Use Of Automatic Linear Modeling And Decision Trees In Body Weight Prediction In Normal Feathered Chickens

Isaac, U. C. and Ernest, C. U.*

[1] Department of Animal Science, Faculty of Agriculture, Nnamdi Azikiwe University, Awka

KEYWORDS

Body weight,
Data mining algorithms,
Local chicken.
Prediction,

***CORRESPONDING AUTHOR**

uc.isaac@unizik.edu.ng
+2348064368510

ABSTRACT

The aim of the study was to predict body weight (BWT) from linear body measurements of 200 mature normal feathered chickens, comprising 150 males and 50 females randomly selected in Amansea, Awka North LGA, Anambra State, Nigeria. Automatic Linear Modeling (ALM) and decision tree algorithms involving Chi-square Automatic Interaction Detector (CHAID), Exhaustive CHAID, and Classification and Regression Trees (CART) were used in the prediction. Adjusted coefficient of determination (R^2adj.) and Alkaike's Information Criterion corrected (AICc) were used to evaluate the predictive ability of ALM in males, females and pooled sexes. Males were significantly ($p<0.05$) superior to females in BWT (1.26±0.02 kg vs 1.0.5±0.03 kg) and many linear parameters. The ALM indicated that body length (BL), shank length (SL) and breast width (BW) had the highest significant ($p<0.001$) fractional importance in BWT prediction in males (0.793), females (0.721) and pooled sexes (0.501), respectively. Highest predictive ability of ALM was achieved in pooled sexes with the smallest AICc (-0.572.92), and predicted BWT value of 1.49 kg. CHAID and Exhaustive CHAID revealed that highest BWT of 1.220 kg could be predicted with BW>9.000 cm. With CART, no variable was identified as important in BWT prediction. The study revealed ALM as the best model for predicting BWT using BW of pooled sexes in normal feathered local chicken.

INTRODUCTION

The productivity of the local chickens has remained low due to limited improvement of their genetic constitution and environmental factors affecting them (Ojedapo *et al*., 2019; Isaac and Ezejesi, 2023). The body weight is an important economic trait, which determines the market price of the poultry and is useful in management decisions in poultry production (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 such as the breast width and shank length are not biased by the gut fill (Isaac and and Adeolu, 2023), and for this reason, prediction of body weight from them is preferred.

Conventional multiple linear and stepwise multiple linear regressions have been employed in prediction experiments (Nwakpu *et al*., 2020; Isaac and Adeolu, 2023). However, these methods have not been found very efficient due to their inability to handle interdependency or multicolinearity exiting among the linear body measurements or predictors which can lead to biased estimates and reduced predictive accuracy (Adeyinka *et al*., 2017; Xi *et al*. 2024).Recently, modern regression algorithms which are more robust and efficient in prediction have been in use. Among these are the automatic linear modeling (ALM) and decision trees. Automatic linear modeling (ALM) has emerged as a powerful screening analytical tool in prediction experiments, automating the process of selecting the most suitable subset of predictors, which is particularly crucial when dealing with a large number of predictors (Yang, 2013; Oshima and Dell-Ross, 2016; Genç and Mendeş, 2021).

Decision trees present prediction results in tree graphics which have nodes and terminal nodes that represent points of decision. Each decision tree result can be produced through chi-square automatic interaction detection (CHAID), exhaustive CHAID and classification and regression trees (CART) growing algorithms. These algorithms can construct 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). One important feature of the decision trees is their ability to handle both categorical and continuous variables (Wray and Byers, 2020; Razi and Athappilly, 2005).

Assan (2015) had employed ALM to accurately predict body weight in Nigerian indigenous cocks. Yakubu *et al*. (2020) and Yakubu *et al.* (2021) had used ALM and decision trees algorithms in prediction studies in livestock. The aim of the study was to predict body weight from linear body measurements using automatic linear modelling (ALM) and decision trees algorithms in normal feathered chickens.

MATERIALS AND METHODS

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 of approximately 150 – 200 m. The area has a typical semi-tropical rainforest vegetation, characterized by freshwater swamps. It has a humid climate with an average daily temperature of about 30.6 ℃ and a rainfall between 152 cm and 203 cm. The area has two distinct seasons: the wet season ranging 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.

