Journal of Agriculture, Food, Environment and Animal Sciences
Tarım, Gıda, Çevre ve Hayvancılık Bilimleri Dergisi
http://www.jafeas.com, ISSN: 2757-5659

J. Agric. Food, Environ. Anim. Sci.
6(2): 419-440, 2025

Utilizing Data Mining Algorithms to Predict Body Weight in Nigerian Normal Feathered Local Chickens

Ugwumba Chijioke ISAAC[1*], Confidence Ugochiziterem ERNEST[2]

[1, 2] Department of Animal Science, Faculty of Agriculture, Nnamdi Azikiwe University P.M.B. 5025, Awka, NIGERIA

[1]https://orcid.org/0000-0002-7729-4516, [2]https://orcid.org/0009-0003-0821-4345

[*]Corresponding Author: uc.isaac@unizik.edu.ng

Research Article

Article History:
Received: 20 June 2025
Accepted: 22 September 2025
Published online: 01 December 2025

***Keywords*:**
Automatic Linear Modelling
Body Weight
Chicken
Decision Trees
Linear Body Measurements

ABSTRACT

This study aimed to predict body weight (BWT) from linear body measurements in normal feathered chickens. A total of 200 chickens aged between 8 and 10 months old, comprising 150 males and 50 females were randomly selected from the local market in Amansea, Awka North LGA, Anambra State, Nigeria. Automatic linear modelling (ALM), chi-square automatic interaction detection (CHAID), exhaustive CHAID, and classification and regression trees (CART) were employed for the prediction. The adjusted coefficient of determination (R^2adj.) and Akaike's Information Criterion corrected (AICc) were used to assess the predictive performance of ALM. Males exhibited significantly ($p<0.05$) greater BWT (1.26±0.02 kg vs 1.05±0.03 kg) and breast width (BW) (9.59±0.07 cm vs 9.15±0.15 cm) compared to females. Body length (BL) showed the strongest positive correlation with BWT in females ($r = 0.46$). BL, shank length and BW recorded the most significant ($p<0.001$) fractional importance in males (0.793), females (0.721), and pooled sexes (0.501), respectively using ALM. The highest predictive accuracy for ALM was observed in pooled sexes, indicated by the lowest AICc (-572.92) and a predicted BWT of 1.49 kg. Whereas CART did not identify any variable as influential, CHAID and exhaustive CHAID recognised BL as the most critical predictor, yielding a maximum predicted BWT of 1.32 kg when BW exceeded 9.00 cm. The findings highlight BW as the best predictor of BWT in a mixed flock of normal feathered chickens. It is recommended that breast width and body length be prioritised in selection strategies aimed at improving body weight in chicken.

To Cite : Isaac UC, Ernest CU., 2025. Utilizing Data Mining Algorithms to Predict Body Weight in Nigerian Normal Feathered Local Chickens. Agriculture, Food, Environment and Animal Sciences, 6(2): 420-441

INTRODUCTION

Indigenous poultry breeds play a significant role in the livelihoods of rural farmers in Nigeria, contributing to food security, income generation, and cultural heritage. The genetic potential of these breeds has remained largely untapped due to little or no improved genetic and environmental factors influencing their performance (Ojedapo et al., 2019; Isaac and Ezejesi, 2023).

Measurement of body weight of animals is vital in determining their market price and some management practices (Nwaogwugwu et al., 2018). Determination of accurate body weight value is often a critical challenge in rural places where many farmers lack scale for measurement. Even where scales exist, direct measurement of body weight of chickens is highly biased from the gut fill of the animals. The linear body measurements are not biased by the gut fill (Isaac and Adeolu, 2023), and for this reason, prediction of body weight from them is preferred.

Different methods, including ordinary multiple linear and stepwise multiple linear regressions have been employed in prediction experiments (Nwakpu et al., 2020; Isaac and Adeolu, 2023). Conventional regression models often fail to capture nonlinear relationships and interactions among predictors. This limitation can result in biased estimates and reduced predictive accuracy (Adeyinka et al., 2017). This has necessitated the use of automatic linear modelling (ALM) and decision trees in body weight prediction. ALM is used in prediction experiments, especially in cases of multi-predictors, automating the process of selecting the most suitable subset of predictors, which is particularly crucial when dealing with a large number of predictors (Oshima and Dell-Ross, 2016; Genç and Mendeş, 2021).

A decision tree is a predictive modelling tool used in machine learning and statistics involving mapping from observations about an item to conclusions about its target value (Matzavela and Alepis, 2021). It presents a predictive result in a tree-like structure with nodes and terminal nodes. Decision trees work by recursively partitioning the data into subsets based on the values of input features (Mushkudiani et al., 2008). The three growing methods or algorithms that make up the decision trees are Chi-square Automatic Interaction Detection (CHAID), exhaustive CHAID and classification and regression trees (CART). CHAID is a growing method of the decision tree algorithm used for classification and regression tasks. It recursively partitions the data into mutually exclusive and exhaustive subsets based on categorical predictor variables, aiming to maximize the homogeneity of the outcome variable within each subset (Seok and Kang, 2015). Exhaustive CHAID is a variant and extension of CHAID (Hothorn et al., 2006).

Classification and regression trees (CART), according to Breiman et al. (1984) is a predictive algorithm widely used in machine learning. CART models can be categorized based on the dependent variable. Categorical outcome variables require the use of a classification tree, while continuous outcome requires regression trees to identify the relationships among variables (Wray and Byers, 2020). CART constructs a binary decision tree structure where each fork represents a predictor variable, and each node provides a prediction for the target variable (Lee et al., 2010; Ali et al., 2015; Wray and Byers, 2020).

Ogah et al. (2019) reported that in predicting body weight, automatic linear modelling utilizes advanced statistical tools to analyze the correlation between different quantitative traits and body weight in the Nigerian indigenous cocks. Yakubu et al. (2021) has used ALM to study the factors influencing reproductive traits in livestock species, by analyzing data on factors such as age, weight, and environmental conditions. Yakubu et al. (2022) has used multivariate adaptive regression splines (MARS), ALM, CART, CHAID and exhaustive CHAID algorithms to predict body weight in goats. Prediction of body weight of Nigerian indigenous normal feathered chickens using data mining algorithms is scarce in the literature.

The present study was conducted to predict body weight from linear body measurements in Nigerian normal feathered local chickens using ALM, CHAID, exhaustive CHAID and CART data mining algorithms.

MATERIAL and METHOD

Ethical Approval

This study was approved by the Departmental Research Committee, Department of Animal Science, Nnamdi Azikiwe University, Awka, Nigeria (Ref: NAU/ANS/DRC/2024/001; Approval Date: 9 April 2024).

