RESEARCH ARTICLE

PREDICTION OF BODY WEIGHT FROM MORPHOMETRIC TRAITS OF WHITE LEGHORN USING DATA MINING ALGORITHM

Hlamulo Luke Matsambu[1], Raisibe Lisbert Mahlo[1], Victoria Rankotsane Hlokoe[1], Humbulani Jedidiah Phaduli[1], Lebo Trudy Rashijane[1], Pfunzo Xina Rammbebu[1], Madumetja Cyril Mathapo[1], Thobela Louis Tyasi[1*]

[1]School of Agriculture and Environmental Sciences, Department of Agricultural Economics and Animal Production, University of Limpopo, South Africa

[*]**Corresponding author**:
Dr. sc. Thobela Louis Tyasi

Address: X1106, Sovenga, 0727, Limpopo, South Africa
ORCID: 0000-0002-3519-7806
E-mail: louis.tyasi@ul.ac.za

Original Submission: 19 February 2025
Revised Submission: 23 June 2025
Accepted: 2 July 2025

How to cite this article: Matsambu HL, Mahlo RL, Hlokoe VR, Phaduli HJ, Rashijane LT, Rammbebu PX, Mathapo MC, Tyasi TL. 2025. Prediction of body weight from morphometric traits of White Leghorn using data mining algorithm. Veterinaria, 74(2), 159-67.

ABSTRACT

The study aimed to establish a predictive model for body weight in White Leghorn using morphometric traits through different data mining algorithms. Data was collected from 100 chickens, including body weight (BW), beak length (BKL), body length (BL), keel length (KL), chest girth (CG), body girth (BG), shank length (SL), back length (BCL), shank circumference (SC) and wing length (WL). Chi-Squared Automatic Interaction Detection (CHAID), Classification and Regression Trees (CART) and Exhaustive chi-squared Automatic Interaction detection (EX-CHAID) were used for data analysis. Based on goodness of fit criteria, CART model was the best model for prediction of body weight in White Leghorn chickens with higher values of correlation coefficient ($r = 0.84$) and coefficient of determination ($R^2 = 0.71$), and lower root mean square error (RMSE = 0.18), Akaike information criterion (AIC = -341.77) and Bayesian information criterion (BIC = -339.16). CART model identified CG, BL, and WL as key contributors to BW variation, suggesting that focusing on these traits can assist in BW prediction and support farmers in improving their chickens.

Keywords: CHAID, CART, Exhaustive CHAID, goodness of fit

INTRODUCTION

White Leghorn is a white feathered chicken breed, which is also known as an egg layer. It originates from Egypt, and they are valued for their egg-laying abilities and adaptability (Ewonetu, 2017). Morphometric traits have emerged as effective tools for predicting the body weight, that is critical for livestock management and breeding (Ebong et al., 2023). Data mining algorithms offer robust models for body weight prediction by leveraging morphometric traits (Tyasi et al., 2020). Body weight and morphometric traits have a significant relationship in measuring the growth of the domesticated chickens (Nosike et al., 2017). However, accurate animal weighing and assessment may be hampered by a lack of technical knowledge; small-scale farmers in rural areas are facing difficulty on accessing weighing scales, which makes this method challenging (Negash, 2021). Morphometric traits have been used to determine body weight in poultry and livestock species (Assan, 2013). Several studies focus on using data mining algorithms to predict body weight in chickens. Ogunshola et al. (2017) reported that morphometric traits of a chicken could be used to predict the ration of body weight at any period of age from 17-25 weeks old of Fulani ecotype chicken. Dalal et al. (2020) concluded that morphometric traits had an excessive effective and crucial correlation with 40 weeks body weight, indicating the effectiveness of the use of morphometric traits in predicting body weight in synthetic White Leghorn strain.

However, according to the author's knowledge, there is limited information on the prediction of body weight from morphometric traits of White Leghorn using data mining algorithms. Therefore, the objectives of the study were as follows: 1) To determine the relationship between the body weight and morphometric traits of the White Leghorn chicken breed, 2) To establish a model for prediction of body weight from morphometric traits of White Leghorn chicken breed using data mining algorithm. This study will help local farmers to gain knowledge about the relationship between the body weight and morphometric traits of White Leghorn and understand which morphometric traits can be used as a selection criterion to estimate body weight of White Leghorn chickens.