Data Collection and Measurements of Quantitative Traits

Data were collected from randomly selected 200 normal feathered chickens (150 males and 50 females) reared extensively in Amansea, Awka North in Anambra State, Nigeria. Body weight and linear body measurements (LBMs) were measured on each individual chicken. The body weight was measured in kilogram (kg) using an analog kitchen measuring scale (model KCA) with a capacity of 5 kg and sensitivity of 2 g. The linear body measurements (LBMs), namely body length (BL), body length (BL), Breast Width (BW), Keel length (KL), Shank length (SL), Thigh circumference (TC) and Wing length (WL) were measured in centimeter (cm) using a measuring tape described according to Isaac *et al.* (2022), Yakubu *et al.*, (2022) and Isaac *et al.* (2023).

STATISTICAL ANALYTICAL PROCEDURES

The mean and standard error of mean of each quantitative trait for male and female normal feathered chickens were computed and significant difference tested using independent t test. The formula for the t statistic used for the comparison was according to Isaac *et al.* (2023). Automatic linear modelling and decision trees were used to predict body weight from the LBMs of male, emale and pooled (mixed) sexes of 200 samples of the normal feathered chickens. The decision trees analyses were done using Chi-square automatic interaction detection (CHAID), exhaustive CHAID and classification and regression trees (CART) growing methods according to previous authors (Eyduran *et al*., 2016; Celik *et al*., 2017; Yakubu *et al*., 2022. The three growing methods used cross validation to estimate error (Ali *et al*., 2015; Eyduran *et al*., 2017; Yakubu *et al*., 2022).The predictive performance of each of algorithm was determined previous authors (Yakubu *et al*., 2022; Eyduran *et al*., 2017; Celik *et al*., 2017). All analyses were carried out using IBM SPSS Statistics (2017) computer software.

RESULTS AND DISCUSSION

Effect of sex on quantitative traits of normal feathered chickens

The effect of sex (mean ±se) on 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 for BWT and BW compared to their female counterparts.

Table 1: Effects of sex (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	7.5±0.09	0.19	0.853

$^{a,\,b}$ Means on the same row were significantly ($p<0.05$) different.

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

The regression coefficients and fractional importance of traits in body weight prediction in normal feathered chickens is presented in Table 2. Based on the positive coefficients, significance and fractional importance, BL, SL and BW were the most influential traits in predicting body weight in males, females and pooled sexes, respectively. The SL and BW however, had negative impact on body weight of the female chickens.

Table 2: 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	0.110	0.279
	Intercept	–0.79	0.013	
Pooled	BW	0.09	0.000	0.501
	BL	0.07	0.000	0.499

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 3.These criteria were for the most important traits in predicting body weight, which were shown in Table 2 as BL, SL and BW for males, females and pooled data, respectively. The result of Table 3indicated that although, the adjusted coefficient of determination contributed by BL in the male chickens (0.306) were higher than the other two groups but the AICc (-573.921) of the pooled sexes contributed by BW were comparatively smaller and better than the male and female chickens.

Table 3: 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 weight of males, females and pooled sexes of the normal feathered chickens are presented in Table 4. The lowest and highest observed and predicted BWT were: male (0.70 vs 0.89), female (0.68 vs 0.87) and pooled (0.81 vs 0.82). 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 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 sexes compared to male and female chickens.

Table 4: Observed and predicted body weights in normal feathered chickens using automatic linear modeling.

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

Prediction of body weight from linear body measurements of normal feathered local chickens using decision trees

Figure 1is the graphical representation of body weight using CHAID algorithm. BL was more important (i.e. nodes at higher positions are more important than those at lower position) in BWT prediction. A total of 3 terminal nodes (nodes 1, 3 and 4) were generated in the prediction of BW. Among the terminal nodes, chickens with BW>9.000 cm had the highest BWT of 1.320 kg [variance: $(0.231)^2 = 0.053$] in node 4 followed by node 3 (1.174 kg). However, chickens with BW<=9.000 cm had the smallest body weight of 1.092 kg

