The Scientific Temper (2024) Vol. 15 (3): 2740-2747
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.3.43
https://scientifictemper.com/

RESEARCH ARTICLE

Crop yield prediction in diverse environmental conditions using ensemble learning

M. Menaha[1*], J. Lavanya[2]

Abstract

Precise assessment of crop yield is a vital component in agricultural planning and decision-making, having immediate consequences for food security and the allocation of resources. This study presents a new approach for predicting agricultural output in different climatic conditions by integrating the xgboost algorithm with the Whale Optimization Algorithm (WOA). XGBoost is a kind of ensemble learning method that enhances the accuracy of predictions by combining the results of many weak learners. However, the performance of the system may be significantly affected by the selection of suitable hyperparameters and feature subsets. To address this problem, we use the WOA algorithm, a nature-inspired optimization approach that mimics the foraging behavior of humpback whales. This technique is used to improve the parameters of xgboost and discover the most influential features. We evaluate the proposed model by using extensive datasets that include a diverse array of crops, soil compositions, climatic conditions, and geographic regions. The results suggest that the xgboost-WOA model outperforms traditional machine learning models in terms of both projected accuracy and efficiency. Furthermore, the suggested method showcases robust and reliable performance across different environmental circumstances, highlighting its potential for practical use in precision agriculture. This research emphasizes the effectiveness of combining AdaBoost with WOA for forecasting agricultural output. Furthermore, it contributes to the development of advanced predictive systems to support sustainable agricultural operations in adapting to climate variations and changing environmental conditions.

Keywords: Machine Learning, Crop Yield, Optimization, AdaBoost, WOA.

Introduction

Agriculture has a vital role in ensuring food security and sustaining economies worldwide. Accurate assessment of crop yield is essential for effective agricultural planning, resource allocation, and risk management. It enables farmers, policymakers, and stakeholders to make educated decisions that optimize production and avoid the adverse effects of environmental unpredictability. Predicting agricultural yields is a complex task due to several factors, including soil characteristics, weather conditions, farming methods, and geographical differences.

Machine learning (ML) techniques are used in several fields, including forecasting customer phone usage and analyzing consumer behavior in supermarkets. The use of machine learning in the agricultural sector has become widespread. Predicting agricultural yield is a difficult endeavor in the field of horticulture, and several studies have been conducted and assessed so far. This topic requires the use of many datasets since crop yield is affected by numerous variables, including soil quality, seed type, fertilizer treatment, and meteorological conditions. This demonstrates that the task of forecasting agricultural production is not a simple exchange but rather a sequence of complex procedures. Existing yield expectation models possess the ability to effectively ascertain the real yield, nevertheless, there is a want for enhanced yield forecasting.

Regression methods are used to forecast future outcomes, while descriptive models are utilized to extract insights from gathered data and provide explanations on past events. Developing a prediction model with extraordinary performance in the area of machine learning

[1]Department of Computer Science, Adaikalamatha College, Vallam, Affiliated to Bharathidasan University, Tiruchirappalli, India.

[2]Department of Artificial Intelligence and Machine Learning, Queens College of Arts and Science for Women, Punalkulam, Pudukottai, India.

***Corresponding Author:** M. Menaha, Department of Computer Science, Adaikalamatha College, Vallam, Affiliated to Bharathidasan University, Tiruchirappalli, India., E-Mail: Menaha.m1989@gmail.com

How to cite this article: Menaha, M., Lavanya, J. (2024). Crop yield prediction in diverse environmental conditions using ensemble learning. The Scientific Temper, **15**(3):2740-2747.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.3.43

Source of support: Nil

Conflict of interest: None.

Received: 06/07/2024 **Accepted:** 10/08/2024 **Published** : 16/09/2024

has several obstacles. Traditional statistical methods and simple machine learning models often struggle to capture the intricate relationships among these diverse components, leading to suboptimal predictive accuracy. To address these challenges, advanced machine learning methods, such as ensemble techniques, have gained more prominence in recent years. The xgboost method is a machine learning strategy that seeks to improve prediction accuracy by leveraging the strengths of several weak learners. The efficacy of xgboost is highly dependent on the meticulous choice of hyperparameters and pertinent attributes. The efficacy of xgboost is significantly impacted by the meticulous choice of hyperparameters and pertinent attributes.

This research introduces a novel approach that integrates xgboost with the Whale Optimization Algorithm (WOA) to enhance the precision of forecasting agricultural output across various climatic conditions. The WOA algorithm, which is influenced by the unique foraging behavior of humpback whales, is well-known for its remarkable effectiveness in finding optimal solutions inside complex search contexts. The aim of this research is to enhance the predictive accuracy of the model by tweaking the hyper parameters of xgboost and doing feature optimization using the WOA algorithm.