MATERIALS AND METHODS

Study area

The study was carried out at the University of Limpopo Experimental Farm, Limpopo province, South Africa. The farm is situated 10 kilometres north-west of the University of Limpopo (23°49' S; 29°41' E). The area has a semi-arid climate with an average temperature ranging from 10° C to 36° C in summer and 5° C to 28° C in winter, the farm receives an annual rainfall of less than 400 mm (Molabe et al., 2024).

Experimental animals, management and study design

The study used a total of 100 White Leghorn chickens. The chickens were raised following the ordinary husbandry practices of feeding systems, housing, vaccination, and health care as described by Alabi (2012). The chickens were housed under intensive production conditions. The chicken house was cleaned seven days before the chickens arrived and disinfected with Virokill disinfectants to avoid transmission of pathogenic diseases to the chickens. The biosecurity protocols were followed in the area, where the footbaths with disinfectant were placed at the door for disinfecting before entering the chicken house. The study used a cross-sectional study design with one-time measurements per bird.

Data collection

The data was collected through measurements of the body weight (BW) and morphometric traits. Morphometric traits such as beak length (BKL), body length (BL), keel length (KL), back length (BCL), chest girth (CG), shank length (SL), shank circumference (SC) and wing length (WL) of White Leghorn chickens were measured using a measuring tape calibrated in centimetres (cm), while the body weight of each chicken was measured in kilograms (kg) using a weighing scale as described by Tyasi et al. (2021). The measurements were collected by one person throughout to avoid individual variation on measuring.

Statistical analysis

The data collected was analysed using Statistical Package for Social Sciences (IMB SPSS, 2023) version 29.0 software. Pearson's correlation was used to examine the association between the body weight and morphometric traits, while CHAID, CART and Ex-CHAID were used to develop the models to predict body weight from morphometric traits. The predictive performance of CHAID, CART and Ex-CHAID was done using the goodness of fit criteria.

The following goodness of fit criteria were used in the study:

Correlation coefficient

$$R = \frac{cov\ (Yi\ ,Yip\ )}{Syi\ SYip}$$

Coefficient of determination

$$R^2 = \left[1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\right] \times 100$$

Adjusted coefficient of determination

$$R^2_{Adj} = \left[1 - \frac{\frac{1}{n-k-1}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\right] \times 100$$

Root mean square error

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}$$

Akaike information criterion

$$AIC = NLn\ (\frac{SSE}{N}) + 2p$$

Bayesian information criterion

$$BIC = NLn\ (\frac{SSE}{N}) + pLnN$$

Where:

Y_i, the actual body weight (g) of the White Leghorn;

$\hat{Y}_i$, the predicted body weight value of White Leghorn;

$\bar{Y}$, average of the actual body weight of the White Leghorn;

k, number of significant independent variables in the model;

and n, the total number of White Leghorns.

The residual value of each White Leghorn is expressed as $\mathcal{E}_i = Y_i - \hat{Y}_i$.

RESULTS

Table 1 below represents descriptive statistics of BW and morphometric traits of White Leghorn chicken breed. The findings displayed that the mean values of the morphometric traits ranged from 3.17 to 42.64 cm, with the chicken height having the highest mean value. The minimum and maximum values range from 1.20 to 50.00.

Table 1 Descriptive statistics for body weight and morphometric traits

Traits	Mean	SE	SD	Minimum	Maximum
BW	2.09	0.03	0.27	1.20	2.57
BL	24.78	0.22	2.16	20.00	29.00
WL	20.07	0.17	1.66	16.00	23.00
CG	37.46	0.27	2.74	31.00	43.00
BKL	3.17	0.07	0.72	2.00	5.00
BCL	6.44	0.13	1.28	4.00	10.00
CH	42.64	0.35	3.51	36.00	50.00
TBL	26.52	0.22	2.16	13.00	31.00

SE: stand. error, SD: stand. deviation, BW (kg): body weight, BL (cm): body length, WL (cm): wing length, CG (cm): chest girth, BKL (cm): beak length, BCL (cm): back length, CH (cm): chicken height, TBL (cm): tail-back length.

