Journal of Information Systems and Informatics

Vol. 5, No. 2, June 2023 e-ISSN: **2656-4882** p-ISSN: **2656-5935**

DOI: 10.51519/journalisi.v5i2.506 Published By DRPM-UBD

Rice Yield Forecasting: A Comparative Analysis of Multiple Machine Learning Algorithms

Eli Adama Jiya[1,*], Umar Illiyasu[2], Akinyemi Mudashiru[3]

[1,2]Department of Computer Science, Federal University Dutsinma, Katsina State, Nigeria.
[3]Department of Agricultural Economics, Federal University Dutsinma, Katsina State, Nigeria.
Email: [1]jiyaeli@fudustinma.edu.ng, [2]uiliyasu@fudustinma.edu.ng, [3]makiyemi@fudustinma.edu.ng

Abstract

Agriculture plays a crucial role in Nigeria's economy, serving as a vital source of sustenance and livelihood for numerous Nigerians. With the escalating impact of climate change on crop yields, it becomes imperative to develop models that can effectively study and predict rice output under varying climatic conditions. This study collected rice yield data from Katsina state, spanning the years 1970 to 2017, sourced from the Nigeria Bureau of Statistics. Additionally, climatic data for the same period were obtained from the World Bank Climate Knowledge portal. Logistic Regression (LR), Artificial Neural Network (ANN), Random Forest (RF), Random Trees (RT), and Naïve Bayes (NB) were employed to develop rice yield prediction models utilizing this dataset. The findings reveal that random forest and random trees exhibited superior classification performance for yield prediction. The developed models offer a promising tool for predicting future rice yields, facilitating proactive measures to ensure food security for the people of the state.

Key word: Machine Learning, Yield prediction, crop yield prediction, Rice yield in Katsina state.

1. INTRODUCTION

Agriculture plays a crucial role in Nigeria's economy, serving as a primary source of sustenance for the population and a significant means of livelihood for numerous Nigerians [1], [2]. Among the diverse forms of agriculture, crop cultivation stands out as the predominant agricultural activity in Nigeria [3]. Consequently, any failures in crop production have profound implications for families and the overall economy. Notably, rice cultivation emerges as one of the prominent agricultural practices in Nigeria, spanning across almost every state in the country.

Rice holds a prominent position as a staple food in Nigeria, being widely consumed by both the impoverished and affluent segments of society [4]. Its significant consumption in terms of tons per year highlights its vital role in the country [4]. Consequently, any failures in rice production would have a substantial

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

impact on Nigeria's food security status [5]-[8]. The thriving of agriculture depends on favorable climatic and environmental conditions, which must be in proper proportion [9]. However, the escalating effects of climate change, including erratic rainfall patterns, increasing droughts, and rising temperatures, have rendered rice yield highly uncertain, exposing nations to food insecurity and hunger.

Katsina State, which experiences the dual challenges of climate change and encroaching desertification [10], is particularly affected by these issues. Moreover, it houses a significant population of rice consumers. Although some research has been conducted in other states such as Kogi [11] and Ebonyi [7], limited studies have focused on predicting rice yield in Katsina using climate variables. Therefore, to facilitate effective policy formulation, planning, and intervention for food security in Katsina State, the development of rice yield prediction models incorporating climate variables becomes crucial [12]. This study aims to accomplish this by leveraging machine learning algorithms, which can provide accurate and cost-effective estimations of farm output in the face of changing climatic conditions. Such models will greatly assist policymakers in preparing for potential rice failures and ensuring food sufficiency [13].

Crop yield is influenced by various factors, and the impact of climate change on crop production has become an increasingly researched topic [14]-[18]. Previous studies, such as [19] and [14], have employed machine learning (ML) techniques to identify the factors influencing crop yield. ML algorithms, including artificial neural networks, fuzzy logic, support vector machines, k-nearest neighbor, genetic algorithms, and more, offer computational tools that yield cost-effective and accurate estimates while extracting vital information for yield prediction [18]-[24]. The accuracy of each developed model is evaluated using metrics such as Root Mean Square Error (RMSE), Root-Relative Square Error (RRSE), and Mean Absolute Error (MAE) to assess their efficacy.

Numerous studies have been conducted to predict crop yield using climatic, soil, or remote sensing data [14], [15], [25]–[27]. For instance, [28] formulated a prediction model for rice yield in Jigawa State, Nigeria, using fuzzy logic. The model incorporated rainfall, land, and previous rice yield as basic features, successfully forecasting the required amount of rain for achieving optimal future yields. Another study by [29] explored the relationship between rice yield and climate change, utilizing precipitation, evaporation, temperature, wind, and sunshine as climatic variables. The findings suggested that the model's accuracy was sufficiently reliable for future rice yield forecasts in Sri Lanka. Similarly, [11] employed the Random Forest algorithm to model yield in Kogi State, Nigeria, with rainfall, wind, and temperature as input features.

