012). IOP Publishing.

Ahmed, A., Adewumi, S. E., and Yemi-peters, V. (2023). Seasonal Crop Yield Prediction in Nigeria Using Machine Learning Technique. *Journal of Applied Artificial Intelligence*, 4(1), 9-20.

Ahmed, I., ur Rahman, M. H., Ahmed, S., Hussain, J., Ullah, A., and Judge, J. (2018). Assessing the impact of climate variability on maize using simulation modeling under semi-arid environment of Punjab, Pakistan. Environmental Science and Pollution Research, 25, 28413-28430.

Ansarifar, J., and Wang, L. (2019). New algorithms for detecting multi-efect and multi-way epistatic interactions. *Bioinformatics*, *35*, 5078–5085.

Atzberger, C., Vuolo, F., Klisch, A., Rembold, F., Meroni, M., Marcio Pupin, M., and Formaggio, A. (2016). *Remote Sensing Handbook* (Agriculture. In: Thenkabail, P.S. (Ed.)). CRC Press.

Awad, M. (2019). Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture*, *9*(3), 54.

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y.,and You, L. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agric. For. Meteorol.*, *274*, 144–159.

Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., Warren, R., Qian,B., Daneshfar, B., Bedard, F., and Reichert, G. (2015). Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. *Agricultural and Forest Meteorology*, *206*, 137–150. https://doi.org/10.1016/j.agrformet.2015.03.007

Chlingaryan, A., Sukkarieh, S., and Whelan, B. (2018). Machine learning approaches for cropyield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, *151*, 61–69. https://doi.org/10.1016/j.compag.2018.05.012

Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environmental Research Letters, 13(11), 114003.

Cravero, A., Pardo, S., Sepúlveda, S., and Muñoz, L. (2022). Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review. Agronomy, 12(3), 748.

Goldstein, A., Fink, L., Meitin, A., Bohadana, S., Lutenberg, O., and Ravid, G. (2018). Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge. Precision agriculture, 19, 421-444.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. Deep Learning. MIT Press. Retrieved December 3, 2022, from https://www.deeplearningbook.org/

Gopal, P. S. M., and Bhargavi, R. (2019). Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Machine Learning Algorithms. Applied Artificial Intelligence, 33(7), 621–642. https://doi.org/10.1080/08839514.2019.1592343

Factors that influence crop yield. (2017). Retrieved from https://www.fertilizer.co.za/en/public-relations/news/2017/259-factors-that-influence-crop-yield

Filippi, P., Jones, E. J., Wimalathunge, N. S., Somarathna, P. D., Pozza, L. E., Ugbaje, S. U., ... and Bishop, T. F. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agriculture, 20, 1015-1029.

Javadinejad, S., Eslamian, S., and Ostad-Ali-Askari, K. (2021). The analysis of the most important climatic parameters affecting performance of crop variability in a changing climate. International journal of hydrology science and technology, 11(1), 1-25.

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., and 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

Khaki, S., Wang, L., and Archontoulis, S. V. (2020). A cnn-rnn framework for crop yieldprediction. *Front. Plant Sci.*, *10*, 1750.

Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., and Shearer, S. (2018). Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and electronics in agriculture*, *153*, 213-225.

Kuradusenge, M., Hitimana, E., Hanyurwimfura, D., Rukundo, P., Mtonga, K., Mukasine, A., ... and Uwamahoro, A. (2023). Crop yield prediction using machine learning models: case of Irish potato and maize. Agriculture, 13(1), 225.

Lokers, R., Knapen, R., Janssen, S., van Randen, Y., and Jansen, J. (2016). Analysis of Big Datatechnologies for use in agro-environmental science. *Environmental Modelling andAmp; Software*, *84*, 494–504. https://doi.org/10.1016/j.envsoft.2016.07.017

MARSWiki, 2020. MARS Crop Yield Forecasting System. https://marswiki.jrc.ec.europa. eu/agri4castwiki/index.php/Welcome_to_WikiMCYFS (Last accessed: May 11, 2020).

