Deep Learning Based Neural Network Modelling for Cassava Yield Prediction

[1,2]Samuel S. Udoh*, [1,2]Udoinyang G. Inyang, [1,2]Uduak A. Umoh,
[1,2] Uduak D. George, [3]Unyime R. Etuk,[4]Theresa B. Akpan, [4]Etekamba M. Bassey

[1]*Department of Computer Science, Faculty of Science, University of Uyo, Nigeria*
[2]*Tetfund Centre of Excellence in Computational Intelligence Research, University of Uyo, Nigeria*
[3]*Department of Agricultural Economics and Extension, Faculty of Agriculture, University of Uyo, Nigeria*
[4] *Department of Computer Science, School of Science, National Open University of Nigeria, Nigeria*

**Corresponding author: email: samueludoh@uniuyo.edu.ng*

Abstract

Predicting the yield of cassava is necessary to guide decisions towards its availability for the ever-growing number of people depending on it for food. The dynamism and complexity of cassava yield (CY) prediction make it difficult for linear models to produce accurate predictions. Many CY prediction models are based on linear relationships with attendant inability to extract hidden interactions existing among CY features thereby providing insufficient, inefficient and inaccurate information for CY prediction. In consideration of the need to properly analyze the nature, trend, impact and parameter combination for optimal CY prediction, this work proposes a Deep Learning (DL) CY prediction model with capabilities of deciphering hidden and non-linear relationships among CY parameters. A 5-layer (14-6-4-3-1) DL neural network model was designed. The cassava dataset with 2500 samples was collected and used for cassava yield model building and estimation. Hyperbolic transfer function was deployed in the DL hidden layers while Sigmoid transfer function was used in the output layer to produce the least average training and testing errors of 0.0024 and 0.037 respectively. Investigations regarding CY based on the number of cassava cultivars planted per stand revealed that the cultivation of one stem per stand had a higher contributory effect on average CY than cultivation of more stems per stand. The DL model earned 93.60% and 95.00% for accuracy and precision metrics respectively, indicating improved performance. As further work, other variants of DL would be investigated and compared with the proposed DL model with a view to improving CY prediction.

Keyword: Cassava yield prediction, Deep learning, Data mining, Neural network, Agriculture

1. Introduction

A significant percentage of global cassava production is contributed by Nigeria [4]. The crop is mostly cultivated in the southern and middle belt regions of the country [43]. It is the crop of choice for small and large-scale farming and is considered a staple food for more than a billion people in 105 countries of the world [9]. Cassava is an important source of vitamins and micronutrients [21]. Cassava leaves are also used as protein supplement for livestock [23]. Although Nigeria produces a significant percentage of cassava in the world, the production potential has not been harnessed for significant economic growth in the country [13]. Timely and accurate prediction of crop yield is of great importance in strengthening global food security. It plays a major role in determining the strategic plans for agricultural products' import-export policies.

Although crop yield prediction (CYP) is a complex and challenging task, machine learning and data mining approaches have recently shown predictive prominence. Alison *et al.* in [3] presented cassava root yield prediction using the linear regression modelling approach. Prediction accuracy was only based on correlation coefficient. The model lacked the ability to learn cassava yield patterns from previous data. Ikuemonisan & Akinbola in [43] deployed ARMA/ARIMA tools to forecast cassava production using indicators in Nigeria. The forecasts were made from ARIMA (5,1,0), ARMA (1,1) and ARIMA (1,1,3) to fit production series, yield series and harvested series respectively. Findings showed that 84% of total cassava output was projected to be available for consumption and that 29% of the 84% was lost during post-harvest activities. The accuracy of the model was not evaluated in addition to non-inclusion of learning modules to guide intelligent decisions. Sadenovaa *et al.* in [44] used mathematical model to predict crop yield. Dynamic characteristics of predictors were studied to predict productivity with correlation coefficient (CC) as the metric of choice. The CC between the calculated yield values and the official statistics was 0.84.

DL evolves from artificial neural network (ANN) for the simulation of the mechanism of learning in biological organisms [1], [39]. DL is more attractive than conventional learning methods primarily when sufficient data and computational power is available. Recent years have seen an increase in data availability and computational intelligence, which have led to investigations of many hidden layers in DL technology. DL has strengths especially learning from previous data, memorizing salient features, recalling data patterns and generalizing into future outcomes using unknown examples [39]. DL algorithms have surpassed human performance in some tasks such as image processing [5], [16], pattern classification [40] and exponential computation [41]. This work aimed at adopting DL methodology for CYP with capabilities to decipher the hidden and non-linear relationships among cassava yield parameters to activate accurate prediction and decision processes. CYP algorithm would be formulated to guide the sequence of operations in the prediction process. This work would provide a guide to agronomic strategic planning towards increasing cum predicting the yield of cassava in the agricultural sector and provision of food as well as income for economic sustenance. The remainder of the work is organized as follows: Section 2 focusses on review of related works carried out in cassava production and DL technologies. Methodological framework of DL model for CYP is presented in Section 3. In Section 4, CYP model implementation and results are presented, while Section 5 deduces the conclusion and recommendations for future research.

2. Related works

Research on improved agronomic practices in cassava production, such as identifying optimum planting time, tillage operation, preparation of planting material, weed control, intercropping and soil erosion control has resulted in better understanding of opportunities to improve crop production and yield [17]. Some researchers had worked on development of improved varieties for crop production. Chetty *et al*. in [9] established a robust transformation platform for the production of transgenic industry-preferred cassava using biotechnology. Ceballos *et al*. in [8] deployed conventional breeding, marker-assisted selection, genomic selection and inbreeding to study the methods of improving cassava varieties. Knowledge and tools for genetic composition to develop resistant varieties and improve CY are reported in [22], [13].

Reinhard in [31] experimented on agronomic practices for sustainable production of cassava. The work showed that cassava yield is negatively affected when either rainfall or temperature are low in the first 3-5 months after planting. It was shown that planting in ridges is better in the rainy season whereas planting on flat surface is better in the dry season. Consideration of varieties of cassava cultivars, soil types, variation in soil properties and differences in atmospheric conditions could be incorporated to substantiate the experiment. Odubanjo *et al*. in [26] investigated the effect of different amounts of supplemental drip irrigation on the tuber yield of cassava and showed that the soil with 100% drip irrigation produced 600% cassava root yield compared to the soil with no irrigation.

A spatial model to assess the suitability of land for supporting sustainable cassava production was reported in [30]. It was shown that cassava grow very well in the tropics, between 30°N and 30°S in areas where annual rainfall is greater than 500 mm and mean temperature is greater than 20 °C. Ezedinma *et al*. in [11] presented the opportunities and challenges for production of cassava in Nigeria. Lack of exploitation of the potentials of cassava production in contributing to economic growth motivated the work. Amanda *et al*. in [4] examined physiological processes underlying productivity in cassava and provided some strategies for yield improvement through genetic alterations. The study revealed that although informed use of fertilizer could lead to increase in yield, the genetic yield potential of cassava sets the ceiling on what may be produced at a given location.

Tunrayo *et al*. in [37] developed site-specific recommendations for cassava production in Nigeria and Tanzania. Geospatial information obtained were related to climate, soil and remote sensing data. Spatial multivariate analysis was used to delineate agricultural extension partners' operational area into homogeneous clusters. Multivariate cluster analysis provided unbiased guide for site selection for technological innovations. This approach ensured representativeness and maximized unbiasedness while at the same time maximizing operational efficiency [37]. Deep learning algorithms are integral part of machine learning method. It incorporates many hidden layers to improve the learning process

[10]. Deep learning algorithm learns decision boundaries from non-linear data and automates the process of extracting features that improve model performance [6], [12], [18], [24], [34]. [35], [37].