Location of the Study

The study was carried out at Amansea, which is located in Awka North LGA, Anambra State, Nigeria. Amansea is situated within the Awka capital territory and is bounded by Awka Town to the south, Mamu Rivers and Ebenebe Town to the north, Mgbakwu to the west, and Ezinato/Ubibia stream to the east. It is within the rainforest area of Nigeria and experiences an annual rainfall of 1000 – 1500 mm. Amansea has a latitude of 6°21'40'' N and a longitude of 6°51'38'' E at altitude ranging from 150 to 200 m above sea level. The area has a typical semi-tropical rainforest vegetation, characterized by freshwater swamps. It has a humid climate with an average temperature of about 30.6 ℃ (87 °F) and a rainfall between 152 and 203 cm. The area has two distinct seasons: a wet season from April to October and a dry season from November to March. The major occupation of the inhabitants of Amansea is trading, animal farming and crop farming.

Stock Selection and Data Collection

A random sample of two hundred (200) healthy normal feathered local chickens made up of 150 males and 50 females were selected at Gariki local market in Amansea, Awka North, Anambra State, Nigeria. The chickens aged between 8 and 10 months old. Data were collected on body weight (BWT) and linear body measurements (LBMs) of each individual chicken. The BWT was measured in kilogram (kg) using an analog kitchen measuring scale (model KCA) with a capacity of 5 kg and sensitivity of 1 g. The linear body measurements (LBMs) were measured in centimeter (cm) using a measuring tape. The measurements were described according to Isaac et al. (2022), Yakubu et al. (2022) and Isaac et al. (2023) as follows.

Body length (BL): The distance from the base of the comb over the neck through the body trunk to the base of the tail around the uropygial or preen gland.

Breast Width (BW): The circumstance of the breast around the deepest region, ie. from the point of depression under the wing to the sharp edge of the keel bone.

Keel length (KL): The distance from the V-joint to the end of the sternum.

Shank length (SL): The distance from the hock joint to the tarso-metatarsus pad-Digit three joint.

Thigh circumference (TC): The widest point of the thigh.

Wing length (WL): The distance between the carpo-humerus joint to the tip of the phalanges (digits) of the wing.

Statistical Analysis

Descriptive Statistics

The mean and standard error of mean of each quantitative trait for male and female normal feathered chickens were computed and tested for any significant difference using independent t test. The t statistics used for the comparison is given in expression (1), according to Isaac et al. (2023).

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S^2p}{n_1} + \frac{S^2p}{n_2}}} \quad (1)$$

where,

$\bar{X}_1$ and $\bar{X}_2$ are means of observations from male and female normal feathered chickens, respectively, n_1 and n_2 are number of observations for male and female normal feathered chickens, respectively, S^2p is the pooled variance.

The linear relationship existing between any two traits was established using Pearson Product Moment Correlation. The magnitude of such correction was computed by correction coefficient (r) as given in expression (2) according to Isaac and Obike (2022).

$$r = \frac{\sum xy - \sum x \sum y/n}{\sqrt{[\sum x^2 - (\sum x)^2/n] - [\sum y^2 - (\sum y^2)/n]}} \quad (2)$$

where r is the phenotypic correlation, x and y are any two correlating traits such as body weight and breast width and n is the number of measurements taken for males or females.

Automatic Linear Modelling for Predicting Body Weight from Linear Body Measurements

Automatic linear modelling (ALM) is an automated model selection used in identifying the most important predictor variables for inclusion in a regression model. The ALM technique was selected to evaluate the relationships between body weight and the linear body measurements (predictors). Forward stepwise method was employed as a model selection method. The ALM algorithm automatically transformed the selected predictors using logarithmic transformation and excluded outliers that could adversely affect prediction. The ALM has the capacity to handle missing data, although there were no cases of missing data recorded in the study. The model was used to predict the body weight of male, female and pooled (mixed) sexes of 200 samples of the normal feathered chickens.

The predictive equation is given in expression (3).

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_n X_n + \varepsilon \quad (3)$$

where,

Y is the predicted body weight, β_0 is the intercept, representing the expected body weight when all predictor variables are zero, β_1, β_2,…, β_n are the regression coefficients, indicating the change in body weight for a unit change in a predictor, X_1, X_2, X_3,..., X_n are the linear body measurements/predictors measured for each chicken and ϵ represents the error term.

Additionally, graphs of the observed versus predicted body weight were generated to assess the model fit while the graphs of studentized distribution of residuals (errors) were used to verify the assumptions of residual normality.

Decision tree algorithms indicated predictor traits (linear body measurements) and their contributions in prediction of body weight. The results were in form of graphic illustrations showing nodes to represent the most important and less important variables with its significant percentage. The prediction was done using Chi-square automatic interaction detection (CHAID), Exhaustive CHAID and classification and regression trees

(CART) growing methods. CHAID is a tree-based model proposed by Kass (1980) with merging, partitioning and stopping stages that recursively uses multi-way splitting procedures to form homogenous subsets. The Bonferroni adjustment was applied to control Type I error arising from multiple comparisons, thereby reducing the likelihood of false significance in the splitting process (Orhan et al., 2016; Celik et al., 2017). The Exhaustive CHAID, as a modification of CHAID algorithm, applies a more detailed merging and testing of predictor variables (Yakubu et al., 2022; Celik et al., 2017).

The three growing methods used cross validation to estimate error (Yakubu, 2012; Ali et al., 2015; Eyduran et al., 2017; Yakubu et al., 2022). The number of sample folds used for cross validation was ten. The maximum three depths were automatically set by the algorithm to 3 for CHAID and exhaustive CHAID and 5 for CART. Minimum parent node size for the three algorithms was 50, as set by the software used for the analysis.

The predictive performance of ALM was determined using expressions (4), (5), (6) and (7) (Celik et al., 2017; Eyduran et al., 2017; Yakubu et al., 2022).

1. Coefficient of determination:

$$R^2 = 1 - \frac{\sum_i^n = 1 (Y_i - \hat{y})^2}{\sum_i^n = (Y_i - \hat{y})^2} \quad (4)$$

2. Adjusted Coefficient of determination:

$$\text{Adj. } R^2 = \frac{\frac{1}{n-k-1} \sum_i^n = 1 (Y_i - \hat{y})^2}{\frac{1}{n-1} \sum_i^n = 1 (Y_i - \hat{y})^2} \quad (5)$$

3. Akaike’s Information Criterion (AICc):

$$AIC = n\log\left(\frac{RSS}{n}\right) + 2K \quad (6)$$

4. Akaike’s Information Criterion Corrected (AICc):

$$AICc = AIC + \frac{2k(k+1)}{n-k-1} \quad (7)$$

where n is the number of cases in a set, k is the number of model parameters, Y_i is the observed value of BWT, Y_{ip} is the predicted value of BWT, RSS: Residual sum of squares.