Newlands, N. K., Zamar, D. S., Kouadio, L. A., Zhang, Y., Chipanshi, A., Potgieter, A., Toure, S., and Hill, H. S. J. (2014). An integrated, probabilistic model for improved seasonal forecasting of agricultural crop yield under environmental uncertainty. *Frontiers in Environmental Science*, *2*. https://doi.org/10.3389/fenvs.2014.00017

Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., and Athanasiadis, I. N. (2021). Machine learning for large-scale crop yield forecasting. *Agricultural Systems*, *187*, 103016. https://doi.org/10.1016/j.agsy.2020.103016

Phalan, B., Green, R., and Balmford, A. (2014). Closing yield gaps: perils and possibilities for biodiversity conservation. Philosophical Transactions of the Royal Society B: Biological Sciences, 369(1639), 20120285.

https://doi.org/10.1098/rstb.2012.0285

Ranjan, A. K., and Parida, B. R. (2019). Paddy acreage mapping and yield prediction using sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India). Spatial Information Research, 27(4), 399-410.

Ransom, C. J., Kitchen, N. R., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernández, F. G., Franzen, D. W., Laboski, C. A., Myers, D. B., Nafziger, E. D., Sawyer, J. E., and Shanahan,

J. F. (2019). Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations. Computers and Electronics in Agriculture, 164, 104872. https://doi.org/10.1016/j.compag.2019.104872

Romero, J. R., Roncallo, P. F., Akkiraju, P. C., Ponzoni, I., Echenique, V. C., and Carballido, J. A. (2013). Using classification algorithms for predicting durum wheat yield in the province of Buenos Aires. Computers and Electronics in Agriculture, 96, 173–179. https://doi.org/10.1016/j.compag.2013.05.006

Shahhosseini, M., Martinez-Feria, R. A., Hu, G., and Archontoulis, S. V. (2019). Maize yield and nitrate loss prediction with machine learning algorithms. Environmental Research Letters, 14(12), 124026. https://doi.org/10.1088/1748-9326/ab5268

Taherei Ghazvinei, P., Hassanpour Darvishi, H., Mosavi, A., Yusof, K. B. W., Alizamir, M., Shamshirband, S., and Chau, K. W. (2018). Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network. Engineering Applications of Computational Fluid Mechanics, 12(1), 738-749.

USDA-NASS, 2012. The Yield Forecasting Program of NASS. Technical Report. United States Department of Agriculture (USDA). https://www.nass.usda.gov/Education_and_Outreach/Understanding_Statistics/Yield_Forecasting_Program.pdf

Van der Velde, M., Nisini, L., 2019. Performance of the MARS-crop yield forecasting system for the European Union: assessing accuracy, in-season, and year-to-year improvements from 1993 to 2015. Agric. Syst. 168, 203–212. https://doi.org/ 10.1016/j.agsy.2018.06.009.

Willcock, S., Martínez-López, J., Hooftman, D. A., Bagstad, K. J., Balbi, S., Marzo, A., Prato, C., Sciandrello, S., Signorello, G., Voigt, B., Villa, F., Bullock, J. M., and Athanasiadis, I. N. (2018). Machine learning for ecosystem services

Xu, X., Gao, P., Zhu, X., Guo, W., Ding, J., Li, C., Zhu, M., and Wu, X. (2019). Design of an integrated climatic assessment indicator (ICAI) for wheat production: A case study in Jiangsu Province, China. Ecological Indicators, 101, 943–953. https://doi.org/10.1016/j.ecolind.2019.01.059

Zhao, Y., Potgieter, A. B., Zhang, M., Wu, B., and Hammer, G. L. (2020). Predicting Wheat Yield at the Field Scale by Combining High-Resolution Sentinel-2 Satellite Imagery and Crop Modelling. Remote Sensing, 12(6), 1024. https://doi.org/10.3390/rs12061024

Zhong, H., Li, X., Lobell, D., Ermon, S., and Brandeau, M. L. (2018). Hierarchical modeling of seed variety yields and decision making for future planting plans. Environment Systems and Decisions, 38, 458-470