In the realm of climate change and its impact on crop yields, [30] utilized weather forecast data to study the effects. They found that rainfall and air temperature

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

played more significant roles compared to other predictors such as solar radiation, air humidity, soil moisture, and wind speed. [31] examined the predictive accuracy of machine learning and regression approaches for crop production prediction using ten agricultural datasets. The M5-Prime and k-closest neighbor models demonstrated high levels of accuracy among all the methods considered. Specifically, the M5-Prime model achieved RMSEs of 5.14, 79.46%, and 18.12%, as well as RRSEs of 79.46% and MAEs of 18.12%.

Prediction models for multiple crop yields have also been explored. [32] conducted a study using Gradient Boosting, Support Vector Regression (SVR), and k-Nearest Neighbors, along with crop models, to predict yields for various crops in the Netherlands, Germany, and France. The models incorporated weather, remote sensing, and soil data as input features. Similarly, [33] utilized random forest techniques to forecast cotton yield in Maharashtra, India, highlighting the wide usage of soil, climate, and solar parameters in predicting crop yield, among other factors. [34] focused on predicting potato tuber yield, employing four machine learning algorithms: linear regression, elastic net, k-nearest neighbor, and support vector regression. Their results indicated that Support Vector Regression performed the best, with RMSE values of 5.97, 4.62, 6.60, and 6.17 t/ha for different years.

In Rwanda, [35] used the Aqua Crop model to predict maize yield under rainfed agriculture in the Eastern province. The study analyzed various climatic parameters, including temperature, rainfall, evapotranspiration, and maize yield. Notably, the research revealed no significant impact of rainfall trends on crop yield during the considered study period. Additionally, [36] developed a Deep Neural Network-based solution to predict and assess the yield of corn hybrids using environmental and genotype data as part of the 2018 Syngenta Crop Challenge. Their model demonstrated high accuracy, achieving an RMSE of 12% of the average yield and 50% of the standard deviation for the validation dataset, utilizing predicted weather data. Although numerous machine learning predictions have been conducted on rice yield, further research is still needed in this field.

2. METHODS

The methods adopted in this work follows the regular data mining procedure which include data collection, data preprocessing, choosing learning algorithm and training them to produce rice yield prediction model. Briefly, the diagram in Figure 1 captures the method conceptually.

2.1 Data source

The study location is katsina state in northern Nigeria, it has border with the republic of Niger and is a semi-desert area. Data on rice yield from katsina state in

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Nigeria was collected from Nigeria Bureau of statistics (NBS) while some climatic data was downloaded from World Bank climate knowledge portal. The data from NBS contains annual records of rice yield in katsina state from 1970 to 2017 with the following attributes: elevation, max temperature, min temperature, wind, relative humidity, and Yield/metric tons. While the World Bank dataset contains precipitation, and average temperature.

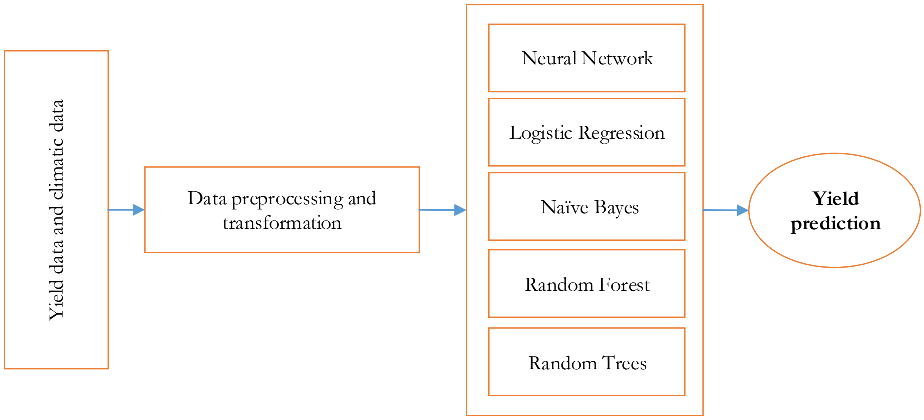

Figure 1. The conceptual model of yield forecasting

2.2 Feature selection

One crucial aspect of the project is the feature selection. In this work we use the entire features or attributes earlier introduced. The features focus on climatic and environmental variables which affect yield in the area. These features were also used to study the relationship between climate change and yield in southern part of Nigeria [4]. At a continental level, [25] and [1], used these variables. Table 1 give a statistical detail of the data. In the table, the minimum and maximum value for each of the attributed are given with the mean and standard deviation.