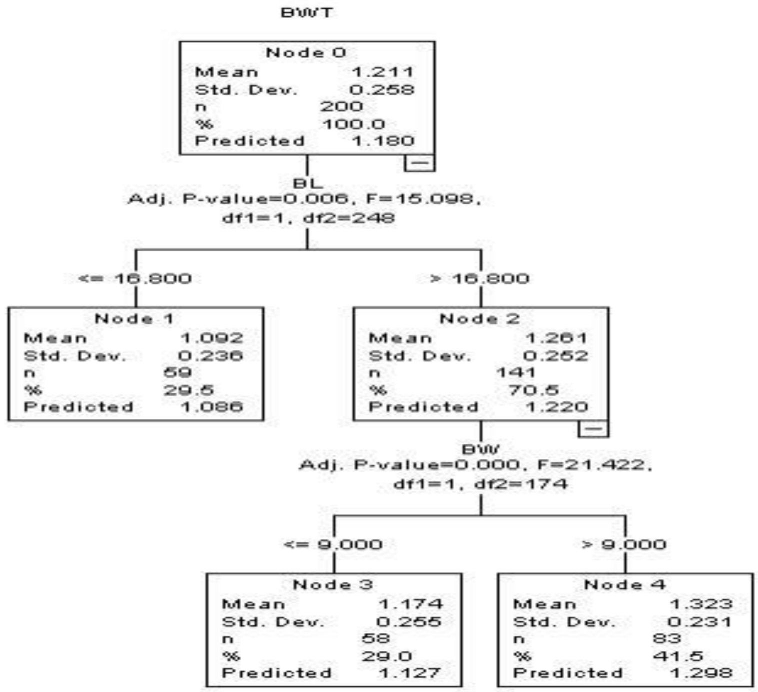


Fig. 1 A graphical representation of body weight prediction using CHAID algorithm

The results of Exhaustive CHAID were the same as those of CHAID. Three (3) terminal nodes (nodes 1, 3 and 4) were generated (Fig. 2). Body length was also the superior trait.

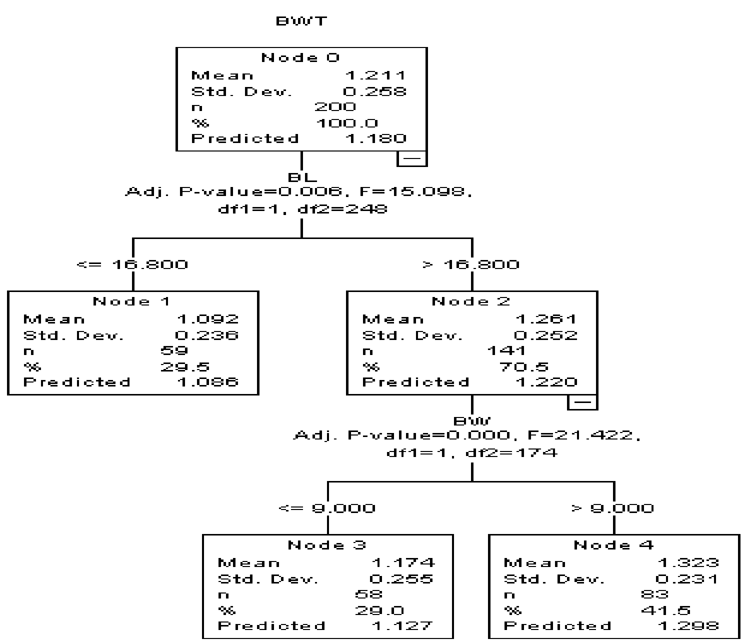
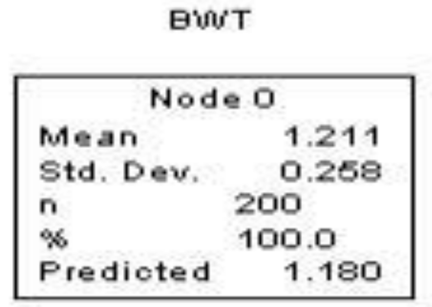


Fig. 2. A graphical representation of body weight prediction using Exhaustive CHAID algorithm

In CART analysis, only the root node (node 0) with no terminal node was generated (Fig.3). None of the linear body measurements was important in predicting BWT as there was no terminal node produced. With this CART, the predicted mean body weight was 1.18 kg and accuracy of prediction was 100%.