We conduct extensive experiments using diverse datasets, including a wide array of crops, soil compositions, meteorological circumstances, and geographical regions. The research results suggest that the xgboost-WOA model outperforms traditional machine learning approaches in terms of accuracy. In addition, the model consistently performs well under different environmental situations. This research emphasizes the potential of improving agricultural output forecasting by combining ensemble learning with nature-inspired optimization. This study contributes to the broader goal of promoting sustainable agriculture and enhancing food security by providing a more reliable forecasting tool in the face of climate change and environmental uncertainty.

Millet crop yield predicting technique was developed by using a large amount of complex data (Champaneri, M., *et al*., 2020). By using the Random Forest Classifier, we achieved a prediction accuracy of 99.74% for millet crop output. This was accomplished by taking into account several input variables, including soil type, minimum temperature, maximum temperature, humidity, rainfall, and other relevant factors.

An innovative research project is found to examine the prediction of crop yields that would result in high production in a certain location by considering the climate and soil conditions (Chourasiya, N.L., *et al*., 2019). The current research on seed classification is considered unproductive since it lacks a validation technique. Throughout this inquiry, the author made an effort to create a prediction model utilizing a machine-learning algorithm. The purpose of this model was to help farmers estimate seed planting, ultimately leading to higher agricultural productivity. This event employs analytical machine learning algorithms to extract insights from data, enabling predictions, realistic simulations, and the identification and categorization of patterns in the input file. A human-developed neural network is used to model complex relationships between input and output variables or to identify patterns in datasets. The purpose of this study is to get a thorough comprehension of the machine learning method that utilizes neural networks and create a model capable of predicting seed categories by using machine learning techniques. The model undergoes testing using a seed dataset and thereafter makes predictions on seed categories.

A new attempt was made to predict agricultural output by analyzing the factors of rainfall, crop type, climatic conditions, area, production, and yield (Sharma, P., *et al*., 2023). These characteristics have been recognized as crucial issues that pose a threat to the long-term viability of agriculture. Crop yield prediction is a sophisticated technology that employs machine learning (ML) and deep learning (DL) approaches to aid in decision-making about crop selection and the management of its development throughout the course of the season. It has the capacity to choose the specific crops to cultivate and make determinations about the tasks to be performed throughout the whole growth cycle of the crop. ML and DL algorithms are used in crop selection to mitigate agricultural output losses, even in the face of various distractions. Machine learning algorithms, including decision trees, random forests, and XGBoost regression, as well as deep learning techniques like conventional neural networks (CNN) and long-short-term memory networks, have been used to forecast agricultural production.

The massive use of an ANN architecture was proposed for data modeling in predictive tasks (Bohra, R., *et al*., 2024). The approach has four stages: historical data analysis, data preparation, data modeling, and performance estimate. Firstly, classify the data based on several attributes. Regression analysis uses CNN to investigate the correlation between an independent (predictor) variable and a dependent (target) variable. The model will use relation training to precisely predict agricultural output production. ANN is used to predict agricultural output by using many crop performance measurements as input parameters.

The benefit of machine learning methods was recommended to examine various data sets with the aim of predicting agricultural yield (Pujitha, B., *et al*., 2024). The main aim is to assess soil parameters such as rainfall, temperature, humidity, pH levels, nitrogen (N), phosphorous (P), and potassium (K). This comprehensive research has established a basic correlation, allowing accurate predictions in the ever-changing agricultural industry. The proposed technique utilizes the Light GBM algorithm, a kind of ensemble learning, which obtained a prediction accuracy rate of 99%. This proposed technique would function as a significant asset for farmers to optimize their agricultural production.

The machine learning method was proposed by employing Decision Trees (DT), Support Vector Machines (SVM), Naive Bayes (NB), K-Nearest Neighbors (K-NN), Extreme Gradient Boosting, and Random Forest (RF) (Swathi, T., & Sudha, S. 2023). The dataset used in this research was acquired from the Kaggle website and comprises six unique crop kinds, each linked to 11 specific nutrients. The models are trained using 80% of the dataset and evaluated using the remaining 20%. The results indicate that extreme gradient boosting and Naive Bayes perform better than other models, with accuracy scores of 0.994 and 0.993, respectively.

Support Vector Machines (SVM) and Artificial Neural Networks (ANN) was proposed and crop prediction is performed by considering variables such as precipitation levels, minimum and maximum temperatures, soil composition, humidity, and soil pH levels (Fegade, T.K., & Pawar, B.V. 2019). The data is sourced from the official agriculture website of Maharashtra and is categorized into nine agricultural zones. A user interface is being created to enable farmers to enter essential data for crop prediction. The neural network has a prediction accuracy of 86.80%.