Many researchers have deployed various methods in the task of CYP. Oni & Akanle in [25] predicted cassava production (CP) using moving average, trend analysis and smoothing models. The work compared exponential smoothing models and inferred that CP in Nigeria is not affected by season. Time forecasting was very practical and accurate compared with other models. The study showed that the simple exponential smoothing model is better in predicting the yield of cassava compared to Winter's additive, Winter's multiplicative and Holts' trend models. Alison *et al.* in [3] presented cassava root yield prediction in different water regimes. An agronomic and physiological data of final root yield obtained under two water regimes were tested using four prediction models: linear regression with backward selection, linear regression with stepwise selection, partial least square and Bayesian ridge regression. There were differences in the predictive ability of the models due to early traits of the crop regardless of the water condition. Thomas *et al.* in [35] investigated the best time for cassava planting to achieve maximum yield. A growth simulation model for cassava called SIMulation of CASsava (SIMCAS) was trained to predict cassava growth yield. It was observed that the predicted and observed values were reasonably close. SIMCAS was considered a good model that could predict correct planting time and potential yield of cassava at a given location. Hajir in [15] compared the yield of crops using climatic factors such as: sunlight, humidity, temperature and rainfall. The data was pre-processed from its raw format to a numerical form and was split into training and testing datasets. A regression model was deployed to determine the relationship between the input variables and CY.

Al *et al.* in [2] presented breeding and agronomic research on cassava production. The breeding program successfully realized high-yielding cultivars with significant economic benefits. Building resistance to invasive pests and diseases have become a top priority due to the emergent threat of cassava mosaic disease [2]. Further exploration in data-driven agriculture is needed to guide researchers and farmers towards sustainable navigation in innovative technology. It was reported in [14], [7] that cassava yield depends on various factors such as water, soil type, soil nutrients, climate and the environmental factors. However, experiments could be carried out to substantiate the relationship and the nature of correlation between CY and the aforementioned variables. Kintché *et al.* in [20] studied CY loss in two provinces of the Democratic Republic of Congo. Boundary line approach was used to investigate the CY loss. Forty-two cassava farms in Kongo central and thirty-seven farms in Tshopo were analyzed to find out how soil fertility, farmers' cultivation practices as well as pest and disease infestation affected CY. Logistic regression modelling revealed that pests and diseases played a sparse role in the yield losses. Low soil fertility led to decrease in cassava yield in many farms.

Patricia *et al.* in [27] reviewed the growth and yield of cassava crops. The goal was to study Eighteen published cassava models and gain more insights on their capability to simulate storage root biomass and to categorize them into dynamic and static models. Fourteen models were dynamic while four models were static. The dynamic models simulated the growth process and provided the yield estimates but lacked ability to predict maturity date and were less-accurate in simulating the detailed development of nodal units and determining the final yield. The four static models were useful in estimating cassava yield. However, the models were evaluated using a limited range of predictors thereby hindering comprehensive assessment of non-linear relationships between input variables and CY.

Petteri *et al.* in [28] deployed DL based on convolutional neural networks (CNNs) for CY prediction. The CNN methodology tested the selection of the training algorithm, network depth, hyper parameters on the regularization strategy and the prediction efficiency. Saeed *et al.* in [31] hybridized CNNs and recurrent neural networks (RNNs) for CY prediction. CNN's processed multiple data arrays, while RNN captured sequential data based on time dependencies. The model was compared with fully connected neural networks (FNN) and random forest (RF). It was reported that CNN-RNN model outperformed other tested models in CY prediction. It captured time dependent environmental factors, generalized the yield prediction to environments that were not part of the test and revealed the extent to which weather conditions could affect CY. However, no precise algorithm was presented for CY prediction.

Summary of limitations of the works reported in [20], [27], [14], [17], [31] are lack of concise algorithms to guide other researchers on the sequence of CY prediction operations. Deployment of few prediction variables which hinder proper investigation of non-linear interaction. Deployment of small dataset and few evaluation metrics which impede overall test of accuracy, efficiency and reliability of the model. The current research incorporates a concise CY prediction algorithm which catalogs the sequence of prediction operations. Large dataset with fourteen prediction variables are incorporated to investigate hidden patterns and non-linear interaction in the prediction process. Fifteen decision metrics are deployed to evaluate and decipher the accuracy, efficiency and reliability of the model.

3.0 Materials and methods

The methodological workflow for the prediction of CY comprises four main stages namely: Agronomic dataset, pre-processing, DL modelling and model evaluation as shown in Fig. 1. The agronomic dataset holds agro-climatic variables, location-based factors, farmers' parameters and other Cassava production-related data. Pre-processing task includes feature scaling, filling of missing values, encoding of categorical variables as well as identification and extraction of relevant features for CYP. At this stage, the extraction of variables that influence CY is performed and the dataset is split into three groups for training, testing and validation operations. DL modelling is performed using NeuroSolutions 7.0 software tools and the result is evaluated using confusion matrix parameters and its derivatives such as accuracy, specificity, recall and precision. The model with least training and testing errors would be used for prediction of CY and extraction of relationships, trends and other relevant information for CYP. Details of the methods involved in the modelling are described in the following sections.

3.1 Cassava yield dataset characterization

Data totaling 2500 samples were collected over a period of one year (March, 2021 to March 2022) from Akwa Ibom North-West Agricultural zone, southern Nigeria. The study area covered forty (40) cassava farms, comprising four (4) farms from each of the ten (10) Local Government Area making up the Agricultural zone. Twenty (20) of the forty (40) farms were planted on flat surface while the other 20 were cultivated on ridges. Each of the 40 farms was divided into five (5) plots. Each of the plots was used in cultivating a particular variety of cassava. The cassava stem was cut 30cm long and planted with 1m x 1m intra and inter-row spacing. There were six (6) rows with ten (10) plants per row, making it sixty (60) plants per plot. Ikot Ekpene, the main study area is located at Latitude 5.183^{o}N and Longitude 7.715^{o}E with Elevation of 75.68m. The work in [42] showed that the average annual rainfall in the main study area is 2007.49mm with concentration in the rainfall occurring between April and October. The study area has an average day temperature of 32.12 ^{o}C and average night temperature of 23.67^{o}C. It has a gentle undulating land with diversity in soil nutrients for cassava production.

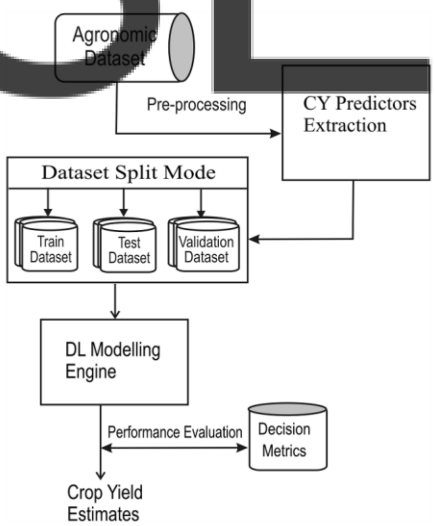


Fig. 1. Methodological workflow for CYP

Table 1
Description of Cassava yield input/output variables

SN	Variable	Code	Data Type / Size	Remarks
1	Rainfall/ Irrigation	RFI	FLOAT (6)	RFI was measured in millimetres per day. Average monthly RFI in the study area was computed as 2007.49 mm/day.
2	Temperature	TMP	FLOAT (4)	TMP was measured in degree Celsius. Average monthly day and night TMP in the study area was evaluated as 27.90 ^{0}C.
3	Number of Stems Cultivated	NSC	INTEGER (1)	NSC value of 1 or 2 or 3 was assigned as input to the Deep Learning Model based on the number of cassava stems (cultivars) planted per stand.
4	Fertilizer Application	FZA	BINARY (1)	FZA value of 1 was assigned if fertilizer was applied and 0 otherwise.
5	Manure Application	MNA	INTEGER (1)	MNA had the values such as: Farmyard MNA(3), Compost MNA(2) and Green MNA(1)
6	Number of Stands at Harvest	NSH	INTEGER (4)	Values of 60, 50, 40 and so on were assigned to NSH depending on the number of stands of cassava per plot as at the time of harvest.
7	Soil Type	STP	INTEGER (1)	STP values of 1, 2, 3, 4 were assigned to Sandy, Clay, Silt and Loamy soils respectively
8	Crop Variety	CVT	INTEGER (1)	Values of 1, 2, 3, 4, 5 were assigned to Poundable, Hope, Game-changer, Baba-70 and Obasanjo-2 cassava varieties respectively.
9	Pest Control	PTC	INTEGER (1)	Value of 1 or 2 was assigned to traditional or pesticide operations
10	Disease Control	DSC	BINARY (1)	DSC value of 1 was assigned if diseases was controlled and 0 otherwise.
11	Weed Control	WDC	INTEGER (1)	Value of 1 or 2 was assigned to traditional or herbicide operations respectively.
12	Crop Duration	CRD	INTEGER (2)	Values of 6, 8, 12 were assigned to CRD based on the duration of crop from planting time to harvesting time.
13	Planting Method	PLM	INTEGER (1)	PLM had value of 1 for crops planted on flat surface and 2 for those planted on ridges.
14	Unforeseen Event	UFE	INTEGER (2)	UFE could take values between (-1,1), The values could be negative depending on factors such as total lock-down, terrorist attacks on farmers, massive death of farmers and so on. It could also be positive based on Government incentives to farmers, favourable policies on agricultural production and so on.
15	Crop Yield	CY	FLOAT (6)	CY measured in tones/hectare, served as the output of the system.