Fig. 3. A graphical representation of body weight prediction using CART algorithm



The means of the quantitative traits obtained for the chickens in Table 1 are within the normal ranges reported for local chickens (Dankoli *et al*., 2021), thus confirming that the chickens are of indigenous origin, which are normally characterized by poor growth, low productivity and economic value due to little or no genetic improvement and harsh environmental conditions to which they are exposed (Isaac and Ezejesi, 2023).The superior significant body weight (BWT) and breast width (BW)means recorded in the male against the female chickens is in agreement with the reports of previous authors (Faith *et al*., 2018), who reported higher body weight and some morphometric traits in favour of male chickens and goats This favourable superior growth of the male chickens is attributable 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 BL and BW presented as most important traits in body weight prediction in males and pooled sexes by ALM agrees with the findings of Isaac and Adeolu (2023), who reported that body girth, body length and body width were among the best partial predictors of body weight in crossbred local chickens. These results corroborate with the findings of Smith and Jones (2020) on the role of skeletal measurements in body weight prediction.

The greater predictive accuracy obtained from the pooled data supports the work of Jensen *et al*. (2020), who reported that pooling data across sexes can yield more comprehensive models in growth studies by encompassing a broader range of predictor variations, thus enhancing predictive power. These results are further supported by the findings of Roberts and Wilson (2022), who highlighted models' general accuracy in estimating means while noting challenges in predicting extremes, advocating for pooled datasets to improve predictive performance across a wider range of values.

CHAID and Exhaustive CHAID both identified the BL as a more important trait for predicting body weight. These findings validate the hierarchical importance of BL as reported by Kass (1980) and Eyduran *et al*. (2017).However, the heaviest body weight predicted with specific level of BW (>9.000 cm) not only indicates the BW as an indicator of large body size in chickens (Behiry, *et al*. 2019), but suggests that CHAID or Exhaustive CHAID may be a better algorithm when specific level of a trait is needed for selection for heavy body weight in chickens. The identification of chest circumference in goats by exhaustive CHAID algorithm as an important trait for body weight prediction as reported by Yakubu *et al*. (2020), confirms breast width as an important predictor of body weight in chicken. The result of CART analysis with no node or terminal node contradicts the findings of Yakubu *et al*. (2020). This suggests that results of body weight prediction using data mining algorithms may defer according species of animals involved.

The findings emphasize the significance of ALM, CHAID, Exhaustive CHAID, and CART algorithms for body weight prediction in poultry.

CONCLUSION

The study revealed the automatic linear modelling as the best algorithm for predicting body weight using the breast width from the pooled data. However, using CHAID or Exhaustive CHAID algorithm body length will be preferred, because it requires specific value that can predict a particular body weight. The study is useful in selection of chickens for improvement of body size using these data mining algorithms.

REFERENCES

Adeyinka, I. A., Imumorin, I. G., and 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.*

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

Ali, M., E. Eyduran., M.M. Tariq., C. Tirink., F. Abbas., M.A. Bajwa *et al* (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. *Pak. J. Zool, 47:1579-1585.*

Assan, N. (2015). Methodology and factors influencing the association of body weight performance parameters with linear body measurements assessment in poultry. *Scientific Journal of Pure and Applied Sciences 4(10):200- 210,* https://:dio.org/10.14196/sjpas.v4i10.1950.

Behiry, F. M., Hassanin, M.N.F., Ali Abd El- Az., E.M., El-Kamash, M,Bahnas, M.(2019).Using some body measurements as predictors of live body weight and carcass traits in four broiler strains .*Egyptian Poultry Science*, 39 (4): 835-849.

Celik, S., Eyduran, E., Karadas, K., Tariq, M.M. (2017). Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. Rev. Bras. Zootech. 46(11): 863–872.

Dankoli, Z. A., Muhammad, I. R. Ogah, D. M. (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.

Eyduran, E., Keskin, I., Erturk, Y.E., Dag, B., Tatliyer, A., Tirink, C., Aksahan, R., Tariq, M. M., (2016). Prediction of leece weight from wool characteristics of sheep using regression tree method (CHAID Algorithm). *Pakistan Journal of Zoology, 48: 957-960.*

Eyduran, E., Zaborski D., Waheed A., Celik S., Karadas K., and 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

Genç, S and M. Mendes (2021). Evaluating performance and determining optimum sample size for regression tree and automatic linear modeling. *Arq. Bras. Med. Vet. Zootec. 73, 1391−1402.* https://doi:10.1590/1678-4162-12413.

IBM SPSS Statistics (2017). IBM Corp. Release 2017. IBM SPSS Statistics for windows, version 25.0. Armonk. NY: IBM Corp.