A meticulous model that accurately predicts soil series by taking into account the kind of land (Rahman, S.A., *et al.*, 2018). Moreover, it has the capability to provide suggestions for appropriate crops depending on the projections. Various machine learning techniques, including weighted k-NN, Bagged Trees, and Gaussian kernel-based SVM, are used for soil classification. The empirical evidence suggests that the SVM-based strategy outperforms other existing techniques.

A new methodology that examines past data in order to forecast agricultural output. ML approaches, such as support vector machines (SVM) and random forest (RF), are used to analyze agricultural data (Bondre, D.A., & Mahagaonkar, S. 2019). The algorithms thereafter provide suggestions for the optimal fertilizer to be used for each individual crop. The work focuses on developing a prognostic model that may be used to estimate forthcoming crop yields. This is a succinct analysis of forecasting agricultural output using the use of machine learning.

The nutritional deficiency in a paddy crop was investigated (Shidnal S., *et al.*, 2019). TensorFlow, the machine learning framework developed by Google, is used to build a neural network capable of autonomously classifying plants into categories of nitrogen, potassium, phosphorus deficiencies, or healthy. It is crucial to maintain a suitable balance of nitrogen, potassium, and phosphorus levels. The Tensor Flow model identifies deficiencies by evaluating a set of photographs. The result is entered into a layer powered by machine learning to quantify the degree of inadequacy. It employs the k-means clustering method. The estimation of the cropland's yield is determined by analyzing it using a rule matrix. Two fatigued machine learning algorithms produced a pretty precise prognosis of 76 to 77%.

Proposed Work

Accurately forecasting crop production under varying environmental circumstances is essential for guaranteeing food security, improving agricultural methods, and reducing economic risks for farmers. Traditional approaches to yield prediction often fail to include the complex and ever-changing connections among climatic factors, soil attributes, and crop traits. This work presents a novel approach that combines the whale optimization algorithm (WOA) with the xgboost ensemble learning methodology to enhance the accuracy and dependability of agricultural output forecasts.

The primary objectives of this study are:

- The goal is to develop a hybrid model that integrates the WOA with the xgboost Classifier in order to predict agricultural yield under different environmental situations.
- The efficacy of the proposed model will be evaluated by comparing it to existing prediction models.
- The objective is to identify the key environmental factors that impact agricultural productivity and analyze their consequences.

Steps in Classification

The proposed methodology aims to provide accurate and reliable predictions of crop yield, aiding farmers and policymakers in making educated decisions for sustainable agricultural methods. The method for the recommended work is explained in Figure 1.

Input Data Collection

Collect previous from many sources, including weather stations, agricultural research stations, remote sensing platforms, and Internet of Things devices. Make careful to include elements such as soil moisture, temperature, rainfall, humidity, crop type, planting date, fertilizer levels, and other pertinent aspects.

Pre-Processing

Apply techniques such as mean/mode imputation to handle missing data. Replace the missing numbers in the column with the average and middle values, which can be calculated using equations (1) and (2).

$$x_i = \frac{1}{N}\sum_{j=1}^{N} x_j \tag{1}$$

for all x_i that are missing

$$x_i = median(x_1, x_2, \dots \dots, x_n) \tag{2}$$

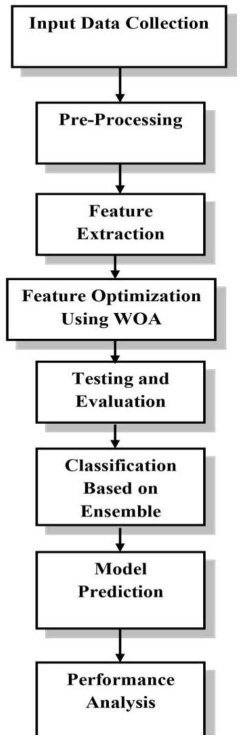

Figure 1: Classification steps in proposed work

Feature Extraction and Selection

Feature extraction is a crucial part of the preliminary process for forecasting crop yields. Data mining involves extracting significant new features from the original data in order to better grasp underlying patterns and connections. This approach improves the effectiveness of machine learning models by providing them with more informative inputs. The model is trained and evaluated using 10 characteristics derived from the dataset.

Crop

The name of the crop cultivated.

Crop_Year

The year in which the crop was grown.

Season

The specific cropping season (e.g., Kharif, Rabi, Whole Year).

State

The Indian state where the crop was cultivated.

Area

The total land area (in hectares) under cultivation for the specific crop.

Production

The quantity of crop production (in metric tons).