Table 1. Statistical description of the data

	Max temp	Min temp	wind	Humidity	precipitation	Yield/metric ton
Mean	34.24873	13.95221	3.149485	0.142122671	750.6760417	489.9427067
Std	2.672609	1.862362	0.607981	0.05043383	114.023639	80.07498758
Min	30.09	10.445	1.756891	0.061683155	508.88	352.0555556
Max	40.249	17.644	4.52383	0.266961099	994.37	663.6944444

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

2.3 Data preparation

A further transformation of data was carried out with each attribute values reduced to maximum of 1, this was done by dividing by the highest values of each attribute. For the training of the model, the first set contains only the data with the six attributes which were initially introduced, and a class attribute based on yield data. The class attribute has 3 categorizations: low, moderate, and high. Any yield less 600 metric tons is categorized as Low yield, any yield less 700 metric tons is categorized moderate yield and yield greater or equal 800 metric ton is categorized high yield. The data was divided into 75% for training and 25% for testing.

2.4 Model development environment

Weka (Waikato Environment for Knowledge Analysis) was used to develop the model. This tool is a data mining suite with several machine learning algorithms for carrying out classification, clustering, association task. The software also has some data preprocessing tools and visualization tools and other functionality for data transformation into proper form for mining.

2.5 Prediction Algorithms

A total of 5 predictor were used in this work: Logistic Regression (LR), Artificial Neural Network (ANN), Random Forest (RF), Random Trees (RT) and Naïve Bayes (NB). These classifiers have their strength and weakness. LR is an algorithm that can be used for classifying an object into various label groups and can used as a regression technique [1]. It predicts the probability of an event taking place using a logistic function that whose value range between 0 and 1. If the probability is higher than certain threshold, then the object belongs to the class, else it is not. ANN is a computing algorithm with large number of interconnecting artificial neurons. Neural network works analogously like human brain [37]–[39]. It consists of computational nodes which receive input and a processing layer that sums up the input to produce the output. There are several architectures of neural network, however, all have basic 3 layers. it maps input to output to find patterns in the training data, with this, it generalizes training set of the input value already classified in predefined class. RF is a form of learning algorithm that generates a tress based on the attributes from the dataset, where each tree is itself a classification tree [40]. Several random samples are generated from which randomized trees are developed. It initially randomly samples the complete data set, following which many decision trees are generated. Each tree is trained using a random sample from which it was built. All of the decision trees' predictions are then combined into a single tree for a single output. If multiple trees are trained and a greater number of them predict that an object belongs to class Y, and one says no, the final random forest prediction will be class Y. RT Is a tree is formulated from a random sample of other trees, with each trees possessing n

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

number of random features at each node [41]. The general theory is that each tree has probability of being sampled. NB is an algorithm that uses a function $g(x)$ to performs mapping of certain input x into output class, where the class have label 1……n. As a classifier, NB forecast the predict the probability that an item belongs to a certain class [42]. It uses the bayes theorem based on the following equation. the probability of Y given X can be expressed.

$$P(Y|X) = \frac{P(X|Y)\,P(Y)}{P(X)}. \tag{1}$$

2.6 Accuracy metric

A good number of metrics methods exist to measure the accuracy of machine learning model. The first metric used in this work measure the accuracy of the models at the class level. True Positive (TP), False Positive (FP), precision, Recall, F-Measure and Receiver Operating Characteristic (ROC) area. ROC area measures the usefulness of a model. The second is the metric that measure the error rate of the model. They include RMSE, RRSE and MAE. Mathematical equations of the matrices are given in the following equations 2 to 6.

$$Precision = \frac{True_{positive}}{True_{positive} + False_{positive}} \tag{2}$$

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}} \tag{3}$$

$$Recall = \frac{True_{positive}}{(True_{positive} + False_{negative})} \tag{4}$$

$$MAE = \frac{1}{n}\sum_{j=1}^{n} |y_j - \hat{y}_j| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{6}$$

3. RESULT AND DISCUSSION

In this section, we present the outcomes obtained from several rice yield prediction models that we have developed. The results are depicted in Figure 2, illustrating the correlations among the different variables employed in the models.

Journal of Information Systems and Informatics
Vol. 5, No. 2, June 2023

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

3.1 Prediction Model

Predictive modeling is a powerful mathematical approach employed to forecast future events or outcomes through the analysis of patterns and trends within a given dataset. It involves the application of various algorithms and techniques to identify relationships and dependencies among the input variables, ultimately enabling accurate predictions to be made based on these patterns. By leveraging historical data and statistical methods, predictive modeling provides valuable insights and predictions that can aid decision-making processes. The following is an overview of the steps involved in constructing the Rice Yield Forecasting model.