3.2 CYP modelling

The DL framework for CY prediction (Fig. 2) was adapted from demand prediction model and disease diagnosis model reported in [39] and [41] respectively. The CY model comprises input layer, hidden layers and output layer. The input layer accepts values of variables such as Rainfall/Irrigation (RFI), Temperature (TMP), Number of Stems Cultivated (NSC), Fertilizer Application (FZA), Manure Application (MNA), Number of Stands at Harvest (NSH), Soil Type (STP), Crop Variety (CVT), Pest Control (PTC), Disease Control (DSC), Weed Control (WDC), Crop Duration (CRD), Planting Method (PLM), Unforeseen Event (UFE), and so on that influence the yield of cassava and other crops.

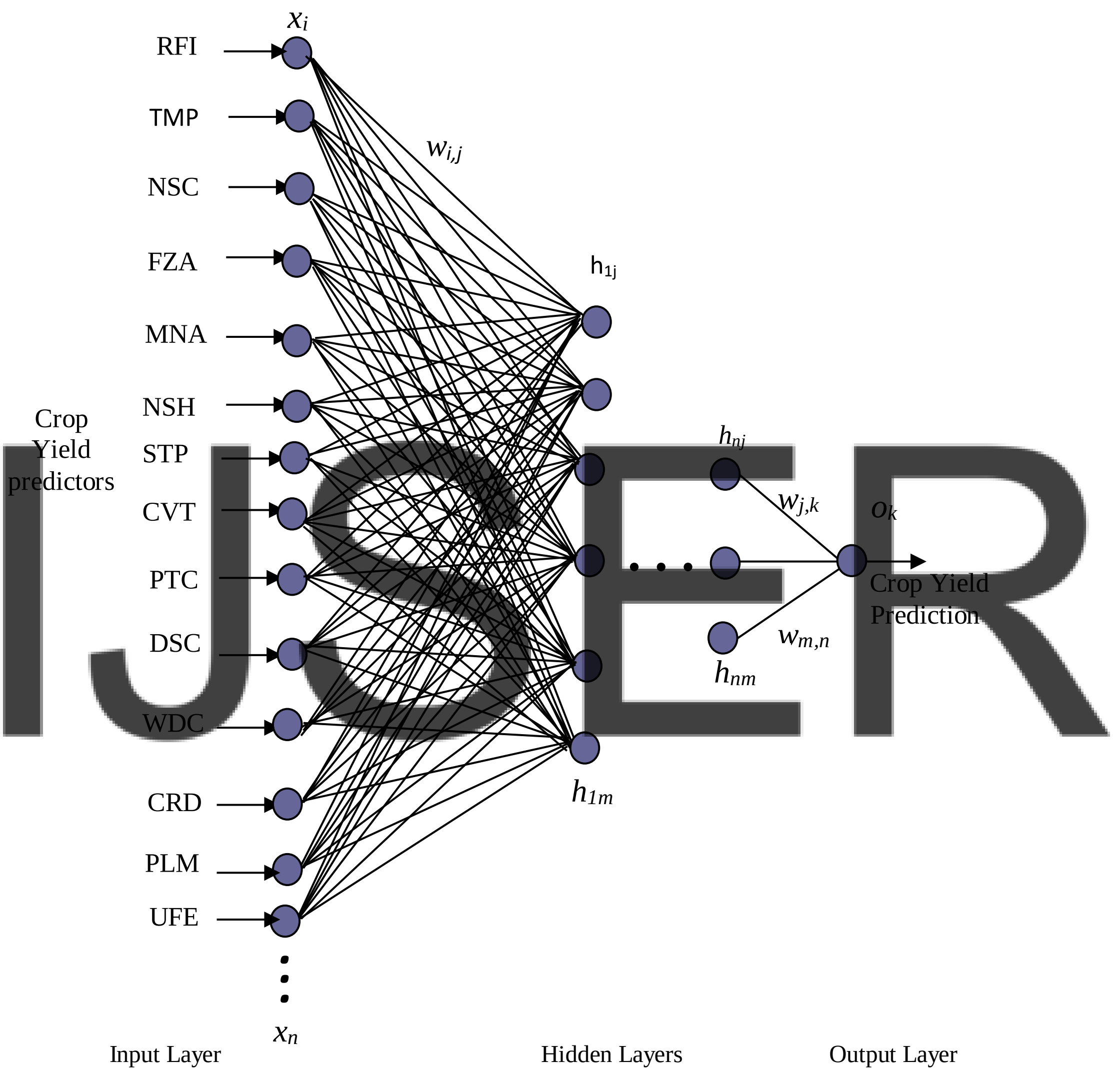


Fig. 2. DL Schema for CYP

The value of each node in the first hidden layer is the sum of products of inputs that influences CY and their respective weights. The value of each hidden layer node generates the output for that node via the activation function while the output of the first hidden layer becomes the input to the second hidden layer and so on. The process continues till the final hidden layer sends its results as input to the output layer which computes its output (CYP value) via output layer activation function. In the DL schema of Fig. 2, let the layers be represented as follows:

i. CY input (variable) layer x_i : $i = 1,2,...,n$

ii. CY hidden (processing) layer h_j : $j = 1,2...,m$

iii. CY output (prediction) layer o_k : $k = 1,2,...,l$

A system of equations is formulated for the input and hidden layers in Eq. (1) and compressed in Eq. (2), where $W_{i,j}$ represents the matrix of weights on the connection from the *ith* node in the input layer to the *jth* node in the hidden layer. W is the matrix of weights, x_i is the crop yield input vector and h^* is the hidden layer pre-output. The actual output of hidden layer node h_j is obtained by subjecting the pre-output in Eq. (3) to the hyperbolic transfer function as shown in Eq. (4). Similarly, the computation in the output layer node is performed. $W_{j,k}$ is a matrix of weights that connects *jth* node in the hidden layer to the *kth*-output layer. The output layer equation is formulated as shown in Eq. (5), it is composed in vector form Eq. (6) and compressed as shown in Eq. (7)

$$\begin{array}{ccccccccc} w_{1,1}x_1 & + & w_{1,2}x_2 & + & \cdots & + & w_{1,n}x_n & = & h_1^* \\ w_{2,1}x_1 & + & w_{2,2}x_2 & + & \cdots & + & w_{2,n}x_n & = & h_2^* \\ & & & & \vdots & & & & \\ w_{m,1}x_1 & + & w_{m,2}x_2 & + & \cdots & + & w_{m,n}x_n & = & h_m^* \end{array} \tag{1}$$

$$\begin{array}{ccccc} W & & x & & h^* \\ \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ & & \vdots & \\ w_{m,1} & w_{m,2} & \cdots & w_{m,n} \end{bmatrix} & & \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} & = & \begin{bmatrix} h_1^* \\ h_2^* \\ \vdots \\ h_m^* \end{bmatrix} \end{array} \tag{2}$$

$$h_j^* = \sum_{j=1}^{m} \sum_{i=1}^{n} W_{i,j} x_i \tag{3}$$

$$h_j = \frac{e^{h_j^*} - e^{-(h_j^*)}}{e^{h_j^*} + e^{-(h_j^*)}} \tag{4}$$

$$w_{1,1}h_1 + w_{1,2}h_2 + \cdots + w_{1,m}h_m = o_k^* \tag{5}$$

$$\begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,m} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix} = o_k^* \tag{6}$$

$$o_k^* = \sum_{k=1}^{p} \sum_{j=1}^{m} w_{j,k} h_j \tag{7}$$

$$o_k = \frac{1}{1 + e^{-(o_k^*)}} \tag{8}$$

$$e_k = d_k - o_k \tag*{(9}$$

$$\mathrm{SSE} = \sum_{k=1}^{p} (e_k)^2 \tag{10}$$

$$e_j = h_j\left(1 - h_j\left(\sum_{j=1}^{m} w_{jk} e_k\right)\right) \tag{11}$$

$$w_{jk}(n+1) = w_{jk}(n) + \beta e_k h_j \tag{12}$$

$$w_{ij}(n+1) = w_{ij}(n) + \beta e_j x_i \tag{13}$$

$$P(x_i, w_{ij}, w_{jk}, o_k) = \begin{cases} High\ (H) & , \quad o_k > \theta \\ Moderate\ (M), & \quad o_k = \theta \\ Low\ \ (L) & , \quad o_k < \theta \end{cases} \tag{14}$$