Annual_Rainfall

The annual rainfall received in the crop-growing region (in mm).

Fertilizer

The total amount of fertilizer used for the crop (in kilograms).

Pesticide

The total amount of pesticide used for the crop (in kilograms).

Yield

The calculated crop yield (production per unit area).

The research utilizes the principle component analysis (PCA) methodology for feature selection. PCA is a technique used to decrease the dimensionality of a dataset by transforming a group of potentially interrelated variables into a fresh set of uncorrelated variables, referred to as principle components. Principle components have the potential to capture the most degree of variability in the data. PCA may be used to perform feature selection by retaining just the most significant components.

To choose or make a selection choose the k primary components that have the greatest variation. The cumulative explained variance ratio may be used to determine the number of components.

$$Explainedvarianceratio = \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j} \quad (3)$$

$$CummulativeExplainedvarianceratio = \sum_{i=1}^{k} \frac{\lambda_i}{\sum_{j=1}^{p} \lambda_j} \quad (4)$$

Feature Optimization

This research suggests using the bio-inspired WOA to optimize features. Whales engage in collective behavior to overcome the challenges of finding resources, performing three specific tasks: surrounding, narrowing, and catching prey. Hunting is used for the purpose of discovery, whereas encircling and shrinking operations are utilized for the purpose of exploitation. Issues pertaining to dimensional optimization (DO) are addressed using cutting-edge xth-generation methods.

Step 1: Encircling Operation

The search space is initialized with a stochastic distribution of whale locations (Xi), where m denotes the whole population. The selected position of the search agent is regarded as the optimal place for finding prey. Alternative methods use the following equation to modify their positions in relation to the most optimal search agents.

$$X_{i+1} = X_i - P|QX_i^* - X_i| \quad (5)$$

The optimal whale placement is denoted by Xi*, while the whale placement at the end of iteration Xi+1 denotes me.

$$P = 2mr_1 - m \quad (6)$$

$$Q = 2r_2 \quad (7)$$

Gradually decrease m from 2 to 0 in repetitions, with random r1 and r2 values.

Step 2: Bubble net prying

Humpback whales use a bubble net predation approach with a shrink encircling mechanism and spiral ascent, with a half likelihood for local optimization. The shrink-wrapping mechanism is derived by changing P values in equations (7), (8) and (9).

$$X_{i+1} = \begin{cases} X_i - P|QX_i^* - X_i|, & th < 0.5 \\ |X_i^* - X_i|e^{ab}\cos(2\Pi b) + X_i^* & th \geq 0.5 \end{cases} \quad (8)$$

Randomly selected th, b from [0,1], and a controls spiral shape with integer control.

Step 3: Food Searching Phase

The process changes P, and random solutions are used for updating whale locations, causing the whale to get $|\vec{P}|$ far from the best possible solution. The WOA algorithm's worldwide search is modeled in Equation 21.

$$X_{i+1} = X_{rand} - P|QX_{rand} - X_i| \quad (9)$$

where Xrand is a random point within the whale population.

The WOA uses a random sample and a mathematical model for bubble-net and circling hunting to revise agents' coordinates, ensuring convergence and acting as a fulcrum if not:

Pseudo code for WOA
Step 1: initiate population of Whale Xi
Step 2: initiate P, m and Q
Step 3: Evaluate the fitness of search agent
Step 4: Let take i=1
Step 5: whilei $\leq$ max iteration
Step 6: for the entire agent do
Step 7: if |P|=1 then
Step 8: Update the current search position
Step 9: if $|p| \neq 1$ then
Step 10: Select random search agent Xrand
Step 11: Change the current location of the search
Step 12: End if
Step 13: End for
Step 14: revise P, m and Q
Step 15: Update X*
Step 16: i=i+1
Step 17: End While
Step 18: return X*
Step 19: end

Prediction of crop yield

This study proposes the use of an ensemble methodology to forecast agricultural production. XGBoost is a well-known ensemble learning approach that combines many weak classifiers to create a powerful classifier. The technique operates by iteratively training weak classifiers, with each subsequent classifier assigning more significance to the cases that were erroneously classified by the preceding classifiers.

The xgboost technique operates by iteratively training a series of weak classifiers, with each subsequent classifier assigning more significance to the data points that were misclassified by the preceding classifiers. The best model is created by merging the weak classifiers, with each classifier being assigned a certain weight.

Initialize weights

Initially, all training samples are given equal weights by using eqn (10)

$$w_i^t \frac{1}{N} for i = 1,2, \ldots . . N \quad (10)$$

Train Weak Classifier

For each iteration t from 1 to T

Train a weak classifier $h_t(x)$ on the weighted training data.