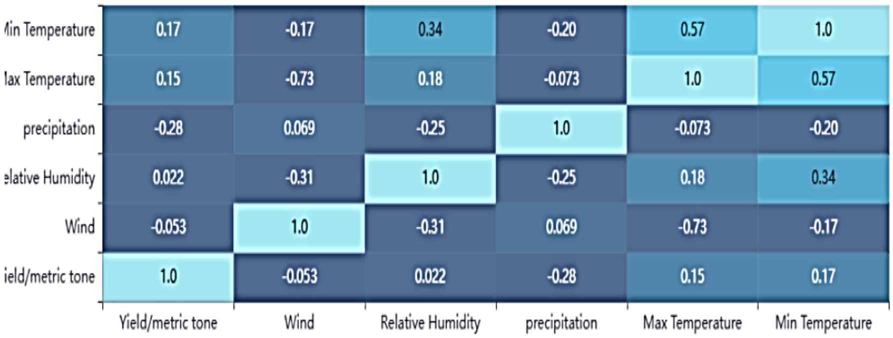

Figure 2. Relationship among the attributes of rice yield model

Figure 2 illustrates the relationships between rice yield and various climatic variables. The strongest correlation observed is between rice yield and maximum temperature, with a coefficient of 0.17. However, it is important to note that this correlation is still considered weak. Interestingly, our findings align with other studies, such as [4], which also reported no relationship between rice yield and rainfall in southeast Nigeria. Similarly, our results indicate a negative correlation between rice yield and precipitation. Although this may seem counterintuitive, similar findings have been documented by [19] in China. Among all the variables considered, temperature and humidity exhibit a positive relationship with rice yield, while wind and precipitation display a negative correlation. Moving on, Figure 3 and Figure 4 depict graphical representations of two of the predictive rice yield models developed in this study.

Figure 3 provides a visual representation of the rice yield prediction models based on Artificial Neural Networks (ANN). The model architecture includes input neurons representing wind, relative humidity, precipitation, maximum temperature, minimum temperature, and rice yield. The ANN consists of one hidden layer for processing the input data, while the output layer categorizes the yield into classes such as low, moderate, or high. On the other hand, Figure 4

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

depicts a graphical representation of the Random Tree rice yield prediction models. The tree structure consists of branches representing wind, relative humidity, precipitation, maximum temperature, and minimum temperature as the input variables. The leaves of the tree represent the output class, which corresponds to different yield categories. These graphical representations provide a clear overview of the model structures and the relationships between the input variables and the predicted rice yield.

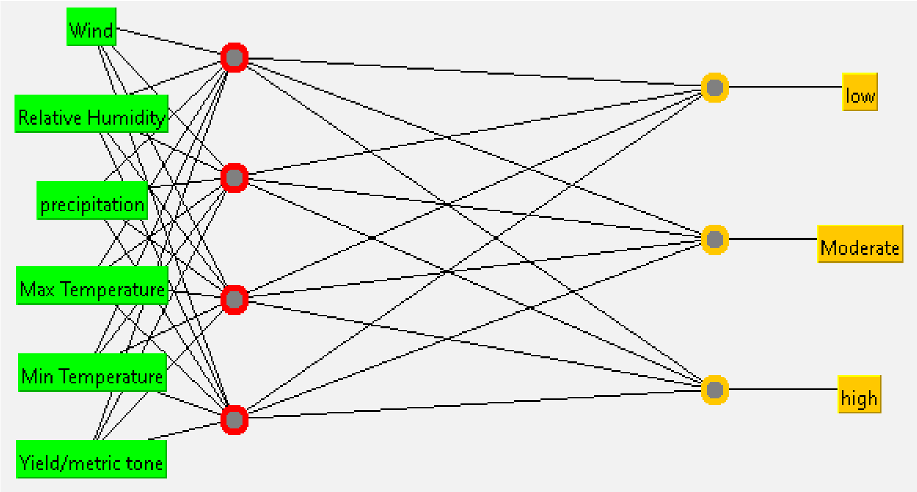

Figure 3. ANN rice yield model

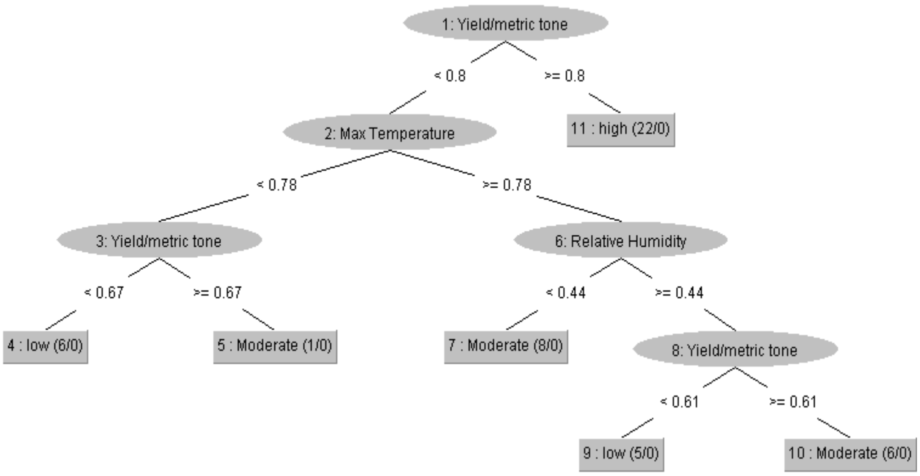

Figure 4. Random tree rice yield model

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

3.2 Model Accuracy Measure

This paper presents the development of five models. A comprehensive overview of these models, along with their respective accuracy measures, can be found in Tables 2 and 3. Table 2 provides class-based accuracy measures for each model, while Table 3 presents the error measures for each model.