The actual value of the output layer node (o_k), is obtained by subjecting the pre-output in Eq. (7) to the sigmoid transfer function as shown in Eq. (8). The final output (o_k), represents the value of crop (cassava) yield prediction. If the difference (error) between the computed output (o_k) and the desired output (d_k) is greater than a predefined value. The error value is deployed to adjust the connecting weights of output and hidden layers as seen in Eq. (12) and Eq. (13) respectively. The process is repeated until the value of sum of squared error (SSE) in Eq. (10) is within a prescribed tolerance. Subsequently, the yield predictors (x_i), adjusted hidden layers weights (w_{ij}) in Eq. (13) adjusted output layer weights (w_{kj}) in Eq. (12), DL computed output (o_k) in Eq. (8) and the crop yield threshold value (θ) obtained from agriculturist are deployed for crop yield prediction in Eq. (14).

3.3 Algorithmic model

The algorithmic description of the pseudo-code for CY prediction is given as follows:

1. CYP (x_i, w_{ij}, t_i, h_j^*, h_j, w_{jk}, o_k^*, β, o_k, d_k, θ, P)
2. BEGIN
3. // CYP parameters description
   x_i is the input variables; $i \in \mathbb{Z}^+$, $i = 1,2,\ldots,n$ // $\mathbb{Z}^+$ is a set of positive integers
   w_{ij} is the matrix of connecting weights between the input layer nodes and hidden layer nodes
   t_i is the number of training instances; $i \in \mathbb{Z}^+$, $i = 1,2,\ldots,n$
   h_j^* hidden layer sum-of-product variable
   h_j is the hidden layer node, $j \in \mathbb{Z}^+$, $j = 1,2\ldots,m$
   w_{jk} is the matrix of connecting weights between the hidden layer and output layer
   o_k^* is output layer sum-of-product variable
   β is a set of training biases
   o_k is the computed output
   d_k is the desired output
   θ is crop yield threshold value
   P is CYP function.
4. // Initialization of variables and weights
5. Initialize hidden and output layer variables for sum-of-product: h_j^*, $o_k^* = 0$
6. Initialize connection weights for all neurons to random numbers between 0 and 1
7. Select the first training input and output pair from the vector (x_i, d_k)
8. GET t
9. DO WHILE $t \leq n$
10. FOR $j = 1$ to m
11. Compute sum-of-products for hidden layer (h_j^*) using Eq. (3)
12. Compute the value of hidden layer node h_j via transfer function using Eq. (4)
13. NEXT j
14. Compute sum-of-product for output layer (o_k^*) using Eq. (7)
15. Compute the value of output layer (o_k) via transfer function using Eq. (8)
16. Evaluate the error between the desired output and the computed output using Eq. (9)
17. Compute the sum of squared error (SSE) using Eq. (10)
18. IF the computed training error is equal to or below a given threshold THEN
19. GO TO step 24
20. ELSE

21. Adjust output layer connection weights using Eq. (12)
22. Adjust the hidden layer connection weights using Eq. (13)
23. GO TO step 8
24. STOP Training
25. END IF
26. END DO
27. Load Test data
28. Compare desired output with computed output using Eq. (9)
29. IF the test error is greater than a given threshold THEN
30. Adjust training parameters
31. GO TO step 8
32. ELSE
33. STOP Testing
34. END IF
35. Evaluate model performance using the metrics in Table 4
36. IF values of performance metrics are satisfactory THEN
37. GO TO step 40
38. ELSE
39. GO TO step 30
40. Predict cassava yield using Eq. (14)
41. Make decisions based on predicted value in step 40
42. END IF
43. END // end algorithm CYP

4. Results and discussion

A 5-layer (14-6-4-3-1), DL model for CYP was implemented using neural network multilayer perceptron paradigm provided by NeuroSolutions version 7.0 as depicted in Fig. 4. The model had fourteen input data variables and one output data variable as described in serial numbers 1-14 and 15 respectively in Table 1. In the hidden layers, varying numbers of nodes (between six and three) were deployed for the processing and modelling of CY parameters, while at the output layer, one node was deployed for CYP. The data totaling 2500 samples were split into training, testing and validation data sets in the ratio of 8:1:1. This translated to 2000, 250 and 250 data samples respectively. The model was trained using back-propagation algorithm. Hyperbolic transfer function was deployed in the hidden layers while Sigmoid transfer function was used in the output layer.

In Fig. 4. values of 0.0024, 0.1244, 0.9568, 4.0875, -432,6789 and -323.7566 were observed for mean squared error (MSE), normalized mean squared error (NMSE), correlation coefficient (r), percentage error (%), akaike information criteria (AIC) and minimum descriptive length (MDL) respectively. The observed training MSE value of 0.0024 and correlation value of 0.9568 showed a satisfactory DL process. However, training parameters could be adjusted to reduce the error and increase the accuracy.

The DL model was designed for predicting the CY and trained with data variables whose levels of contribution to CY were determined as depicted in Fig. 3. FZA had the highest contribution to CY, followed by NSH and STP. The contribution of RFI, WDC and CRD were almost at the same level. MNA took the seventh position followed by CVT, DSC, TMP, PTC, UFE, PLM and NSC. Although FZA was observed as the major contributing factor to the yield of cassava. It was observed that other factors such as UFE, PLM and NSC which occupied the last three positions could cause major deviations in crop yield predictions. For instance, in year 2019 to 2020 there was an unforeseen event (UFE) which seriously affected crop yield. The total lockdown due to corona virus pandemic in the world hindered farming activities which resulted in low CY.

In the agricultural zone in which the study area belongs, there are contentions among farmers on number of stems cultivated (NSC) per stand and planting method (PLM). Many farmers opined that they have high yield from planting two (2) or three (3) stems of cultivar per stand, only few farmers subscribe to planting of one (1) stem per stand. Some farmers orated that they have high yield from planting on ridges while others maintained that they have high yield from planting on flat surface. Hence, there is need to consider the aforementioned predictors in CYP, despite their low grades in predictor's scale of importance.