Calculate the weighted error ϵ_t

$$\epsilon_t = \frac{\sum_{i=1}^{N} w_i^t . I(y_i \neq h_t(x_i))}{\sum_{i=1}^{N} w_i^t} \quad (11)$$

Here I(.) is the indicator function that returns 1 if the condition is true and 0 otherwise.

Compute the classifier weight α_t

$$\alpha_t = \frac{1}{2} In \frac{1-\epsilon_t}{\epsilon_t} \quad (12)$$

Update the weights of the training samples

$$w_i^{t+1} = w_i^t . \exp(\alpha_t . I(y_i \neq h_t(x_i)) \quad (13)$$

Normalize the weights to ensure they sum to 1

$$w_i^{t+1} = \frac{w_i^{t+1}}{\sum_{j=1}^{N} w_j^{t+1}} \quad (14)$$

Create strong classifier

Combine the weak classifiers to form the final strong classifier.

$$H(x) = sign \sum_{t=1}^{T} \alpha_t . h_t(x) \quad (15)$$

Performance Analysis

Dataset

The given CSV file contains agricultural statistics on many crops cultivated in various states of India from 1997 to 2020. The dataset is purposefully constructed to predict crop yields by taking into account several agronomic factors, such as weather conditions, application of fertilizers and pesticides, and other relevant variables. The dataset is presented in a tabular manner, with each row including data pertaining to a specific crop and its corresponding features. The dataset has 19,698 rows and 10 columns, including 9 attributes and 1 label. The information includes fundamental attributes required for forecasting agricultural output, such as crop types, cropping seasons, crop years, states, cultivated areas, production quantities, annual rainfall, fertilizer use, pesticide usage, and calculated yields. Table 1 presents the precise details of the selected sample input taken from the dataset used in this research investigation.

Performance Analysis

Figure 2 presents a sequential analysis of the crops. The data clearly shows that the yield has been increasing throughout the year, but since 2014, it has been consistently decreasing. Potential factors contributing to this issue are climate change and a reduction in soil fertility. Between 2015 and 2020, there was an initial surge followed by a slow decline.

Table 1: Sample data from database

	Crop_Year	*Area*	*Production*	*Annual_Rainfall*	*Fertilizer*	*Pesticide*	*Yield*
Count	19689.00000	1.968900e+04	1.968900e+04	19689.00000	1.968900e+04	1.968900e+04	19689.000000
Mean	2009.127584	1.799266e+05	1.643594e+07	1437.755177	2.410331e+07	4.884835e+04	79.954009
Std	6.498099	7.328287e+05	2.630568e+08	816.909589	9.494600e+07	2.132874e+05	878.306193
Min	1997.000000	5.000000e-01	0.000000e+00	301.300000	5.417000e+01	9.000000e-02	0.000000
25%	2004.000000	1.390000e+03	1.393000e+03	940.700000	1.880146e+05	3.567000e+02	0.600000
50%	2010.000000	9.317000e+03	1.380400e+04	1247.600000	1.234957e+06	2.421900e+03	1.030000
75%	2015.000000	7.511200e+04	1.227180e+05	1643.700000	1.000385e+07	2.004170e+04	2.388889
Max	2020.000000	5.080810e+07	6.326000e+09	6552.700000	4.835407e+09	1.575051e+07	21105.000000