Table 2. Class-based accuracy measure

	TP	FP	Precision	Recall	F-measure	ROC Area
ANN	0.750	0.250	0.83	0.75	0.72	0.89
LR	0.750	0.250	0.83	0.75	0.72	0.95
RF	1	0	1	1	1	1
RT	1	0	1	1	1	1
NB	0.917	0.083	0.929	0.917	0.915	1

Class-based metrics provide insights into how accurately the models predict the actual class instances of the data. The evaluation of each model included measurements such as true positive (TP), false positive (FT), precision, recall, F-measure, and ROC Area. Among the models, Random Forest (RF) and Random Trees (RT) demonstrated superior performance with a TP rate of 1, followed by Naïve Bayes (NB) with a rate of 0.19, while Neural Network (ANN) and Logistic Regression (LR) both achieved a rate of 0.75. Considering other metrics like precision and recall, which are functions of TP and FT, RF and RT still outperformed the rest with a value of 1. Referring to Table 2, RF, RT, and NB all attained a ROC Area value of 1, while ANN scored 0.89 and LR achieved 0.95. This indicates that RF, RT, and NB exhibit a higher capability to correctly predict the class of rice yield compared to ANN and LR.

Interpreting the results, ANN and LR were only able to predict the actual rice yield with 75% accuracy, while NB achieved a prediction accuracy of 91.7%. Both tree algorithms, RT and RF, demonstrated perfect accuracy of 100% in predicting rice yield. Based on these metrics, the tree algorithms (RF and RT) are considered the best models for rice prediction using climatic variables. The error-based metrics can be found in Table 3, providing further insights into the models' performance.

Table 3. The error measure

	MAE	RMSE	RAE	RRSE	Accuracy
ANN	0.16	0.35	39.63%	76.37%	75
LR	0.15	0.3	35.7%	78.2%	75
RF	0.14	0.17	33.02%	38.9%	100
RT	0	0	0	0	100
NB	0.082	0.21	19.22%	46.2%	91.6

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Error-based metrics provide insights into how accurately the models predict the actual numerical values of each data instance. The evaluation of each model included measurements such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Relative Absolute Error (RAE), Relative Root Squared Error (RRSE), and Accuracy. Examining Table 3, Random Trees (RT) demonstrated superior performance with MAE, RMSE, RAE, and RRSE values of 0. On the other hand, Random Forest (RF) achieved an MAE of 0.14 and an RMSE of 0.17. Naïve Bayes (NB) ranked third in performance with an MAE of 0.08 and an RMSE of 0.21, while Neural Network (ANN) and Logistic Regression (LR) both exhibited an MAE of 0.15 and 0.16, and an RMSE of 0.3 and 0.35 respectively.

Interpreting the error measures, ANN and LR were only able to predict the actual rice yield with 75% accuracy, while NB achieved a prediction accuracy of 91.6%. Both tree algorithms, RT and RF, demonstrated perfect accuracy of 100% in predicting rice yield. Based on these metrics, the tree algorithms (RF and RT) are considered the best models for rice prediction, with RT being superior to RF due to its lower error. Considering both Tables 2 and 3, RT exhibits better performance in predicting rice yield in Katsina state compared to other models, while ANN performed less effectively than the other models.

3.3 Discussion

This study examines the performance of five distinct models for predicting rice yield based on climatic variables. The models were evaluated using both class-based and error-based metrics to assess their accuracy in predicting class instances and numerical values, respectively. The class-based metrics revealed that Random Forest (RF) and Random Trees (RT) surpassed the other models by achieving a TP rate of 1, indicating their precise predictions of rice yield classes. Naïve Bayes (NB) followed with a TP rate of 0.19, while Neural Network (ANN) and Logistic Regression (LR) attained a rate of 0.75. Notably, RF, RT, and NB showcased excellent precision and recall, surpassing ANN and LR in these measures.