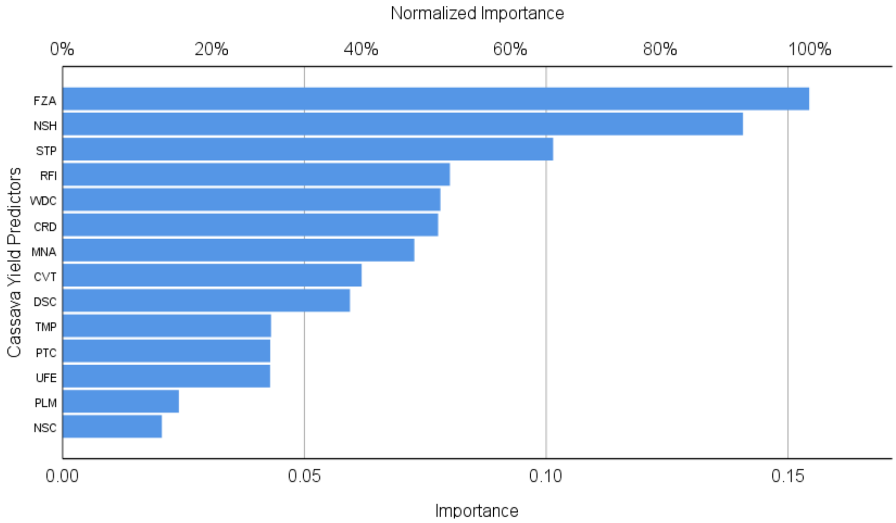


Fig. 3. Cassava yield predictor's importance

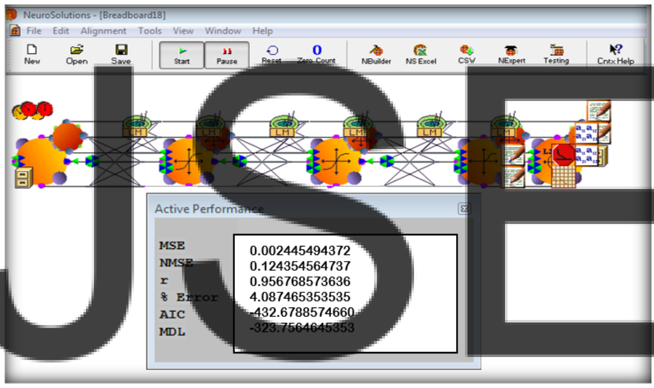


Fig. 4. DL training panel

Table 2
CYP Correlations

CASSAVA YIELD PREDICTORS	YIELD	RFI	TMP	NSC	FZA	MNA	NSH	STP	CVT	PTC	DSC	WDC	CRD	PLM	UFE
YIELD	1	0.253	0.028	0.096	0.921	0.896	1.000	0.743	0.029	0.321	0.422	0.221	0.335	0.348	0.071
RFI	0.253	1	-0.338	0.146	0.100	0.146	-0.251	-0.223	0.000	-0.100	0.100	-0.100	0.035	0.048	0.059
TMP	0.028	0.338	1	0.048	-0.132	0.048	-0.029	0.089	0.000	0.132	-0.132	0.132	0.076	0.044	0.115
NSC	0.096	0.146	0.048	1	0.030	1.000	-0.099	0.014	0.000	-0.030	0.030	-0.030	0.018	0.030	0.061
FZA	0.921	0.100	-0.132	0.030	1	0.030	0.223	-0.179	0.000	-1.000	1.000	-1.000	0.000	0.000	-0.387
MNA	0.896	0.146	0.048	1.000	0.030	1	-0.099	0.014	0.000	-0.030	0.030	-0.030	0.018	0.030	0.061
NSH	1.000	-0.251	-0.029	-0.099	0.223	-0.099	1	-0.004	0.030	-0.223	0.223	-0.223	-0.335	-0.350	-0.070
STP	0.743	-0.223	0.089	0.014	-.0179	0.014	-0.004	1	0.000	0.179	0-.179	0.179	0.017	0.000	0.106
CVT	0.029	0.000	0.000	0.000	0.000	0.000	0.030	0.000	1	0.000	0.000	0.000	0.000	0.000	0.021
PTC	0.321	-0.100	0.132	-0.030	-1.000	-0.030	-0.022	0.179	0.000	1	-1.000	1.000	0.000	0.000	0.387
DSC	0.422	0.100	-0.132	0.030	1.000	0.030	0.223	-0.179	0.000	-1.000	1	-1.000	0.000	0.000	-0.387
WDC	0.221	-0.100	0.132	-0.030	-1.000	-0.030	-0.223	0.178	0.000	1.000	-1.000	1	0.000	0.000	0.387
CRD	0.033	0.035	0.076	0.018	0.000	0.018	-0.335	0.017	0.000	0.000	0.000	0.000	1	0.962	0.509
PLM	0.034	0.048	0.044	0.030	0.000	0.030	-0.350	0.000	0.000	0.000	0.000	0.000	0.962	1	.0387

UFE	0.071	0.059	0.115	0.061	-0.387	0.061	-0.070	0.106	0.021	0.387	0-.387	0.387	0.509	0.387	1

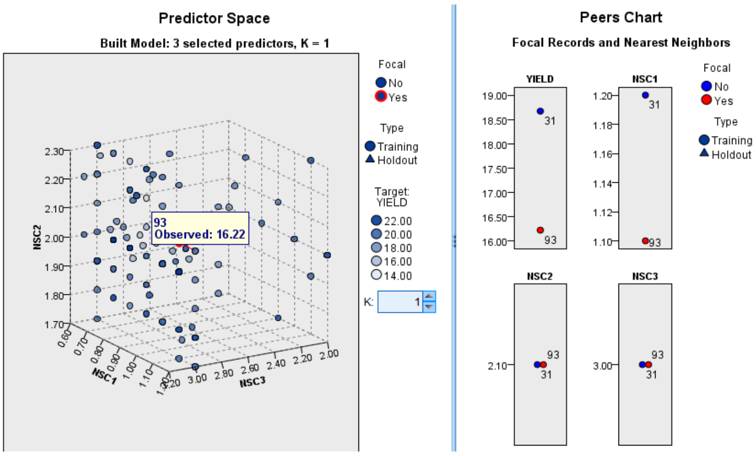


Fig. 5a. Peer chart of cassava yield with number of stems cultivated using K=1

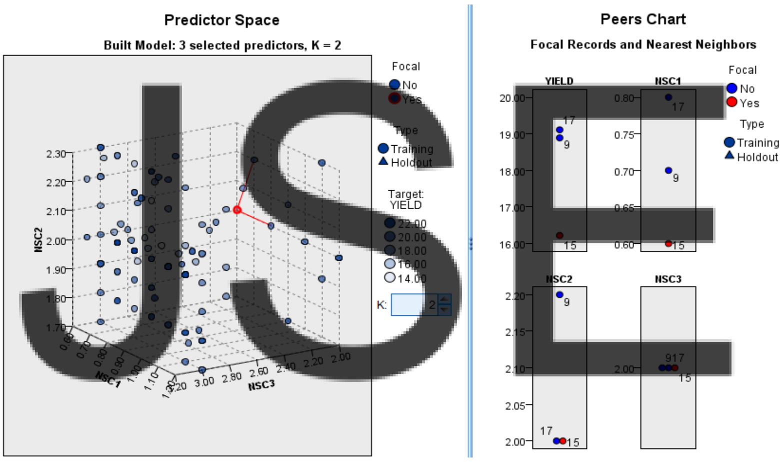


Fig. 5b. Peer chart of cassava yield with number of stems cultivated using K=2

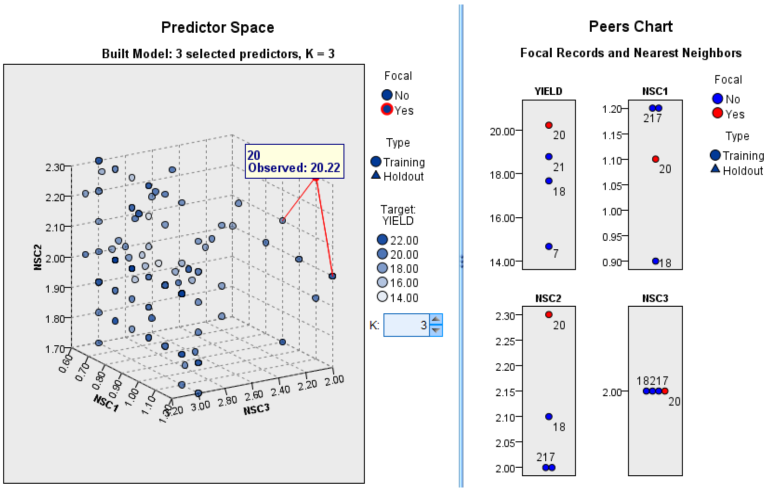
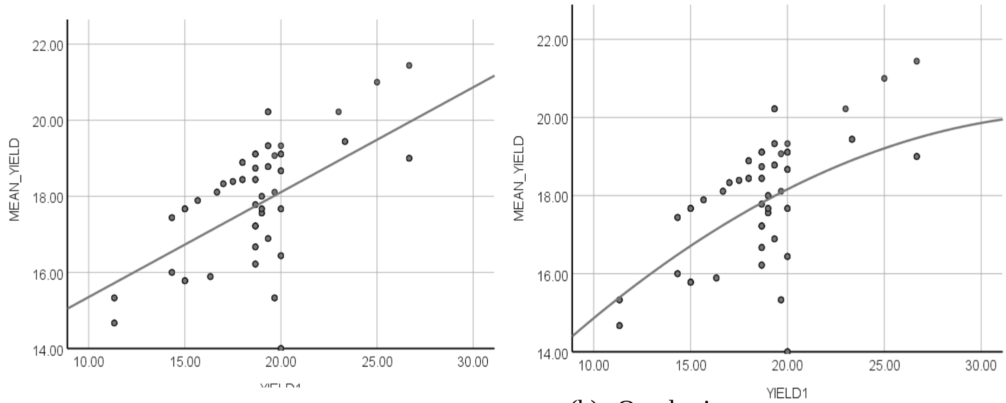