All analyses were carried out using IBM SPSS Statistics (2017) computer software.

RESULTS and DISCUSSION

RESULTS

Descriptive Statistics for Quantitative Traits of Normal Feathered Chickens

The descriptive statistics (mean ± se) for quantitative traits of the normal feathered chickens is presented in Table 1. Sex had significant (p<0.05) influence on body weight (BWT), breast width (BW) and keel length (K) only, with males recording superior means compared to their female counterparts.

Table 1. The descriptive statistics (mean ± se) of quantitative traits of normal feathered chickens

Traits	Sex		t – value	p - value
	Male (n=150)	Female (n=50)		
BWT (kg)	$1.26^{a}\pm0.02$	$1.05^{b}\pm0.03$	5.31	0.000
BL (cm)	17.12±0.1	17.38±0.16	–1.31	0.190
BW (cm)	$9.59^{a}\pm0.07$	$9.15^{b}\pm0.15$	2.94	0.004
KL (cm)	$5.27^{b}\pm0.05$	$5\ 53^{a}\pm0.10$	–2.32	0.022
SL (cm)	4.55±0.04	4.50±0.06	0.65	0.515
TC (cm)	3.13±0.03	3.11±0.04	–0.22	0.826
WL (cm)	7.52±0.06	7.50±0.09	0.19	0.853

[a, b] Means on the same row were significantly different (p<0.05).
BWT= Body weight, BL= Body length, BW= Breast width, KL= Keel length, SL= Shank length, TC= Thigh circumference, WL= Wing length.

The Correlation Coefficients Among Quantitative Traits in Normal Feathered Chicken

Correlation coefficients among quantitative traits in the normal feathered chickens are presented in Table 2. The result recorded both positive and negative correlations in each group (male, female and pooled). The magnitude of the correlation coefficient between KL and BL were larger than the correlation coefficients observed between other pairs of traits. Larger, positive and highly significant (p<0.01) correlation coefficients were observed between KL and BL (r = 0.43), BL and BWT (r = 0.46) and KL and BL (r = 0.47) in males, females and pooled sexes, respectively of the local chickens. In the three groups, the correlation coefficients were higher between KL and BL.

Table 2. Correlation coefficients among quantitative traits in normal feathered chickens

Group		BWT	BL	BW	KL	TC		SL
Male	BWT	1						
	BL	0.27**	1					
	BW	0.27**	-0.16*	1				
	KL	0.04	0.43**	-0.24**	1			
	TC	0.06	0.10	0.25**	0.09	1		
	SL	0.21**	0.19**	0.21**	0.06	0.23**	1	
	WL	0.13	-0.10	0.30**	-0.08	0.31**	0.22**	1
Female	BWT	1						
	BL	0.46**	1					
	BW	0.25**	-0.06	1				
	KL	0.20*	0.40**	-0.11	1			
	TC	0.13	0.11	0.24**	0.09	1		
	SL	0.29**	0.24**	0.29**	0.09	0.31**	1	
	WL	0.22**	-0.09	0.40**	-0.06	0.33**	0.25**	1
Pooled	BWT	1						
	BL	-0.24	1					
	BW	0.12	-0.37**	1				
	KL	-0.26	0.47**	-0.47**	1			
	TC	-0.31*	0.09	0.30*	0.13	1		
	SL	-0.03	0.02	0.02	-0.03	-0.07	1	
	WL	-0.24	-0.14	0.06	-0.14	0.23	0.12	1

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed),

BWT= Body weight, BL= Body length, BW= Breast width, KL= Keel length, SL= Shank length, TC= Thigh circumference, WL= Wing length.

Regression Coefficients and Fractional Importance of Traits in Body Weight Prediction in Normal Feathered Chickens

Table 3 presents the regression coefficients and fractional importance of various traits used to predict body weight in normal feathered chickens. The fractional importance indicates the predictor's relative contribution to reducing error in a model (Yang, 2013). A predictor with a value of 1 is regarded as the most important while one with a value of 0 is regarded as the least important. The analysis indicates that, based on positive coefficients, statistical significance, and fractional importance, BL, SL, and BW emerged as the most influential predictors for males, females, and the combined groups, respectively. However, both SL and BW exerted a negative influence on the body weight of the female chickens.

Table 3. Regression coefficients and fractional importance of traits in body weight prediction in normal feathered chickens

Group	Intercept/Predictor	Coefficient	Significance (p-value)	Importance
	Intercept	–1.52	0.000	
Male	BL	0.10	0.000	0.793
	BW	0.06	0.008	0.113
	WL	0.07	0.020	0.094
	Intercept	2.44	0.000	
Female	SL	–0.14	0.011	0.721
	BW	–0.04	0.11	0.279
	Intercept	–0.79	0.013	
Pooled	BW	0.09	0.000	0.501
	BL	0.07	0.000	0.499

BL= Body length, BW= Breast width, SL= Shank length, WL= Wing length.

Evaluation Criteria for Body Weight Prediction in Normal Feathered Chickens

The evaluation criteria for body weight prediction in normal feathered chickens are presented in Table 4. These criteria were for the most important traits in predicting body weight, which were identified in Table 3 as BL, SL and BW for males, females and pooled data, respectively. The pooled group had a lower AICc (-573.921) compared to males (-460.981) and females (-162.297).

Table 4. Evaluation criteria for body weight prediction in normal feathered chickens

Group	R^2_{adj}	AICc
Male	0.306	– 460.981
Female	0.106	–162.297
Pooled	0.163	–573.921

R^2_{adj} = Adjusted Coefficients of determination
AICc= Akaike's Information Criterion Corrected.

Observed and Predicted Body Weight Means of Normal Feathered Chickens

The observed and predicted body weights of males, females and pooled sexes of the normal feathered chickens are presented in Table 5. The lowest and highest observed and predicted BWT in males, females and pooled group were 0.70 vs 0.89, 0.68 vs 0.87 and 0.81 vs 0.82, respectively. The mean BWT and standard errors of males (1.26 ± 0.25 kg), females (1.05 ± 0.25 kg) and pooled sexes (1.021 ± 0.26 kg) were the same as those of the observed in each case. The lowest (minimum) predicted body weight was closer to its corresponding observed value in pooled group than in male and female chickens. Also, the highest (maximum) predicted body weight was far more than its corresponding maximum observed value in the pooled group compared to the male and female chickens.