Furthermore, the ROC Area, a metric gauging model usefulness, awarded RF, RT, and NB perfect scores of 1, while ANN scored 0.89 and LR achieved 0.95. This suggests that RF, RT, and NB exhibited superior ability in accurately predicting rice yield classes compared to ANN and LR. Analyzing the error-based metrics, which evaluate the models' performance in predicting numerical values, RT demonstrated exceptional results. It achieved the lowest MAE, RMSE, RAE, and RRSE values of 0, indicating its accurate estimation of the true numerical values of rice yield. RF performed well with an MAE of 0.14 and an RMSE of 0.17, slightly higher than RT. NB showcased satisfactory performance with an MAE of 0.08 and an RMSE of 0.21. Conversely, ANN and LR exhibited higher errors, with an MAE of 0.15 and 0.16, and an RMSE of 0.3 and 0.35, respectively.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Interpreting the error measures, both RT and RF achieved perfect accuracy of 100% in predicting rice yield, while ANN and LR reached 75% accuracy. NB demonstrated higher accuracy at 91.6%. Overall, the tree algorithms, specifically RF and RT, displayed exceptional performance in predicting rice yield using climatic variables, with RT being slightly superior due to its lower error.

These findings strongly suggest that the tree algorithms, particularly RT, are the most effective models for accurately predicting rice yield based on climatic variables. They outperformed the other models in both class-based and error-based metrics. This study underscores the significance of utilizing diverse evaluation metrics to comprehensively assess the predictive performance of models.

4. CONCLUSION

This paper developed rice yield prediction model for Katsina state in Nigeria, using climatic data and rice yield data. The work used 5 machine learning algorithms: Logistic Regression (LR), Artificial Neural Network (ANN), Random Forest (RF), Random Trees (RT) and Naïve Bayes (NB). The result of the data analysis reveals that precipitations has no significant relationship with rice output but rather temperature has closer relationship (figure 2). Also, of all the 5 models, Random Trees has the highest accuracy in predicting rice yield with MAE, RMSE, RAE and ERSE all equal to 0 and 100% accuracy while Neural Network performed more poorly than all other models. We therefore recommend that as future katsina state climatic variables are predicted by various agencies within and outside Nigeria, our model can be used alongside to predict rice yield in the state. This will help the government to plan and will also ensure food security for the people of the state.

REFERENCES

[1] L. S. Cedric, W.Y.H. Adoni, R. Aworka, J.T. Zoueu, F.K. Mutombo, M. Krichen, and C.L.M. Kimpolo, “Crops yield prediction based on machine learning models: Case of West African countries,” *Smart Agric. Technol.*, vol. 2, no. December, p. 100049, 2022, doi: 10.1016/j.atech.2022.100049.

[2] N. Ebele and N. Emodi, “Climate Change and Its Impact in Nigerian Economy,” *J. Sci. Res. Reports*, vol. 10, no. 6, pp. 1–13, 2016, doi: 10.9734/jsrr/2016/25162.

[3] National Bureau of Statistics (NBS), *Annual Abstract of Statistics*. Lagos, Nigeria: National Bureau of Statistics, 2010.

[4] P. G. Oguntunde, G. Lischeid, and O. Dietrich, “Relationship between rice yield and climate variables in southwest Nigeria using multiple linear regression and support vector machine analysis,” *Int. J. Biometeorol.*, vol. 62, no. 3, pp. 459–469, 2018, doi: 10.1007/s00484-017-1454-6.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