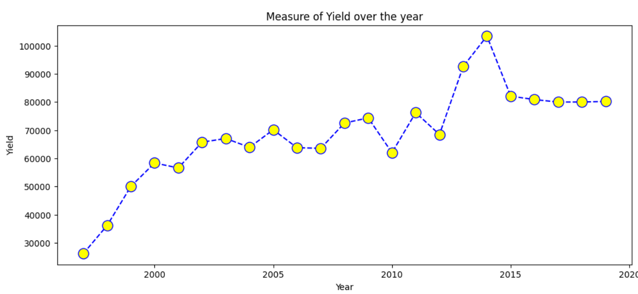

Figure 2: Year-wise analysis of crop yields

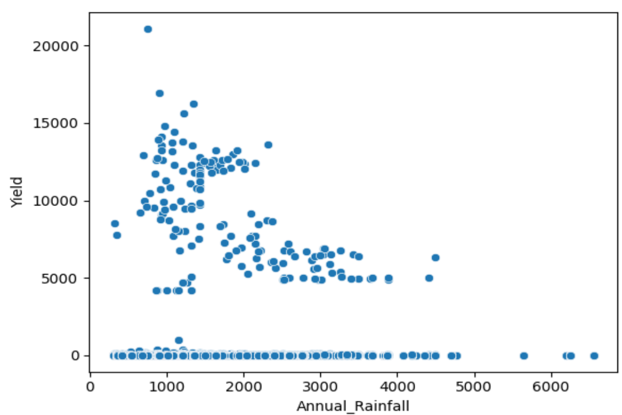

Figure 3: Annual rainfall

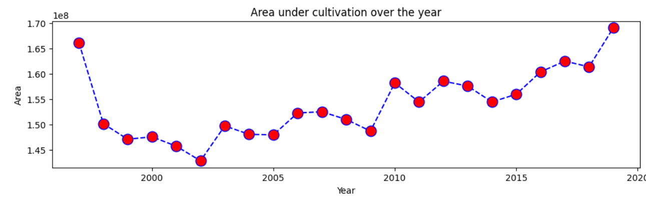

Figure 4: Area under cultivation

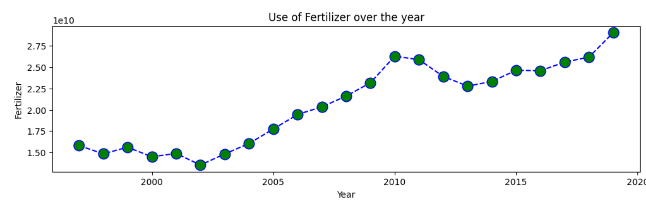

Figure 5: Use of fertilizer

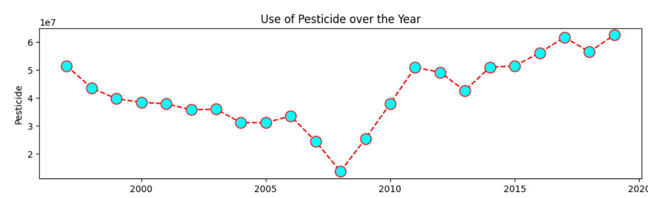

Figure 6: Use of pesticides

The yearly precipitation is a critical determinant in predicting agricultural output since it directly impacts water availability, soil moisture content, and overall plant health. Adequate precipitation ensures that crops get the requisite quantity of water for essential physiological processes, hence influencing germination, growth, and maturation stages. Figure 3 illustrates the projected annual precipitation and its impact on agricultural yield. Research has shown that the yield or production experiences a substantial boost in the presence of an adequate amount of rainfall.

Figure 4 clearly illustrates a substantial increase in the cultivated area. Either fallow land is being farmed via the use of fertilizer and enhanced irrigation, or wooded areas are being transformed for agricultural purposes.

Figure 5 depicts the examination of fertilizer use across the years. Following the year 2006-2007, there was a significant surge in the use of fertilizer with the aim of achieving higher crop yields. Furthermore, there was a noticeable rise in the use of pesticides starting from 2008, as seen in Figure 6.

Fertilizer is used for crops in Gujarat and Bihar, regions characterized by either low or below-average annual precipitation. Diverse methods are being used in the agriculture sector to augment soil moisture levels. Fertilization, as seen in Figure 7, is a specific method or technique.

Figure 8 depicts the application of fertilizer in different types of crops. The Rice Crop requires the largest amount of fertilizer, followed by the Wheat crop. The production of rice and wheat crops occupies a larger land area.