Table 5. Observed and predicted body weights in normal feathered chickens using automatic linear modelling

Male	Lowest	Highest	Mean	Standard error
Observed	0.70	1.60	1.26	0.25
Predicted	0.89	1.53	1.26	0.25
Female				
Observed	0.68	1.33	1.05	0.20
Predicted	0.87	1.21	1.05	0.20
Pooled				
Observed	0.81	1.29	1.021	0.26
Predicted	0.82	1.49	1.021	0.26

Graphs of Body Weight Predictions in Normal Feathered Chickens Using Automatic Linear Modelling

The graphs of the predicted body weight from the observed values in the normal feathered chickens using automatic linear modelling are presented in Figures 1, 2 and 3 for males, females and pooled sexes, respectively. These graphs are meant to demonstrate the model fit by comparing the expected with the observed body weight. Each Figure indicates the estimated lowest and highest observed and corresponding predicted body weight values. The highest predicted body weight values were very close to the observed values. For instance, in male chickens, the highest observed body weight and its corresponding predicted values were roughly 1.60 kg and 1.53 kg.

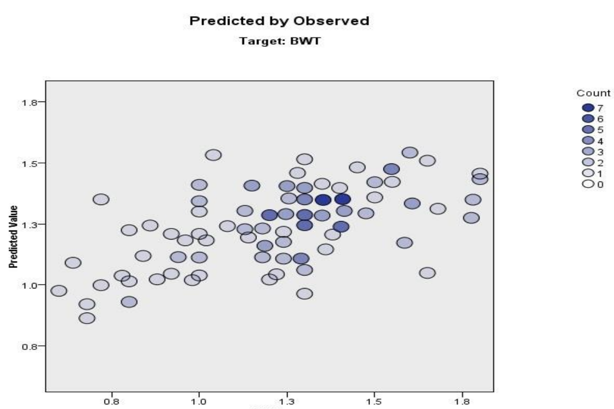

Figure 1. A graph of predicted body weight from the observed values in male normal feathered chickens

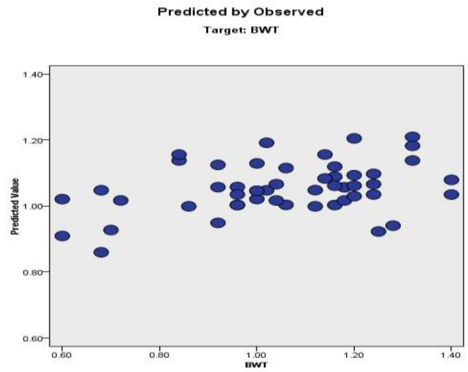

Figure 2. A graph of predicted body weight from the observed values in female normal feathered chickens

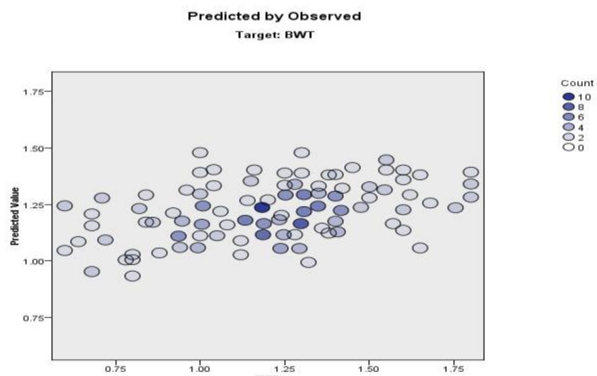

Figure 3. A graph of predicted body weight from the observed values in pooled (combined male and female) normal feathered chickens

Figures 4, 5 and 6 showed the graphs of the histogram of studentized residuals (errors) distribution compared with the normal distribution. Histograms that closely followed the normal curve indicated that the residuals in body weight were approximately normally distributed with zero mean and constant variance. The graphs showed that errors associated with the data were minimal, especially in the male and pooled groups.

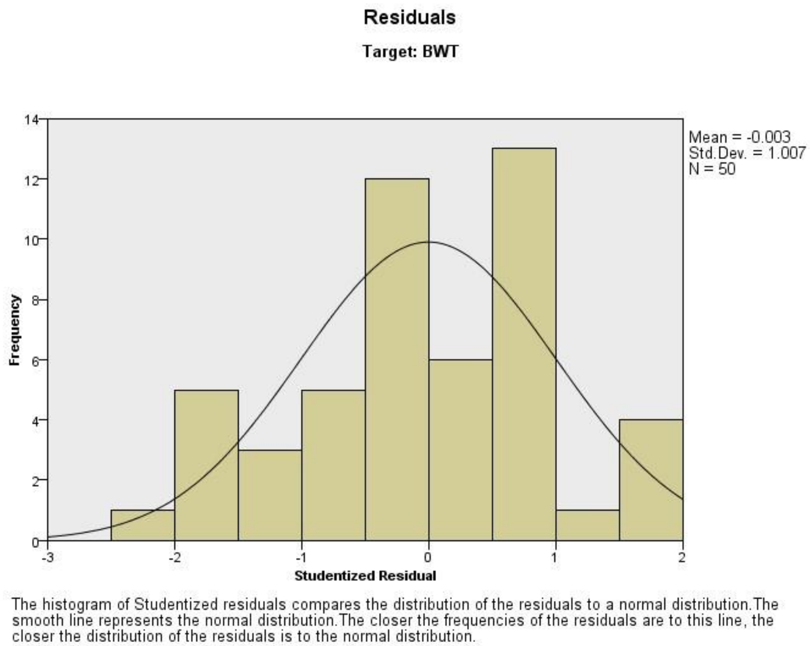

Figure 4. A graph of the histogram of studentized residual showing the distribution of errors associated with data in male normal feathered chickens

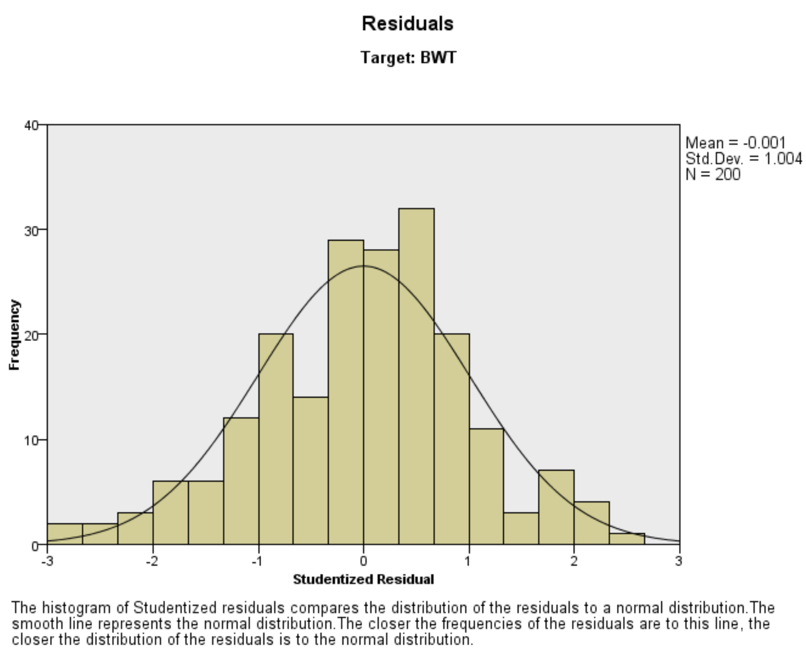

Figure 5. A graph of the histogram of studentized residual showing the distribution of errors associated with data in female normal feathered chickens