[5] S. A. Tiamiyu, J. N. Eze, T. M. Yusuf, A. T. Maji, and S. O. Bakare, "Rainfall Variability and its Effect on Yield of Rice in Nigeria," *Int. Lett. Nat. Sci.*, vol. 49, no. November, pp. 63–68, 2015, doi: 10.18052/www.scipress.com/ilns.49.63.
[6] A. Oluwaseyi, D. Nehemmiah, and S. Zuluqurineen, "Genetic Improvement of Rice in Nigeria for Enhanced Yeild and Grain Quality - A Review," *Asian Res. J. Agric.*, vol. 1, no. 3, pp. 1–18, 2016, doi: 10.9734/arja/2016/28675.
[7] H. O. Edeh and E. C. Eboh, "Analysis of Environmental Risk Factors Affecting Rice Farming in Ebonyi State , Southeastern Nigeria," vol. 7, no. 1, pp. 100–103, 2011.
[8] C. O. Akinbile, O. O. Ogunmola, A. T. Abolude, and S. O. Akande, "Trends and spatial analysis of temperature and rainfall patterns on rice yields in Nigeria," *Atmos. Sci. Lett.*, vol. 21, no. 3, pp. 1–13, 2020, doi: 10.1002/asl.944.
[9] E. A. Jiya, U. Iliyasu, and D. U. Ebem, "Agricultural Research and Food Security under Climate Change : The Place of Machine Learning Models," *J. Adv. Math. Comput. Sci.*, vol. 11, no. 1, pp. 11–20, 2023, doi: dx.doi.org/10.22624/AIMS/MATHS/V11N1P2.
[10] T. Ebenezer, "Drought , desertification and the Nigerian environment : A review," *J. Ecol. Nat. Environ. Rev.*, vol. 7, no. 7, pp. 197–209, 2015, doi: 10.5897/JENE2015.
[11] C. O. Egbunu, M. T. Ogedengbe, T. S. Yange, and T. Gbaden, "Towards Food Security : the Prediction of Climatic Factors in Nigeria using Random Forest Approach," vol. 7, no. December, pp. 70–80, 2021, doi: 10.35134/jcsitech.v7i4.15.
[12] E. I. Papageorgiou, A. T. Markinos, and T. A. Gemtos, "Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application," *Appl. Soft Comput. J.*, vol. 11, no. 4, pp. 3643–3657, 2011, doi: 10.1016/j.asoc.2011.01.036.
[13] J. J. L. Westerveld, M. J. C. van den Homberg, G. G. Nobre, D. L. J. van den Berg, A. D. Teklesadik, and S. M. Stuit, "Forecasting transitions in the state of food security with machine learning using transferable features," *Sci. Total Environ.*, vol. 786, no. May, pp. 147366, 2021, doi: 10.1016/j.scitotenv.2021.147366.
[14] S.H. Qader, C.E. Utazi, R. Priyatikanto, P. Najmaddin, E.O. Hama-Ali, N.R. Khwarahm, A.J. Tatem, and J. Dash, "Exploring the use of Sentinel-2 datasets and environmental variables to model wheat crop yield in smallholder arid and semi-arid farming systems," *Sci. Total Environ.*, vol. 869, pp. 161716, 2023, doi: 10.1016/j.scitotenv.2023.161716.
[15] E. I. Papageorgiou, K. D. Aggelopoulou, T. A. Gemtos, and G. D. Nanos, "Yield prediction in apples using Fuzzy Cognitive Map learning approach,"

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Comput. Electron. Agric., vol. 91, pp. 19–29, 2013, doi: 10.1016/j.compag.2012.11.008.

[16] V. Mamatha and J. C. Kavitha, "Machine learning based crop growth management in greenhouse environment using hydroponics farming techniques," *Meas. Sensors*, vol. 25, no. December 2022, pp. 100665, 2023, doi: 10.1016/j.measen.2023.100665.

[17] J. Manrique-Silupu, J. C. Campos, E. Paiva, and W. Ipanaqué, "Thrips incidence prediction in organic banana crop with Machine learning," *Heliyon*, vol. 7, no. 12, p. e08575, 2021, doi: 10.1016/j.heliyon.2021.e08575.

[18] A. Tripathi, R. K. Tiwari, and S. P. Tiwari, "A deep learning multi-layer perceptron and remote sensing approach for soil health based crop yield estimation," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 113, no. August, p. 102959, 2022, doi: 10.1016/j.jag.2022.102959.

[19] J. Zhi, X. Cao, Z. Zhang, T. Qin, L. Qi, L. Ge,A. Guo, X. Wang, C. Da, Y. Sun, and W. Liu, "Identifying the determinants of crop yields in China since 1952 and its policy implications," *Agric. For. Meteorol.*, vol. 327, no. September, 2022, doi: 10.1016/j.agrformet.2022.109216.

[20] G. Martini *et al.*, "Machine learning can guide food security efforts when primary data are not available," *Nat. Food*, vol. 3, no. 9, pp. 716–728, 2022, doi: 10.1038/s43016-022-00587-8.

[21] E. Wollenberg, S. J. Vermeulen, E. Girvetz, A. M. Loboguerrero, and J. Ramirez-Villegas, "Reducing risks to food security from climate change," *Glob. Food Sec.*, vol. 11, pp. 34–43, 2016, doi: 10.1016/j.gfs.2016.06.002.

[22] J. Schmitt, F. Offermann, M. Söder, C. Frühauf, and R. Finger, "Extreme weather events cause significant crop yield losses at the farm level in German agriculture," *Food Policy*, vol. 112, no. August, 2022, doi: 10.1016/j.foodpol.2022.102359.

[23] M. Singh Boori, K. Choudhary, R. Paringer, and A. Kupriyanov, "Machine learning for yield prediction in Fergana valley, Central Asia," *J. Saudi Soc. Agric. Sci.*, no. xxxx, 2022, doi: 10.1016/j.jssas.2022.07.006.

[24] P. A. J. van Oort, B. G. H. Timmermans, R. L. M. Schils, and N. van Eekeren, "Recent weather extremes and their impact on crop yields of the Netherlands," *Eur. J. Agron.*, vol. 142, no. October 2022, p. 126662, 2023, doi: 10.1016/j.eja.2022.126662.

[25] R. Aworka *et al.*, "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries," *Smart Agric. Technol.*, vol. 2, no. December 2021, p. 100048, 2022, doi: 10.1016/j.atech.2022.100048.