Table 2 below shows the correlation between body weight and morphometric traits of the White Leghorn chicken breed. The results showed that BW had a positive highly significant correlation ($P < 0.01$) with the BL, but positive significant correlation ($P < 0.05$) with CG, WL and CH. It was also found that the BKL, BCL, and TBL had a non-significant correlation ($P > 0.05$) with the BW.

Table 2 Correlation between body weight and morphometric traits

Traits	BW	BL	WL	CG	BKL	BCL	CH	TBL
BW (kg)								
BL (cm)	0.49**							
WL (cm)	0.27*	0.30**						
CG (cm)	0.18*	0.06^{ns}	-0.00^{ns}					
BKL (cm)	0.05^{ns}	-0.02^{ns}	0.06^{ns}	-0.19*				
BCL (cm)	0.12^{ns}	-0.04^{ns}	0.15^{ns}	-0.01^{ns}	0.21*			
CH (cm)	0.21*	-0.09^{ns}	0.16^{ns}	0.05^{ns}	0.04^{ns}	0.04^{ns}		
TBL (cm)	0.14^{ns}	-0.02^{ns}	0.25**	0.00^{ns}	0.08^{ns}	0.19*	0.27**	

BW: body weight, BL: body length, CG: chest girth, BKL: beak length, BCL: back length, CH: chicken height, TBL: tail-back length. **highly significant (P<0.01). *Significant (P<0.05). ns: non-significant.

Figure 1 below represents the CART model. CART model revealed that CG, BL and WL were found to be highly contributing to the variation of the BW of White Leghorn. This data mining algorithm consists of 8 nodes, with node 0 as the root node. Node 0 shows the average mean BW as 2.03 kg. Node 0 was split through CG into two subgroups, node 1 (CG ≤ 42.50 cm) with mean BW of 1.99 kg, and node 2 (CG > 42.50 cm) with mean BW of 3.00 kg. Node 1 was subdivided according to BL into two subgroups, node 3 (BL ≤ 21.50 cm) with mean BW of 1.43 kg, and node 4 (BL > 21.50 cm) with mean BW of 2.03 kg. Node 4 was further subdivided through WL into two subgroups, node 5 (WL ≤ 21.50 cm) with mean BW of 2.00 kg, and node 6 (WL > 21.50 cm) with mean BW of 2.17 kg. Node 6 was further split according to BL into two subgroups, node 7 (BL ≤ 25.50 cm) with mean BW of 2.50 kg, and node 8 (BL > 25.50 cm) with mean BW of 2.00 kg. The predicted values among the all nodes range from 1.43 to 3.00.

Figure 2 below, represents the CHAID model. The findings displayed that BL and BKL were significant variables for prediction of BW of White Leghorn. The CHAID model consists of 7 nodes with node 0 as the root node. The root node showed the mean BW of 2.03 kg. The root node was subdivided through BL into five subgroups, node 1(BL ≤ 21.00 cm) with mean BW of 1.43 kg, node 2 (BL, 21.00-22.00 cm) with mean BW of 2.00 kg, node 3 (BL, 22.00-23.00 cm) with mean BW of 2.00 kg, node 4 (BL, 23.00-24.00 cm) with mean BW of 2.50 kg, and node 5 (BL >24.00 cm) with mean BW of 2.05 kg. Node 5 was subdivided through BKL into two subgroups, node 6 (BKL ≤4.00 cm) with mean BW of 2.03 kg, and node 7 (BKL >4.00 cm) with mean BW of 2.50 kg. All these terminal nodes, node 4 and node 7 produced the maximum predicted value compared to other nodes observed.

Figure 3 below display of Ex-CHAID model. This model revealed BL and WL as significant morphometric traits, which highly contributed to the body weight variation of White Leghorn chicken breed. 11 nodes were obtained with node 0 as the root node. Node 0 as the root node was subdivided through BL into 5 subgroups, node 1(BL ≤21.00 cm) with mean BW of 1.43 kg, node 2 (BL, 21.00-22.00 cm) with mean BW of 2.00 kg, node 3 (BL, 22.00-23.00 cm) with mean BW of 2.00 kg, node 4 (BL, 23.00-24.00 cm) with mean BW of 2.50 kg, node 5 (BL >24.00 cm) with mean BW of 2.05 kg. Node 5 was also subdivided through WL into 6 subgroups, node 6(WL ≤18.00 cm) with mean BW of 2.00 kg, node 7 (WL, 18.00-19.00 cm) with mean BW of 2.00 kg, node 8 (WL, 19.00-20.00 cm) with mean BW of 2.00 kg, node 9 (WL, 20.00-21.00 cm) with mean BW of 2.00 kg, node 10 (WL, 21.00-22.00 cm) with mean BW of 2.33 kg, and node 11 (WL >22.00 cm) with mean