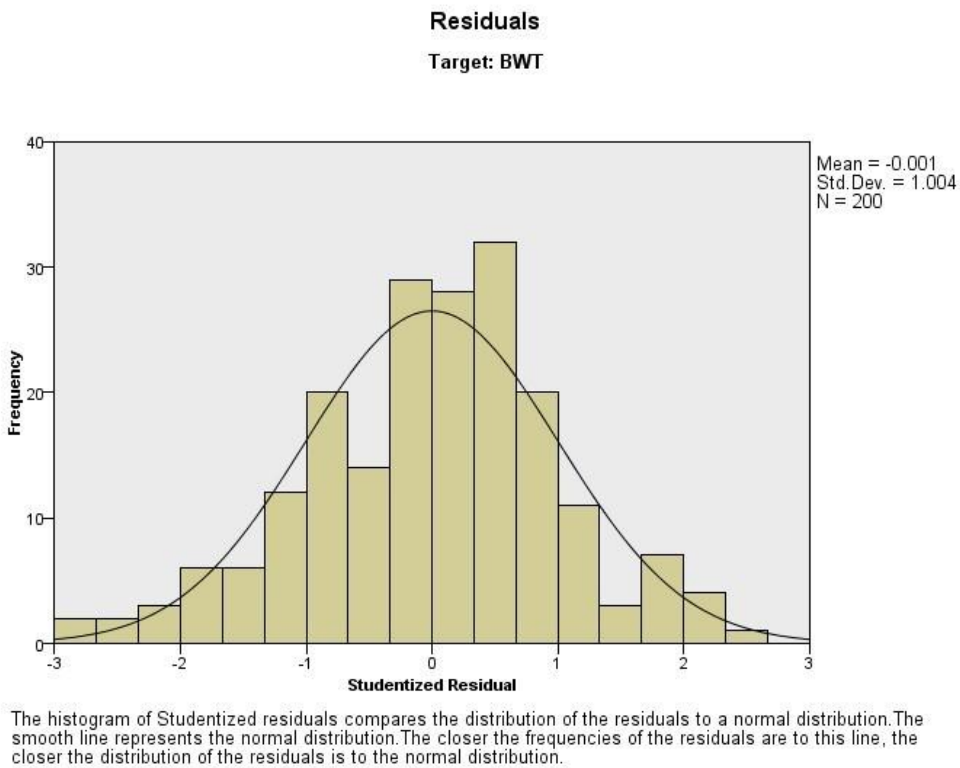

Figure 6. A graph of the histogram of studentized residual showing the distribution of errors associated with data in pooled sexes of normal feathered chickens

Prediction of Body Weight from Linear Body Measurements of Normal Feathered Local Chickens Using Decision Trees

Figure 7 shows the graphical representation of body weight using CHAID. A total of 5 nodes and 3 terminal nodes (nodes 1, 3 and 4) were generated in the prediction of BWT. BL and BW were included as important predictors of BWT. However, based on its higher position, BL was more important than BW in BWT prediction. Nodes at higher positions are judged more important than those at lower position (Yang et al., 2023) because they represent the most significant splits in the data and explain the largest portion of the variation in the target variable. The Figure indicates that 70.50% body weight of the chickens (sexes combined) could be predicted by selecting chickens (sexes combined) with BL > 16.80 cm. The graph also reveals that by selecting normal feathered chickens with BW > 9.000 cm, highest BWT of 1.323 kg can be realized.

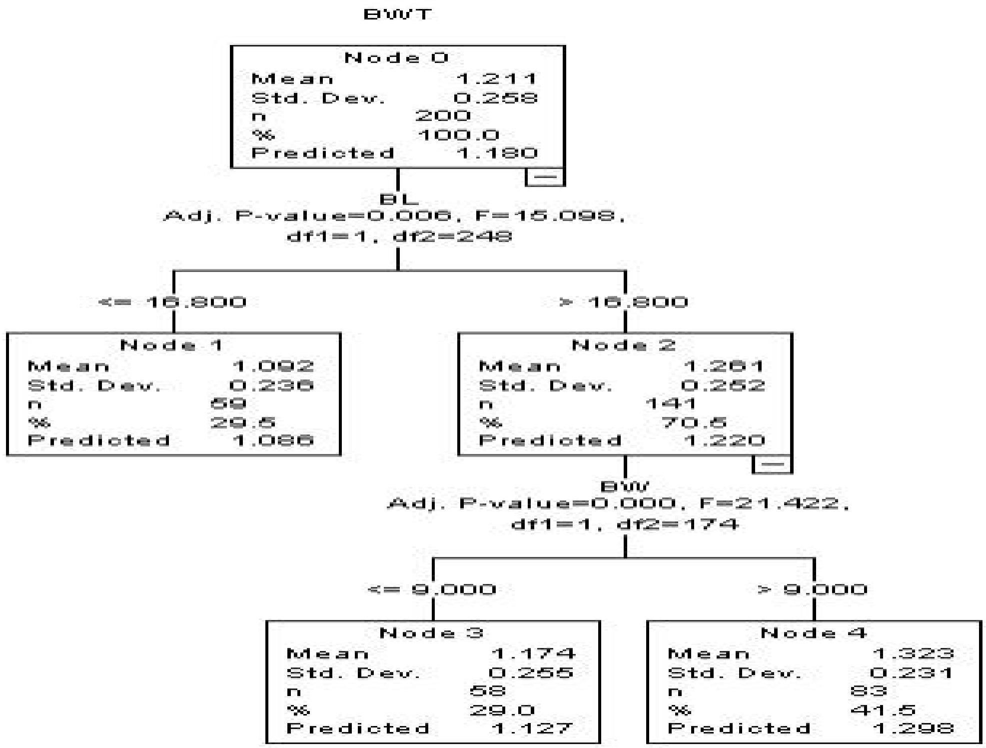

Figure 7. A graphical representation of body weight prediction using CHAID

In the Exhaustive CHAID analysis, 3 terminal nodes (nodes 1, 3 and 4) were also generated (Fig. 8). BL was also the superior trait. The results of the exhaustive CHAID were exactly the same with those of the CHAID.