[26] G. Shao, W. Han, H. Zhang, L. Zhang, Y. Wang, and Y. Zhang, "Prediction of maize crop coefficient from UAV multisensor remote sensing using machine learning methods," *Agric. Water Manag.*, vol. 276, no. November 2022, p. 108064, 2023, doi: 10.1016/j.agwat.2022.108064.

[27] E. I. Papageorgiou, A. T. Markinos, and T. A. Gemtos, "Soft computing technique of fuzzy cognitive maps to connect yield defining parameters

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

with yield in cotton crop production in central Greece as a basis for a decision support system for precision agriculture application," *Stud. Fuzziness Soft Comput.*, vol. 247, pp. 325–362, 2010, doi: 10.1007/978-3-642-03220-2_14.
[28] A. M. Shuaibu, M. N. Muhammad, and Y. Abu-safyan, "Forecasting Rice Production in Jigawa State , Nigeria using Fuzzy Inference System," *Dutse J. Pure Appl. Sci.*, vol. 7, no. 4, pp. 203–213, 2021.
[29] L. Wickramasinghe, R. Weliwatta, P. Ekanayake, and J. Jayasinghe, "Modeling the Relationship between Rice Yield and Climate Variables Using Statistical and Machine Learning Techniques," *J. Math.*, vol. 2021, 2021, doi: 10.1155/2021/6646126.
[30] D. Chakraborty *et al.*, "Usability of the Weather Forecast for Tackling Climatic Variability and Its Effect on Maize Crop Yield in Northeastern Hill Region of India," *Agronomy*, vol. 12, pp. 18-, 2022.
[31] S. Ramesh, A. Hebbar, V. Yadav, T. Gunta, and A. Balachandra, "CYPUR-NN: Crop Yield Prediction Using Regression and Neural Networks," in *Emerging Research in Computing, Information, Communication and Applications, ERCICA 2020.*, Singapore.: Springer, 2022, pp. 193–206.
[32] D. Paudel *et al.*, "Machine Learning for large-scale crop yield forecasting," *Agric. Syst.*, vol. 187, 2921, doi: 10.1016/j.agsy.2020.103016.
[33] A. Prasad, N., Patel, N. R. and Danodia, "Crop yield prediction in cotton for regional level using random forest approach," *Spat. Inf. Res*, vol. 29, no. 2, pp. 195–206, 2021, doi: doi: 10.1007/s41324-020-00346-6.
[34] S. Agarwal and S. Tarar, "A hybrid approach for crop yield prediction using machine learning and deep learning algorithms," *J. Phys. Conf. Ser.*, vol. 1714, no. 1, 2021, doi: doi: 10.1088/1742-6596/1714/1/012012.
[35] C. Rugimbana, "Predicting Maize (Zea Mays) Yields in Eastern Province of Rwanda Using Aquacrop Model," University of Nairob, 2019.
[36] L. Khaki, S. and Wang, "Crop yield prediction using deep neural networks," *Front. Plant Sci.*, vol. 10, 2019, doi: doi: 10.3389/fpls.2019.00621
[37] S. A. Alanazi *et al.*, "Boosting Breast Cancer Detection Using Convolutional Neural Network," *J. Healthc. Eng.*, vol. 2, 2021.
[38] L. He, H. Li, S. K. Holland, W. Yuan, M. Altaye, and N. A. Parikh, "NeuroImage : Clinical Early prediction of cognitive de fi cits in very preterm infants using functional connectome data in an arti fi cial neural network framework," *NeuroImage Clin.*, vol. 18, no. October 2017, pp. 290–297, 2018, doi: 10.1016/j.nicl.2018.01.032.
[39] E. Okewu, S. Misra, L. F. Sanz, F. Ayeni, V. Mbarika, and R. Damaševičius, "Deep Neural Networks for Curbing Climate Change-Induced Farmers-Herdsmen Clashes in a Sustainable Social Inclusion Initiative," *Probl. Sustain. Dev.*, vol. 14, no. 2, pp. 143–155, 2019.
[40] F. Guo *et al.*, "What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests," *Int. J. Wildl. Fire*, vol. 25, no. 5,

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

pp. 505–519, 2016, doi: 10.1071/WF15121.

[41] S. Gupta, S. K. Abraham, V. Sugumaran, and M. Amarnath, "Fault Diagnostics of a Gearbox via Acoustic Signal using Wavelet Features, J48 Decision Tree and Random Tree Classifier," *Indian J. Sci. Technol.*, vol. 9, no. 33, pp. 1–8, 2016, doi: 10.17485/ijst/2016/v9i33/101328.

[42] N. Salmi and Z. Rustam, "Naïve Bayes Classifier Models for Predicting the Colon Cancer," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 546, 2019, doi: 10.1088/1757-899X/546/5/052068.