Isaac U. C. and Adeolu A. I., (2022). Linear models for Predicting body weight in crossbred chickens. *International Journal of Research and Innovation in Applied Sci 8(4):2454-6194.* https://doi.org/10.51584/IJRIAS.2023.8407

Isaac, U. C. and Ezejesi, H. C. (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, U. C., Igbokwe, C. A., Okafor, N. J., Aniemena, C. F., Albert, B. C. and B. C. Nwachukwu (2022). Repeatability of growth traits in three strains of broiler chicken. *Malaysian Animal Husbandry Journal*, 2 (1): 29-33.

Isaac, U. C., Okeke B. U., Nwankwo C. B. (2023). Growth performance of Panda White and Cinnamon Brown strains of Japanese quail. *Proceedings of the 48th Annual Conference of the Nigerian Society* *for Animal Production* (NSAP) held at Federal University of Dutsin-Ma from 18th – 22nd June, 2023. Pp. 1339-1343.

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

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

Lee, B K., J. Lessler and E.A. Stuart (2010). Improving propensity score weighting using machine learning. *Stat. Med*. 29 (3): 337–346. https://doi.org/10.1002/sim.3782.

Mendeş, M. (2021). Re-evaluating the Monte Carlo simulation results by using graphical techniques. *Türkiye Klinikleri J. Biostatistics*, v.13, p.28-38.

Nwakpu, O.F., Nosike, R.J., Isaac, U.C., Nwose, R.N., Nwachukwu, C.C., Egbuedo, U.C., Nwaokoro, C.C., Nathaniel, J., Akinsola, K.L. and Onunkwo, D.N. (2020). Regression equation for selected growth traits prediction in Nigerian Indigenous Cattle (NIC) Genotype. *Nigerian Journal of Animal Science*, 22 (2): 1-6.

Nwaogwugwu, U.C., Udoh, U.H., Obike, O.M. and Ahamefule, G.O. (2018). Body weight prediction and phenotypic correlation with linear body measurements in two strains of Japanese quail. *Int'l J. Agric. And Rural Dev*., 21(1): 3547-3553.

Ojedapo, L. O., Adebambo, O. A., Ilori, L. J., and Ikeobi, C. O. N. (2019). Phenotypic characterization of two Nigerian indigenous chicken populations using body weight and linear body measurements. *Poultry Science, 98(5), 2083-2089.*

Oshima, T. C and T. Dell-Ross (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.

Razi, M.A. and K. Athappilly (2005). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert. Syst.Appl, 29 (1), 65 74.* https://doi.org/10.1016/j.eswa.2005.01.006.

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

Siegel, P. B., Honaker, C. F. (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

Smith, R. L., and Jones, M. P. (2020). Predictive modeling of livestock traits using advanced statistical tools. *Journal of Animal Science, 91(3), e13410.*

Wray, C.M and A.L. Byers (2020). Methodological Progress Note: Classification and Regression Tree Analysis. *J Hosp Med. 15(9): 549-551.* https://doi.org/10.12788/jhm.3366. PMID: 32195654.

Xi, W-F., Jiang, Q-W. and Yang, Ai-M. (2024). Using stepwise regression to address multicollinearity is not appropriate. *International Journal of Surgery*, 110(5), 3122–3123. https://doi.org/10.1097/JS9.0000000000001208

Yawuru, 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, A-R.; Raji, A. O.; Oseni, S. O. (2021). Multivariate principal component analysis of the morphometric traits of some cattle breeds in Nigeria. *Journal of Animal and Plant Sciences, 2021, Vol 31, Issue 3, p681: 1018-708.* http://dio.org/10.36899/JAPS.2021.3.0258.

Yakubu, A., Eyduran, E., Senol C., Juliana, O. I. (2022). Use of linear modeling, multivariate adaptive regression splines and decision trees in body weight prediction in goats. *Genetika, Vol 54, No.3, 1429-144*5. https://doi.org/10.2298/GENSR2203429Y.

Yakubu, A., Salako, A. E., and Imumorin, I. G. (2020). Genetic improvement of indigenous chickens: Challenges and prospects. *Worlds Poultry Science Journal, 76(2), 243-256.*

Yang, H. (2013). The case for being automatic: ıntroducing the automatic linear modeling (linear) procedure in SPSS Statistics. Multiple Linear Regression Viewpoints, v.39: 27-37,