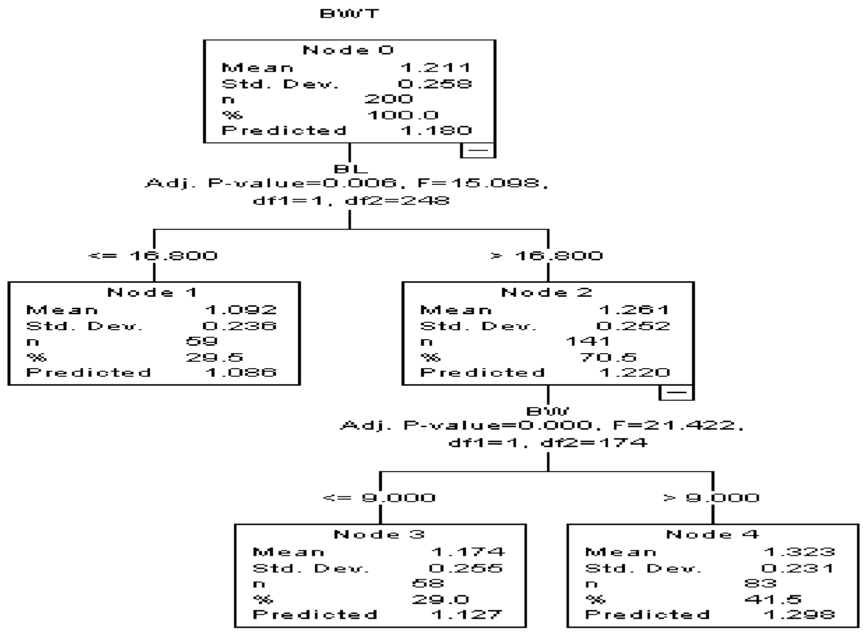

Figure 8. A graphical representation of body weight prediction using Exhaustive CHAID in normal feathered chickens.

In CART analysis, only 1 node was generated (Fig. 9). There was 0 terminal node and hence only a single root node was produced, and no predictor was selected. With this CART, the predicted mean body weight was 1.18 kg.

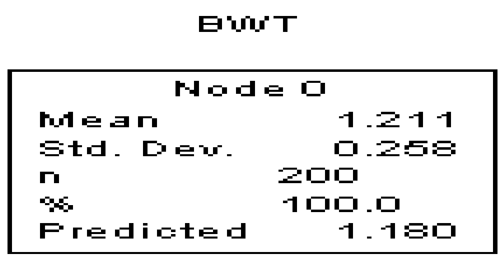

Figure 9. A graphical representation of body weight prediction using CART

DISCUSSION

The means of the quantitative traits obtained for the chickens are within the normal ranges reported for local chickens (Dankoli et al., 2021), thus confirming the fact that the chickens used in this study are of indigenous origin which are normally characterized by poor growth due to lack of genetic improvement, poor nutrition and harsh climatic conditions to which they are constantly exposed (Isaac and Ezejesi, 2023). The superior significant body weight (BWT), breast width (BW) and keel length (KL) means recorded in the male against the female chickens is in agreement with the reports of the previous authors (Faith et al., 2018), who reported higher BWT and some morphometric traits in favour of the male chickens. This higher performance is attributed to the existence of sexual dimorphism in poultry, a phenomenon whereby male chickens manifest greater physical and behavioural traits as the chickens reach sexual maturity (Siegel and Honaker, 2025).

The observed variations in direction and magnitude of the correlation coefficients among the quantitative traits in all groups agree with the report of Nweke-Okorocha, (2022), indicating that some traits may result in increase or decrease in the growth of other ones. In case of negative correlation coefficients, Tongsiri (2014) suggested that traits which exhibit such effect should not be included in selection index for improvement but be improved individually to circumvent the antagonistic relationship between them. The larger, positive and highly significant ($p<0.01$) correlation coefficients observed between KL and BL, BL and BWT and KL and BL in males, females and pooled sexes, respectively of the local chickens compared to other traits, is an indication that by selecting one trait the other trait will be improved through correlated response, as supported by Isaac and Obike (2020). Such selection has an advantage of shortening generation interval and improving annual genetic gain (Nwaogwugwu et al., 2018). BL and BW which had the highest correlation coefficients with BWT were also the best predictors of BWT, indicating that these traits could be very crucial in improvement of BWT of the respective groups of the normal feathered chickens. Other larger, positive and significant correlation coefficients obtained between WL and TC in males, WL and BW in females and between TC and BW in the pooled sexes imply that these traits are important determinants of body size and conformation in local chickens.

The high fractional importance of BL in males indicates the trait's high contribution in reducing error in BWT prediction model. In other words, BWT can be predicted using BL with high accuracy in cocks. Similar results have been reported by Yakubu et al. (2019) in ducks. The negative relationship observed between SL and BWT in the female chickens indicates that an increase in SL can result in decrease in body weight. This relationship is essential in selection of egg-type chickens where body weight reduction is considered an important breeding objective. Debes et al. (2015) reported an association between long SL and a good laying ability in chickens, thus, supporting the importance of SL as an indicator of a good laying breed. The BL and BW which made nearly equal contributions to BWT prediction in the pooled group, support Taylor's (2021) results, which emphasized on combining traits for mixed-sex populations.

The outstanding performance of pooled data in body weight prediction as indicated by the lowest value of its AICc in the evaluation criteria using ALM has been reported

(Akaike, 1974; Jensen et al., 2020; Yakubu et al., 2021; Roberts and Wilson 2022). However, the failure to provide other model evaluation criteria such as root mean-square error (RMSE), mean absolute percentage error (MAPE), mean absolute deviation (MAD) and global relative approximation error (RAE) as in other studies (Celik and Yilmaz, 2017; Celik et al., 2017), to compare the predictive abilities of ALM and the decision trees (CHAID, exhaustive CHAID and CHART) is a limitation of the software used for the analysis of data in this study.

The minimum predicted BWT and its observed value in the pooled sexes which were closely related agrees with the findings of Yakubu et al. (2022) and indicates that ALM is efficient in predicting body weight of chickens.

The identification of BL as a predominant predictor of BWT in CHAID and Exhaustive CHAID algorithms, based on its higher position in the decision tree (Lin and Fan, 2019), is affirmed by previous studies (Kass, 1980; Eyduran et al., 2017; Yakubu et al., 2022).

The heaviest BWT predicted with specific level of BW (>9.000 cm) using the terminal nodes in CHAID and Exhaustive CHAID graphics, suggests a strong correlation between BW and BWT in chickens (Behiry et al., 2019). It further suggests that CHAID or Exhaustive CHAID algorithm is preferred when specific level of a trait is required for selection to get a desired improvement in BWT in chickens. Similarly, chest circumference has been identified as an important trait for BWT prediction in goats using exhaustive CHAID algorithm (Yakubu et al., 2020).