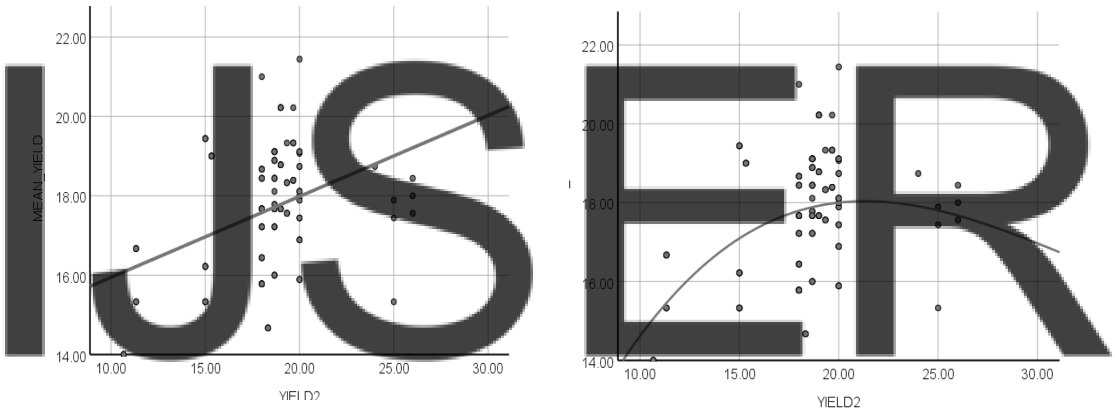
Fig. 5c. Peer chart of cassava yield with number of stems cultivated using K=3

.



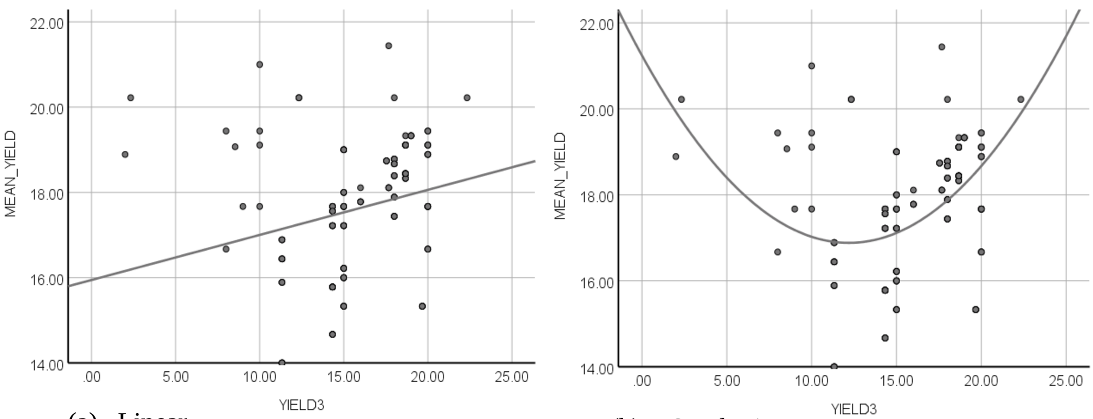
(a) Linear (b) Quadratic

Fig. 6. Graph of mean cassava yield with yield of planting one stem per stand



(a) Linear (b) Quadratic

Fig. 7. Graph of mean cassava yield with yield of planting two stems per stand



(a) Linear (b) Quadratic

Fig. 8. Graph of mean cassava yield with yield of planting three stems per stand

The correlation between cassava yield and predictors (variables) are presented in Table 2. For instance, the correlation between yield of cassava and rainfall (RFI) is valued at 0.253, while that of fertilizer application (FZA) is valued at 0.921. It implies that although both variables contribute positively to the yield of cassava, the contribution of fertilizer is far greater than that of rainfall. A perfect correlation was observed between cassava yield and the number of stands at harvest time (NSH). That means if there are no stands at harvest there are no yield. Cassava yield increases proportionately with the number of stands at harvest. A positive correlation value of 0.096 was observed between cassava yield and the number of stems cultivated (NSC) at planting time. Further investigation to decipher the yield obtained by planting one, two or three stems of cassava cultivars per stand was carried out.

The CY at any point together with its neighbours referred to as k were observed. Using k =1, at data point number 93 as shown in Fig. 5a and in many other data points, it was generally observed that the average cassava yield was closer to the yield obtained from planting one (1) stem of cassava (NSC1) per stand. The yield obtained by planting two (2) stems (NSC2) as well as planting three (3) stems (NSC3) per stand were relatively similar, with NSC2 yield slightly greater than NSC3 yield. Similarly, using k =2 as shown in Fig. 5b, average cassava yield was closer to the yield obtained from planting one (1) stem of Cassava (NSC1) per stand than the yield obtained by planting two (2) stems (NSC2) as well as planting of three (3) stems (NSC3) per stand. At k =3 as shown in Fig. 5c, average cassava yield was closer to the yield obtained from planting two (2) stems of Cassava (NSC2) per stand. Average yield at k =3 deviated significantly from the yield obtained by planting one (1) stem (NSC1) as well as planting of three (3) stems (NSC3) per stand. The deviation in yield patterns could be attributed to interaction of cassava variety, soil type, soil nutrients, planting methods and environmental factors.

Figs. 6 - 8 depict the plots of CY represented by yield1, yield2 and yield3 produced by implementing NSC1, NSC2 and NSC3 respectively at different locations but relatively similar plots of land. In Fig. 6a, the linear interaction between the mean CY with the first yield (yield 1) obtained by cultivating one stem of cassava per stand is presented. The linear graph shows a positive contribution of yield1 to the mean yield of cassava while the quadratic graph in Fig .6b reveals that the contribution of yield1 to mean yield has not reached the maximum point as indicated by the maximum curve that is yet to reach the turning point. Similarly, the linear graph in Fig. 7a shows a positive contribution of yield obtained by planting two cassava stems per stand (yield2) to mean yield. However, the quadratic representation in Fig. 7b shows that the yield2 has maximally contributed to the mean yield as indicated by the turning point. Fig. 8a shows a positive linear contribution of yield obtained by planting three stems per stand (yield3) to the mean yield of cassava. In the quadratic representation of Fig. 8b, the minimum curve shows that planting three cultivars per stand has the least possible contribution to the average yield of cassava.

Evaluation of system performance was carried out using test dataset consisting of 250 samples. Actual values and predicted values obtained from the model were compared. Mean threshold value was obtained from five (5) agriculturists. The threshold value was used to determine the high, moderate and low yield of Cassava prediction. Values above the threshold were considered as High (H), values equal to the threshold value were considered Moderate (M) and values below the threshold were considered as Low (L). The matching of the actual and the predicted values was viewed as a correct prediction and the mismatch was viewed as incorrect prediction.

The matching of high or moderate actual value with high or moderate predicted value served as True Positive (TP). The matching of low actual value with low predicted value was viewed as True Negative (TN). The matching of high or moderate actual value with low predicted value was viewed as False Positive (FP), while the matching of low actual value with high predicted value was referred to as False Negative (FN). Samples of outcome from test dataset matched with mean threshold value of 17 tones/ha of cassava yield is presented in Table 3. The test dataset prediction outcomes for TP, FP, TN and FN were 171, 9, 63 and 7 respectively.