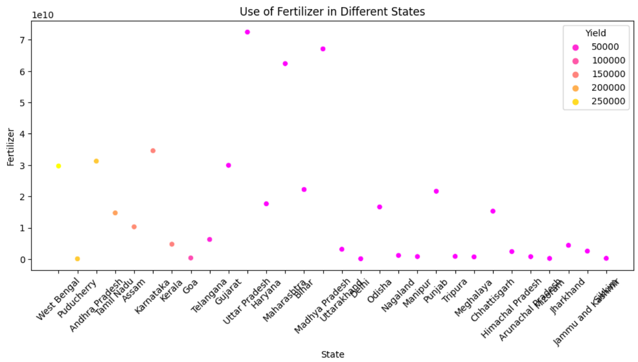

Figure 7: Use of fertilizer across all states

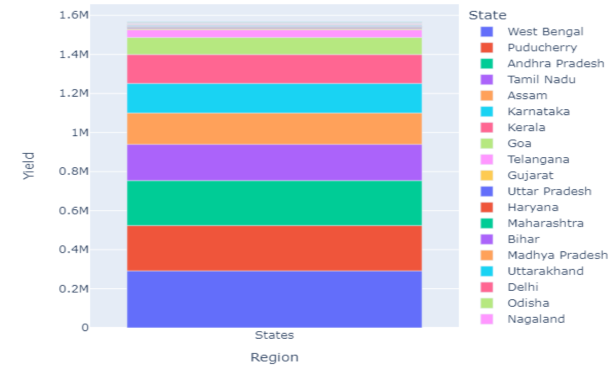

Figure 9: State wise prediction

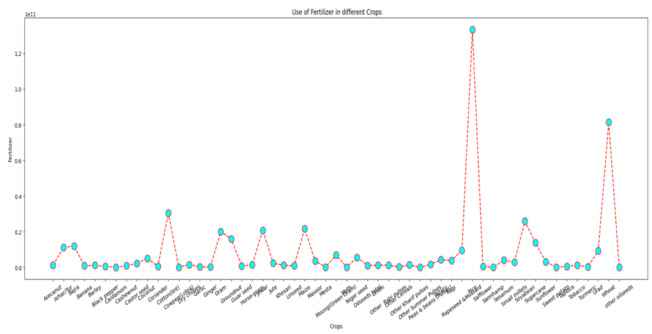

Figure 8: Use of fertilizer on different crops

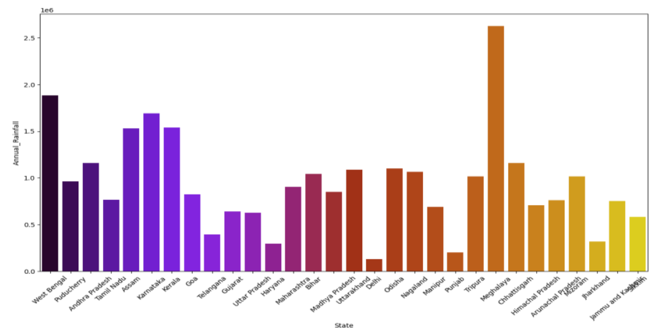

Figure 10: Annual rain across the state

Table 2: Performance comparison

Method	*Training_Acc*	*Testing_Acc*
Linear regression	0.86	0.82
Random forest	0.78	0.81
Supprt vector	0.81	0.76
AdaBoost	0.89	0.97
WOA+Adaboost	0.98	0.97

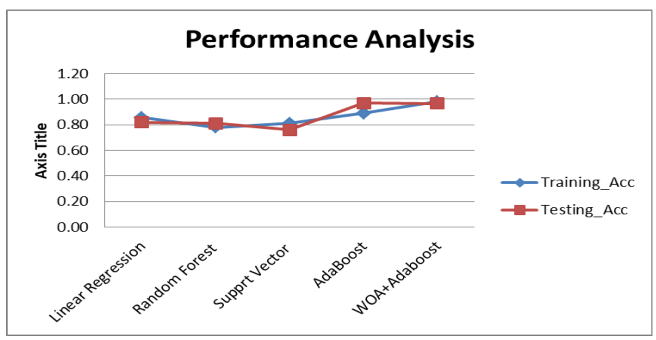

Figure 11: Performance comparison of proposed vs existing

The total yield production across all states is projected and shown in Figure 9. Nagaland, Odisha, and Goa have the highest yield in comparison to the other states, which is clearly apparent. The elevated yearly precipitation in these states, as seen in Figure 10, may be attributed to their geographical location.

Chattisgarh has the maximum yearly rainfall, yet it does not have the highest crop production. West Bengal has the greatest agricultural yield. Uttar Pradesh, Haryana, and Maharashtra are using a substantial amount of fertilizer despite the fact that the agricultural yield remains low, perhaps owing to inadequate annual rainfall.

Table 2 displays the performance analysis of the planned task, whereas Figure 11 presents a graphical representation of the data. The ensemble strategy demonstrates superior performance compared to other machine-learning techniques in crop prediction. The hybrid WOA+xgboost classifier outperforms the xgboost classifier in the ensemble technique. By improving the fundamental characteristics associated with the core, the accuracy performance is improved.

Conclusion

The integration of WOA with the xgboost classifier for agricultural production prediction demonstrates significant improvements in predictive accuracy and processing efficiency. This hybrid technique leverages the capabilities of WOA to both explore and exploit in order to improve the performance of the xgboost classifier. This enhances the classifier's efficiency in handling intricate and non-linear agricultural information. The experimental results show that the WOA-xgboost model outperforms traditional methods by effectively capturing the intricate relationships among several environmental and agronomic factors that impact crop yield.

Furthermore, the adaptability of WOA in investigating the range of possible solutions and its compliance with xgboost's boosting mechanism provide robust model training, minimizing overfitting and improving generalization. This work emphasizes the potential of combining nature-inspired optimization techniques with machine learning algorithms to address challenges in agricultural data processing.

In summary, the WOA-xgboost classifier presents a practical method for accurate agricultural output forecasting, providing valuable insights for farmers, agronomists, and policymakers. Further study might explore the integration of other data sources and the application of this combined model to various crops and geographical regions to validate and enhance its usefulness.