The CART result which showed that no predictor was selected contradicted the findings of Yakubu et al. (2022) in goats, pointing out that species of animals and perhaps the predictors used can influence prediction results. The absence of predictors in the CART model likely reflects its sensitivity to weak or highly correlated variables, its known instability in small datasets (Banerjee et al., 2019) and a constant target variable (King & Restock, 2014; Lin & Fan, 2019; Huynh-Cam, 2021). This instability implies that even minor changes in training data can alter the tree structure.

Comparing the highest predicted body weight of pooled sexes using ALM (1.49 kg) and CHAID or exhaustive CHAID (1.32 kg), it was noted that ALM algorithm had better predicted ability. However, the two decision trees (CHAID and Exhaustive CHAID) specified the breast width (>9.00 cm) that could predict the highest body weight (1.32 kg) which is lacking in ALM. This may make CHAID and exhaustive CHAID more preferable in body weight prediction.

CONCLUSION and RECOMMENDATIONS

This study confirms that body weight in normal feathered chickens can be effectively predicted using the linear body measurements. Sexual dimorphism was evident, with males showing significantly higher body weight and breast width than females. Breast width emerged as the most reliable predictor of body weight in mixed-sex flocks, especially using Automatic Linear Modelling (ALM). Body length was the most critical variable identified by CHAID and Exhaustive CHAID algorithms, particularly when breast width exceeded 9.00 cm. The CART model was less effective, failing to identify any significant predictors.

In a breeding and selection programme aimed at improving body weight using the morphometric traits in chicken, breast width and body length should be prioritized. ALM is recommended for general prediction in a mixed population due to its superior accuracy while CHAID-based models should be applied when targeting specific traits like body length for focused improvement. Sex-specific selection strategies are needed to account for observed differences between males and females.

ACKNOWLEDGEMENT

The authors express their profound gratitude to the local chicken dealers who provided their chickens for this study.

Conflict of interest statement

The authors have declared that there are no competing interests.

Author Contributions

UCI conceptualized the title and design of the study. Data collection was carried out by CUI, while data analysis was done by UCI. The experiment was conducted by CUE under the supervision of UCI. The manuscript was jointly written by CUE and UCI and critically reviewed and approved by UCI.

REFERENCES

Akaike H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6): 716–723. https://doi.org/10.1109/TAC.1974.1100705

Adeyinka IA, Imumorin IG, Olowofeso O., 2017. Body weight prediction in a Nigerian indigenous chicken population using linear body measurements. Iranian Journal of Applied Animal Science, 7(3): 451–457.

Ali M, Eyduran E, Tariq MM, Tirink C, Abbas F, Bajwa MA, Baloch MH, Nizamani AH, Waheed A, Awan MA, Shah SH, Ahmad Z, Jan S., 2015. Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post-weaning period from some biometrical characteristics in Harnai sheep. Pakistan Journal of Zoology, 47: 1579–1585.

Banerjee M, Reynolds E, Andersson HB., 2019. Tree-based analysis: practical approaches to create clinical decision-making tools. Cardiovascular Quality & Outcomes, 12(5): e004879. https://10.1161/CIRCOUTCOMES.118.004879

Breiman L, Friedman JH, Olshen RA, Stone CJ., 1984. Classification and Regression Trees (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315139470

Celik S, Eyduran E, Karadas K, Tariq MM., 2017. Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. Revista Brasileira de Zootecnia, 46(11): 863–872.

Celik S, Yilmaz O., 2017. Comparison of different data mining algorithms for prediction of body weight from several morphological measurements in dogs. The Journal of Plant and Animal Sciences, 27(1): 57–64.

Dankoli ZA, Muhammad IR, Ogah DM., 2021. Use of breast length as a quantitative trait in characterisation of indigenous chickens in semi-arid Nigeria. Nigerian Journal of Animal Production, 48(6): 1–12. https://doi.org/10.51791/njap.v48i6.3271

Debes Hasaan S, El-Rehen H, Eshera A, Bashandy M., 2015. Relationship of shank length with some productive and physiological parameters of Matrouh chicken strains. Journal of Animal and Poultry Production, 6(12): 741–753. https://doi.org/10.21608/jappmu.2015.52959

Eyduran E, Zaborski D, Waheed A, Celik S, Karadas K, Grzesiak W., 2017. Comparison of the predictive capabilities of several data mining algorithms and multiple linear regression in the prediction of body weight using body measurements in the indigenous Beetal goat of Pakistan. Pakistan Journal of Zoology, 49: 273–282. http://dx.doi.org/10.17582/journal.pjz/2017.49.1.273.282

Faith EA, Yakubu A, Agade YI, Abimiku HK, Mohammed J., 2018. Sexual dimorphisms on body weight, morphometric and haematological parameters of indigenous chicken reared in Lafia. Direct Research Journal of Veterinary Medicine and Animal Science, 3(4): 36–41. https://doi.org/10.26765/DRJVMAS2028.6392

Friedman J, Hastie T, Tibshirani R., 2010. Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33(1): 1–22.

Genç S, Mendes M., 2021. Evaluating performance and determining optimum sample size for regression tree and automatic linear modeling. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 73: 1391–1402. https://doi:10.1590/1678-4162-12413

Hothorn T, Hornik K, Zeileis A., 2006. Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics, 15(3): 651–674. https://doi.org/10.1198/106186006X133933

Huynh-Cam TT, Chen LS, Le H., 2021. Using decision trees and random forest algorithms to predict and determine factors contributing to first-year university students' learning performance. Algorithms, 14(11): 318. https://doi.org/10.3390/a14110318

IBM SPSS Statistics., 2017. International Business Machines Statistical Package for Social Sciences. IBM Corp., Version 25.0, Armonk, NY.

Isaac UC, Adeolu AI., 2023. Linear models for predicting body weight in crossbred chickens. International Journal of Research and Innovation in Applied Science, 8(4): 62–70. https://doi.org/10.51584/IJRIAS.2023.8407

Isaac UC, Ezejesi HC., 2023. Genotype impact on body weight and linear body measurements of main and reciprocal crosses of Isa Brown and local chickens. Nigerian Journal of Animal Production, 50(4): 1–9. https://doi.org/10.51791/njap.v50i4.3355

Isaac UC, Igbokwe CA, Okafor NJ, Aniemena CF, Albert BC, Nwachukwu BC., 2022. Repeatability of growth traits in three strains of broiler chicken. Malaysian Animal Husbandry Journal, 2(1): 29–33. https://doi.org/10.26480/mahj.01.2022.29.33

Isaac UC, Okeke BU, Nwankwo CB., 2023. Growth performance of Panda White and Cinnamon Brown strains of Japanese quail. In: Proceedings of the 48th Annual Conference of the Nigerian Society for Animal Production (NSAP), Federal University of Dutsin-Ma, 18–22 June 2023. pp. 1339–1343.