Table 3
Sample of Actual and Predicted Cassava Yield

Test Data ID	Actual Yield	Predicted Yield	Match	Outcome
1	20.22	19.82	HH	TP
2	17.00	17.00	MM	TP
3	19.07	19.19	HH	TP
4	17.22	16.81	HL	FP
5	16.22	16.59	LL	TN
6	21.00	20.82	HH	TP
7	19.11	19.23	HH	TP
8	17.67	18.21	HH	TP
9	16.33	17.19	LH	FN
10	20.22	20.10	HH	TP
11	21.44	20.50	HH	TP
12	18.44	18.30	HH	TP
13	16.67	17.32	LH	FN
14	19.22	18.23	HH	TP
15	18.89	19.29	HH	TP
16	15.33	16.43	LL	TN
17	20.22	18.23	HH	TP
18	17.56	16.23	HH	TP
19	18.74	18.43	HH	TP
20	17.22	16.10	HL	FP
…	…	…	…	…
250	13.44	14.00	LL	TN

Table 4
Evaluation Metrics

SN	Evaluation Metric	Formula	Computation	Value
1	Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	$\frac{171+63}{171+63+9+7}$	0.9360
2	Recall, Sensitivity or True Positive Rate (TPR)	$\frac{TP}{TP+FN}$	$\frac{171}{171+7}$	0.9607
3	Specificity or True Negative Rate (TNR)	$\frac{TN}{TN+FP}$	$\frac{63}{63+9}$	0.8750
4	Precision or Positive Predictive Value (PPV)	$\frac{TP}{TP+FP}$	$\frac{171}{171+9}$	0.9500
5	Negative Predictive Value (NPV)	$\frac{TN}{TN+FN}$	$\frac{63}{63+7}$	0.9000
6	Miss-rate or False Negative Rate (FNR)	$\frac{FN}{FN+TP}$	$\frac{7}{7+171}$	0.0393
7	Fall-out or False Positive Rate (FPR)	$\frac{FP}{FP+TN}$	$\frac{9}{9+63}$	0.1250
8	False Discovery Rate (FDR)	$\frac{FP}{FP+TP}$	$\frac{9}{9+171}$	0.0500
9	False Omission Rate (FOR)	$\frac{FN}{FN+TN}$	$\frac{7}{7+63}$	0.1000
10	F1 score (Harmonic mean of Precision and Recall)	$\frac{2PPV * TPR}{PPV+TPR}$	$\frac{2 * 0.95 * 0.9607}{0.95+0.9607}$	0.9553
11	Critical Success Index or Threat Score (TS)	$\frac{TP}{TP+FN+FP}$	$\frac{171}{171+7+9}$	0.9144
12	Positive Likelihood Ratio (LR+)	$\frac{TPR}{FDR}$	$\frac{0.9607}{0{,}0500}$	19.2135
13	Diagnostic Odds Ratio (DOR)	$\frac{PPV * NPV}{(1-PPV)*(1-NPV)}$	$\frac{0.95 * 0.90}{(1-0.95)*(1-0.90)}$	171
14	Standard error of Log Diagnostic odds ratio (se(logDOR))	$\sqrt{\frac{1}{TP}+\frac{1}{FN}+\frac{1}{FP}+\frac{1}{TN}}$	$\sqrt{\frac{1}{171}+\frac{1}{7}+\frac{1}{9}+\frac{1}{63}}$	0.5251
15	Negative Likelihood Ratio (LR-)	$\frac{FNR}{TNR}$	$\frac{0.0393}{0.8750}$	0.0449

Some metrics reported in [36] such as (accuracy, recall, specificity, precision and many others) could help investigation of prediction parameters to guide decision making. In this work, evaluation metrics are presented in Table 4. The Accuracy value of 93.60% means that more than 93 out of every 100 cases predicted correctly while less than 7 predicted incorrectly. Recall or Sensitivity value of 96.07% implies that less than 4 cases out of 100 cases are miss-labelled as high yield by the program in the task of predicting cassava high yield while more than 96 cases are correctly predicted as cassava high yield. Specificity value of 87.50% means that less than 13 cases out of 100 cases are miss-labelled as cassava low yield while more than 87 cases are correctly predicted as low yield of cassava. Precision value of 95% means that on average, 5 out of 100 cases of cassava high yield predictions are wrongly predicted while 95 cases are correctly predicted. Diagnostic odds ratio value of 171 is greater than 1, it indicates that the model is discriminating correctly in predicting both the high and low CY. The false positive rate and false negative rate values obtained from the prediction model indicate that the failure of the system to correctly predict high yield and low yield of cassava are 0.1250 and 0.0393 respectively.

Sequel to the minimal values of high and low yield prediction errors, the model is poised and could be relied on to guide farmers and agricultural stakeholders in planning and making wise decisions towards storage, purchase or marketing of cassava products for enhancement of food availability and maintenance of economic stability.

Summary of findings in this work are presented as follows:

1. FZA, NSH and STP were observed as three most important variables in CY prediction
2. Other variables such as RFI, WDC, CRD. MNA, CVT, DSC, TMP, PTC, UFE, PLM and NSC arranged in decreasing order of importance were observed to contribute positively to CY prediction
3. Planting of one stem of cassava cultivar per stand was observed to contribute more to average yield than planting two or three stems per stand.
4. Hence, this work scientifically settled the contention that existed among farmers on the number of cultivars to plant per stand in Akwa Ibom North-West Agricultural zone, Nigeria.
5. The DL model could decipher non-linear relationships among predictors and could predict high, moderate and low yield of cassava using CY threshold values, current predictors values and previous knowledge of CY parameters preserved as connection weights in the neural network hidden and output layers.
6. The DL model with three (3) hidden layers generated the least mean squared error value of 0.0024 compared to others.
7. In the DL model, accuracy, recall, specificity and precision metrics values of 93.60%, 96.07%, 87.50% and 95.00% respectively were observed. This indicate high performance capability of the DL model and its suitability for deployment in the prediction process.
8. The 93.60% accuracy of the DL model in this study outperforms the 84 % accuracy of ARIMA model reported in [43] as well as 84 % accuracy of the mathematical model reported in [44].

5. Conclusion

In this work, a DL model for prediction of CY has been presented. Correlation of CY input variables with mean CY was reported. Investigation into the accuracy of the model was carried out by comparing the actual and predicted cassava yield from the test dataset. System performance evaluation was carried out using fifteen metrics as well as comparison with other prediction models. Insignificant prediction errors were observed in the DL model. Based on research findings this work recommends that farmers should plant one stem of cassava cultivar per stand. This work has contributed a DL crop yield prediction algorithm to guide prediction and decision processes in agricultural sector. The model could learn from previous crop yield data and extrapolate into unseen patterns. It could predict the yield of cassava one year into the future. This model is poised to serve as a guide to farmers and agricultural decision makers in planning for storage and marketing of cassava products in event of high yield prediction as well as planning for alternative source of food for the teaming population in event of low yield prediction.

The model would guide stakeholders in the agricultural sector towards making informed decisions about storage, purchase or marketing of cassava products to ensure continuous food supply and unceasing income generation. Detailed DL investigations of cassava yield patterns in response to cassava varieties, soil types, fertilizer and manure types, weather conditions, cultural practices and planting methods are recommended for further research. The flexibility of this algorithm allows adaptability to solve prediction problems in other domains. In order to reduce training time and increase prediction efficiency, investigations of optimization techniques to guide selection of DL layers, training and testing parameters are also recommended for further research.

Acknowledgement

The authors of this work are thankful to the Tertiary Education Trust Fund (TETFund) for their support of this research through the TETFund Centre of Excellence in Computational Intelligence Research, University of Uyo and the University of Uyo management team for the support and provision of enabling environment.

References

[1] Aggarwal C. C. (2018). Neural Networks and Deep Learning. Springer International Publishing. Springer Nature. ISBN 978-3-319-94463-0. https://doi.org/10.1007/978-3-319-94463-0.

[2] Al I., Pasajee K., Vũ A. (2020). Cassava Breeding and Agronomy in Asia: 50 years of History and Future Directions. Journal of Breeding Science, 70(2),145–166.

[3] Alison Borges Vitor, Rafael Parreira Diniz, Carolina Vianna Morgante, Rafaela Priscila Antônio, Eder Jorge de Oliveira. (2019). Early prediction models for cassava root yield in different water regimes. https://www.sciencedirect.com/science/article/abs/pii/ S03784 2901930228X

[4] Amanda P. D, Lynnicia N. M, Deepak J., Siyuan C., Rachel S. & Stephen P. L., (2017). Rooting for cassava: insights into photosynthesis and associated physiology as a route to improve yield potential. Journal of New Phytologist 17 (213), 50–6.

[5] Aparna R. Rout, Sahebrao B. Bagal. (2017): A deep learning model for image classification. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, 4(5) 814-818.

[6] Aravind Pai. (2020). CNN vs. RNN vs. ANN: Analyzing 3 Types of Neural Networks in Deep Learning. *https://medium.com/analytics-vidhya/cnn-vs-rnn-vs.*

[7] Bu F. & Wang X. (2019): A smart agriculture IoT system based on deep reinforcement learning: Future Generation of Computing System. https://www.sciencedirect.com/science/article/pii/S0167739X19307277

[8] Ceballos, H., Kawuki, R. S., Gracen,V.E., Yencho G.C. & Hershey C.H. (2015). Conventional breeding, marker-assisted selection, genomic selection and inbreeding in clonally propagated crops: a case study for cassava. Journal of Theory and Applications Genet. 128(5), 1647–1667.