Acknowledgment

We would like to thank Adaikalamatha College, Vallam, which is affiliated with Bharathidasan University, Tamilnadu, for the helpful support in conducting research in an effective manner.

References

Ayalew, A.T., and Lohani, T.K. (2023). Prediction of Crop Yield by Support Vector Machine Coupled with Deep Learning Algorithm Procedures in Lower Kulfo Watershed of Ethiopia. *Journal of Engineering*.

Beulah.A (2019), "A Survey on Different Data "ining Techniques for Crop Yield Prediction," *International Journal of C"mputer Sciences and Engineering*, Vol.7, Issue.1, pp.738-744.

Bhojani, S.H., & Bhatt, N. (2020). Wheat crop yield prediction using new activation functions in neural network. *Neural Computing and Applications, 32*, 13941 - 13951.

Bohra, R., More, S., Kamat, S., Pathak, R., & Shitole, P.M. (2024). Crop Yield Prediction using CNN. *International Journal of Advanced Research in Science, Communication and Technology*.

Bondre, D.A., & Mahagaonkar, S. (2019). PREDICTION OF CROP YIELD AND FERTILIZER RECOMMENDATION USING MACHINE LEARNING ALGORITHMS. *International Journal of Engineering Applied Sciences and Technology*.

Champaneri, M., Chachpara, D., Chandvidkar, C., & Rathod, M. (2020). Crop Yield Prediction using Machine Learning. *ADALYA JOURNAL*.

Chourasiya, N.L., Modi, P.P., Shaikh, N.F., Khandagale, D.J., & Pawar, S. (2019). Crop Prediction using Machine Learning.

Dharani, M.K., Thamilselvan, R., Natesan, P., Kalaivaani, P.C., & Santhoshkumar, S. (2021). Review on Crop Prediction Using Deep Learning Techniques. *Journal of Physics: Conference Series, 1767*.

Elavarasan, D., Durai, P.M., & Vincent, R. (2021). A reinforced random forest model for enhanced crop yield prediction by integrating agrarian parameters. *Journal of Ambient Intelligence and Humanized Computing, 12*, 10009 - 10022.

Fegade, T.K., & Pawar, B.V. (2019). Crop Prediction Using Artificial Neural Network and Support Vector Machine. International Journal of *Data Management, Analytics and Innovation*.

Jiménez López, F.R., Jiménez López, A.F., & Castellanos Patiño, J.S. (2023). Forecasting irrigation scheduling based on deep learning models using IoT. *Proceedings of the 21th LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI 2023): "Leadership in Education and"Innovation in Engineering in the Framework of Global Transformations: Integration and Alliances for Integral Development"*.

K.S. Leelavathi, D., & Ra"asenathipathi, M. (2023). An EWOA-WTDCNN Model: Effective Whale Optimization Algorithm Incorporating with Weight-Tuned Deep Convolutional Neural Network for Forecasting Indian Crop Yielding Process. *Tuijin Jishu/Journal of Propulsion Technology*.

Muruganandam.C, Maniraj.V (2024). A self-driven dual reinforcement model with meta heuristic framework to conquer the iot based clustering to enhance agriculture production. The Scientific Temper (202 Vol. 15 (2): 2169-2180*.

Nosratabadi, S., Felde, I., Széll, K., Ardabili, S.F., Beszédes, B., & Mosavi, A.H. (2020). Hybrid Machine Learning Models for Crop Yield Prediction. ArXiv, abs/2005.04155.

Pujitha, B., Radha, K., Samitha, K., Prathyusha, N., & Vasanthi, P. (2024). Crop Prediction from Soil Parameters using Light Ensemble Learning Model. *2024 11th International Conference on Signal Processing and Integrated Networks (SPIN)*, 308-313.

Rahman, S.A., Chandra Mitra, K., & Mohidul Islam, S.M. (2018). Soil Classification Using Machine Learning Methods and Crop Suggestion Based on Soil Series. *2018 21st International Conference of Computer and Information Technology (ICCIT)*, 1-4.

Sharma, P., Dadheech, P., Aneja, N., & Aneja, S. (2023). Predicting Agriculture Yields Based on Machine Learning Using Regression and Deep Learning. *IEEE Access, 11*, 111255-111264.

Shidnal, S., Latte, M.V., & Kapoor, A. (2019). Crop yield prediction: two-tiered machine learning model approach. *International Journal of Information Technology, 13*, 1983 - 1991.

Swathi, T., & Sudha, S. (2023). Crop classification and prediction based on soil nutrition using machine learning methods. *International Journal of Information Technology, 15*, 2951-2960.