Isaac UC, Obike MO., 2020. Phenotypic correlations between body weight and egg production traits of local chicken genotypes in humid tropical rain forest of Umudike. International Journal of Agriculture and Biosciences, 9(3): 128–133.

Jensen P, Persson ME, Larsson L., 2020. Integrated models for predicting phenotypic traits in poultry. Journal of Animal Science, 98(12): 1–10.

Kass GV., 1980. An exploratory technique for investigating large quantities of categorical data. Applied Statistics, 20(2): 119–127. https://doi.org/10.2307/2986296

Lee BK, Lessler J, Stuart EA., 2010. Improving propensity score weighting using machine learning. Statistics in Medicine, 29(3): 337–346. https://doi.org/10.1002/sim.3782

Lin CL, Fan CL., 2019. Evaluation of CART, CHAID and QUEST algorithms: a case of construction defects in Taiwan. Journal of Asian Architecture and Building Engineering, 18(6): 539–553. https://doi.org/10.1080/13467581.2019.1696203

King MW, Restock PA., 2014. Data mining in psychological treatment research: a primer on classification and regression trees. Journal of Consulting and Clinical Psychology, 82(5): 895–905. https://doi.org/10.1037/a0035886

Matzavela V, Alepis E., 2021. Decision trees learning through a predictive model for student academic performance in intelligent m-learning environments. Computers & Education: Artificial Intelligence, 2(6): 100035. https://doi.org/10.1016/j.caeai.2021.100035

Mushkudiani NA, Hukkelhoven CWPM, Hernández AV, Murray GD, Choi SC, Maas AIR, Steyedberg EW., 2008. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. Journal of Clinical Epidemiology, 61(4): 331–343. https://doi.org/10.1016/j.jclinepi.2007.06.011

Nwakpu OF, Nosike RJ, Isaac UC, Nwose RN, Nwachukwu CC, Egbuedo UC, Nwaokoro CC, Nathaniel J, Akinsola KL, Onunkwo DN., 2020. Regression equation for selected growth traits prediction in Nigerian indigenous cattle (NIC) genotype. Nigerian Journal of Animal Science, 22(2): 1–6.

Nwaogwugwu UC, Udoh UH, Obike OM, Ahamefule GO., 2018. Body weight prediction and phenotypic correlation with linear body measurements in two strains of Japanese quail. International Journal of Agriculture and Rural Development, 21(1): 3547–3553.

Nweke-Okorocha OG, Obiageri G, Chineke CA, Brilliant OA., 2022. Phenotypic and genetic correlations between body weight and linear body measurements of Nigerian local and improved chickens. International Journal of Innovative Science and Research Technology, 7(2): 423–426.

Ogah DM, Abimiku H, Ejembi C, Okpeku M, De Donato M, Peters SO, Imumorin IG., 2019. Evaluation of body weight prediction models in Nigerian indigenous cocks using automatic linear modeling. Poultry Science, 98(12): 6414–6420.

Ojedapo LO, Adebambo OA, Ilori LJ, Ikeobi CON., 2019. Phenotypic characterization of two Nigerian indigenous chicken populations using body weight and linear body measurements. Poultry Science, 98(5): 2083–2089.

Orhan H, Eyduran E, Tatliye A, Saygici H., 2016. Prediction of egg weight from egg quality characteristics via ridge regression and regression tree methods. Revista Brasileira de Zootecnia, 45(7): 380–385. https://doi.org/10.1590/S1806-92902016000700004

Oshima TC, Dell-Ross T., 2016. All possible regressions using IBM SPSS: A practitioner's guide to automatic linear modeling. Georgia Educational Research Association Conference. Georgia Southern University. Available at: https://digitalcommons.georgiasouthern.edu/gera/2016

Roberts C, Wilson A., 2022. Sex-specific variations in predictive modeling for livestock traits. Journal of Poultry Science, 101(5): 2341–2350.

Seok J, Kang YS., 2025. Mutual information between discrete variables with many categories using recursive adaptive partitioning. Scientific Reports, 5: 10981. https://doi.org/10.1038/srep10981

Siegel PB, Honaker CF., 2025. Sexual dimorphism for juvenile body weight in chickens divergently selected for 8-week body weight. Frontiers in Physiology, 16: PMCID: PMC11788401, PMID: 39902469. https://doi.org/10.3389/fphys.2024.1534334

Taylor JP., 2021. Advances in poultry weight prediction: A comparative analysis of data mining algorithms. Journal of Poultry Science, 100(4): 1209–1220.

Tongsiri S, Jeyaruban MG, Van Der Werf JH, Thummabood S., 2014. Genetic parameters for production traits of Rhode Island Red and White Plymouth Rock breeds selected under tropical condition in Thailand. In: Proceedings of 10th World Congress of Genetics Applied to Livestock Production, pp. 854–856.

Wray CM, Byers AL., 2020. Methodological progress note: Classification and regression tree analysis. Journal of Hospital Medicine, 15(9): 549–551. https://doi.org/10.12788/jhm.3366

Yakubu A., 2012. Application of regression tree methodology in predicting the body weight of Uda sheep. Scientific Papers: Animal Science and Biotechnologies, 45: 484–490.

Yakubu A, Abdullah AR, Raji AO, Oseni SO., 2021. Multivariate principal component analysis of the morphometric traits of some cattle breeds in Nigeria. Journal of Animal & Plant Sciences, 31(3): 1018–1078. http://doi.org/10.36899/JAPS.2021.3.0258

Yakubu A, Dahloum AL, Shoyombo AJ, Yahaya UM., 2019. Modelling hatchability and mortality in muscovy ducks using automatic linear modelling and artificial neural network. Journal of Indonesian Tropical Animal Agriculture, 44(1): 65–76. https://doi.org/10.14710/jitaa.44.1.65-76

Yakubu A, Eyduran E, Senol C, Juliana OI., 2022. Use of automatic linear modeling, multivariate adaptive regression splines and decision trees in body weight prediction in goats. Genetika, 54(3): 1429–1445. https://doi.org/10.2298/GENSR2203429Y

Yang H., 2013. The case for being automatic: introducing the automatic linear modeling (linear) procedures in SPSS Statistics. Multiple Linear Regression Viewpoints, 39(2): 27–37.

Yang Y, Yi D, Deng Y. Sun, Y., 2023. Performance analysis of the CHAID algorithm for accuracy. Mathematics, 11(11): 2558. https://doi.org/10.3390/math11112558