[9] Chetty C. C., Rossin C. B., Gruissem W, Vanderschuren H, Rey M. E. (2013). Empowering biotechnology in southern Africa: establishment of a robust transformation platform for the production of transgenic industry-preferred cassava. New Biotechnology 30: 136–143.

[10] Dankers, Frank & Traverso, Alberto & Wee, Leonard & Kuijk, Sander. (2019). Prediction Modeling Methodology. https://pubmed.ncbi.nlm.nih.gov/31314250

[11] Ezedinma C.I., Kormawa P.M., Manyong V.M. & Dixon A.G. (2007). Challenges, opportunities, and strategy for cassava sub sector development in Nigeria. International Institute of Tropical Agriculture, Ibadan, Nigeria. Proceedings of the 13th ISTRC Symposium, pp.627-640.

[12] Filippi, E.J. Jones, N.S. Wimalathunge, P.D.S.N. Somarathna, L. E. Pozza, S.U. Ugbaje, T.F.A. Bishop. (2019). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. https://www.semanticscholar.org/paper/An-approach

[13] Fondong, V.N. (2017). The search for resistance to cassava mosaic gemini viruses: How much we have accomplished, and what lies ahead. Journal of Science Front Plant 8(1), 408-420.

[14] Frooq M. and Pisante M. (2019). Innovations in Sustainable Agriculture. https://link.springer.com/book/ 10.1007/978-3-030-23169-9

[15] Hajir Almahdi (2020). Predicting crops yield growth. https://towardsdatascience.com/predicting-crops

[16] He K., Zhang X., Ren S., & Sun J. (2016). Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

[17] Howeler, R. H. & Aye, T. M. (2014). Sustainable management of cassava in Asia—From research to practice. International Center for Tropical Agriculture (CIAT), The Nippon Foundation, Vietnam, p. 168

[18] Jin-yue Liu & Bao-ling Zhu. (2012): Research on the Prediction Model of Total Agricultural Output Value based on Wavelet Neural Network. https://www.atlantis-press.com/proceedings/citcs-12/3018.

[19] Kawano, K. (2011). The triumphant cassava chronicled by foresight, political will and accountability. Journal of Root Crops 37 (1),101–110.

[20] Kintché, S. Hauser, N.M. Mahungu, A. Ndonda, S. Lukombo, N. Nhamo, V.N.E. Uzokwe, M. Yomeni, J. Ngamitshara, B. Ekoko, M. Mbala, C. Akem, P. Pypers, K.P. Matungulu, A. Kehbila, B. Vanlauwe. (2017). Cassava yield loss in two provinces of the Democratic Republic of Congo. https://www.semanticscholar.org/paper/Cassava-yield

[21] Latif S. & Meuller J. (2015). Potential of cassava leaves in human nutrition: a review. Trends in Food Science & Technology 44:147–158. https://www.academia.edu/20880191/Potential_of.

[22] Lynam, J. & Byerlee D. (2017). Contributing to a sustainable food future. Publication of International Center for Tropical Agriculture (CIAT), Cali, Colombia. 444(3) 140-150.

[23] Lukuyu B., Okike I, Duncan A., Beveridge M., Bleummel M. (2014). Use of cassava in livestock and aquaculture feeding programs. Nairobi, Kenya: International Livestock Research Institute. https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje

[24] Murthy, V.R.K. (2011). Crop Growth Modeling and Its Applications in Agricultural Meteorology. https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje

[25] Oni O.V. and Y.O. Akanle. (2018). Comparison of Exponential Smoothing Models for Forecasting Cassava Production. https://www.isroset.org/journal/IJSRMSS/full_paper_view.php?paper_id=674

[26] Odubanjo, O. O., Olufayo, A.A. & Oguntunde, P.G. (2011). Water use, growth, and yield of drip irrigated cassava in a humid tropical environment. Soil Water Res., 6(1), 10-20.

[27] Patricia Moreno-Cadena, Gerrit Hoogenboom, James H. Cock, Julian Ramirez-Villegas, Pieter Pypers, Christine Kreye, Meklit Tariku, Kodjovi Senam Ezui, Luis Augusto Becerra Lopez-Lavalle, Senthold Asseng. (2021). A review of the growth, development and yield of cassava models. https://pubmed.ncbi.nlm.nih.gov/34140751.

[28] Petteri Nevavuori, Nathaniel Narra, and Tarmo Lipping. (2019). Crop yield prediction with deep convolutional neural networks, Computers and Electronics in Agriculture. https://europepmc.org/article/AGR/IND606494900

[29] Phnom P. (2015). Cassva Hanbook: China-Cambodia-UNDP Trilateral Cooperation Cassava Project Phase II. (1),1-144.

[30] Purnamasari, R.A., Ahamed, T. & Noguchi, R. (2019). Land suitability assessment for cassava production in Indonesia using GIS, remote sensing and multi-criteria analysis. Asian Journal of Science, 3, 1–3.

[31] Reinhardt, H. H. (2010). Agronomic Practices for Sustainable Production of Cassava in Asia. Publication of Cassava Office for Asia, Department of Agriculture, Chatuchak, Thailand. 1(2), 288-314.

[32] Saeed Khaki, Lizhi Wang, and Sotirios V. Archontoulis. (2020). A CNN-RNN Framework for Crop Yield Prediction. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6993602

[33] Schlenker W., Lobell D. B. (2010). Robust negative impacts of climate change on African agriculture. Research Letters. https://iopscience.iop.org/article/10.1088/1748-9326/5/1/014010.

[34] Sudhanshu Sekhar Panda, Daniel P. Ames, and Suranjan Panigrahi. (2010). Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. https://pdfs.semanticscholar.org/21a8/aaceb

[35] Thomas Van Klompenburg, Ayalew Kassahun & Cagatay Catal (2020). Crop yield prediction using machine learning: A systematic literature review, Computers, and Electronics in Agriculture. https://www.semanticscholar.org/paper/Crop-yield

[36] Tharwat, A. (2020). Classification assessment methods, New England Journal of Entrepreneurship. 17 (1) 168-192.

[37] Tunrayo A., Pieter P., Meklit C., Stefan H., Christine K., & Michael H. (2017). Using Geospatial tools to optimize cassava agronomy trials in Nigeria and Tanzania. International Institute of Tropical Agriculture (IITA), Ibadan. Transforming Agriculture, 50, 1967-2017.

[38] Soria-Ruiz, Fernandes-Ordonez, Quijano-Carranza, Macías-Cervantes, Sauceda, Gonzalez, Quintana. (2012). Remote Sensing and Simulation Model for Crop Management.

[39] S. S. Udoh and O. Olabode (2010). Artificial Neural Network for the Prediction of the Demand of Petrol Using Periodic Review s-S model in a Distribution Depot. Journal of Applied Mathematical and Computational Sciences (AMCOS), 1(2): 151-164.

[40] S. S. Udoh , O. C. Akinyokun , U. G. Inyang , O. Olabode , G. B. Iwasokun. (2017).Discrete Event-Based Hybrid Framework for Petroleum Products Pipeline Activities Classification. DOI: 10.5430/air.v6n2p39. Journal of Artificial Intelligence Research (AIR) 6(2), 39-50.

[41] Samuel S. Udoh, Uduak A. Umoh, Michael E. Umoh & Mfon E. Udo, (2019). Diagnosis of Prostate Cancer using Soft Computing Paradigms. Online ISSN: 0975-4172 & Print ISSN: 0975-4350. Global Journal of Computer Science and Technology (GJCST), Neural and Artificial Intelligence, 19(2), 19-26.

[42] Udosen C. E. (2012). Rainfall Trends in Uyo-Akwa Ibom State and its Implication on Urban Flooding. Journal of Engineering and Applied Sciences. 7(1):79-85.

[43] Ikuemonisan, E.S.,and Akinbola, A.E. (2021). ARIMA Forecasts of Cassava Production Indicators and its Implication for Future Food Supply in Nigeria. Tarım Ekonomisi Araştırmaları Dergisi, 7(1), 14-30.

[44] Sadenovaa, M. A., Beisekenova, N. A., Rakhymberdinaa, M., Varbanovb, P. S. , Klemešb, J. J., (2021). Mathematical Modelling in Crop Production to Predict Crop Yields. Journal of Chemical Engineering Transactions 88(2):1225-1230