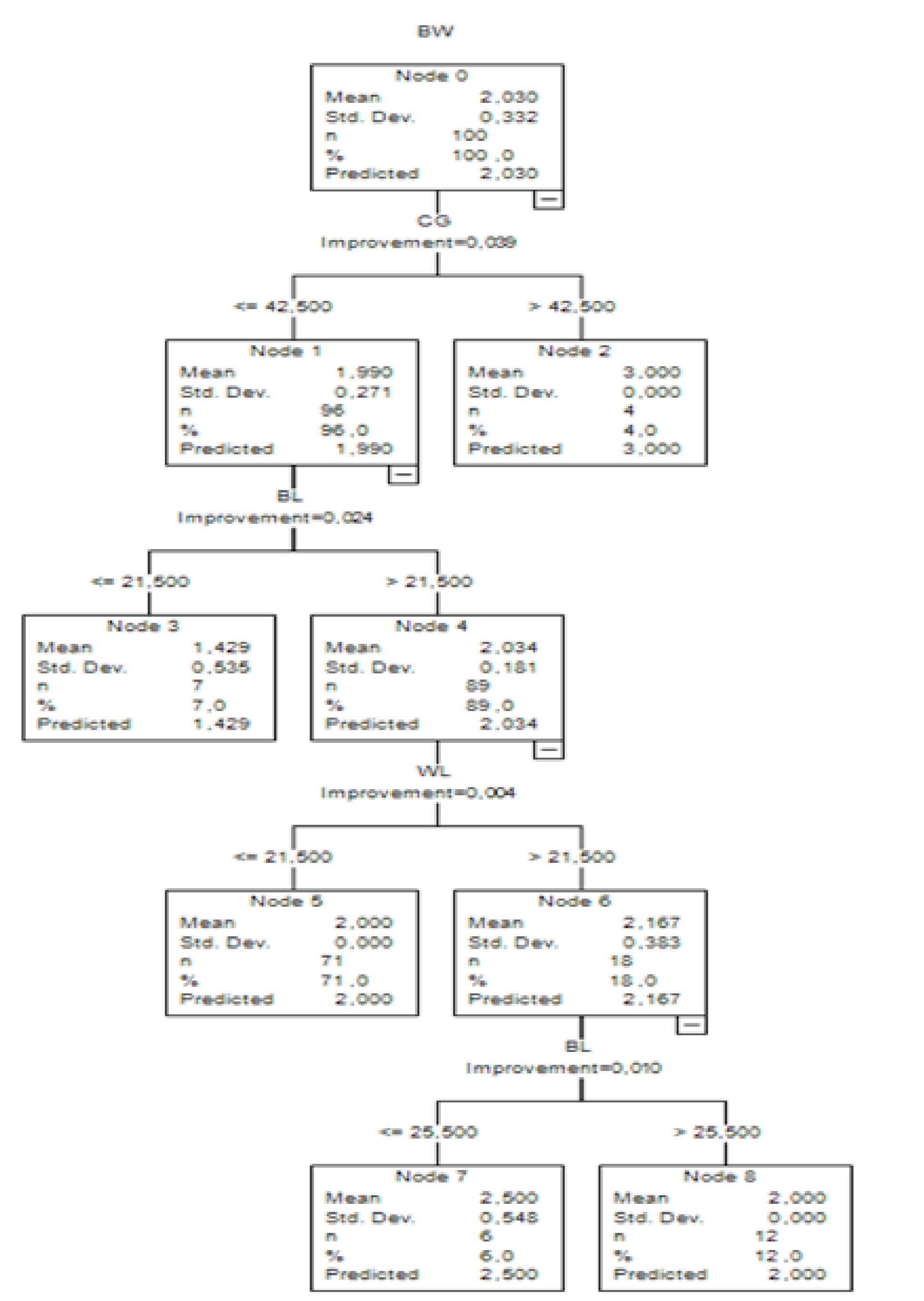

Figure 1 CART model

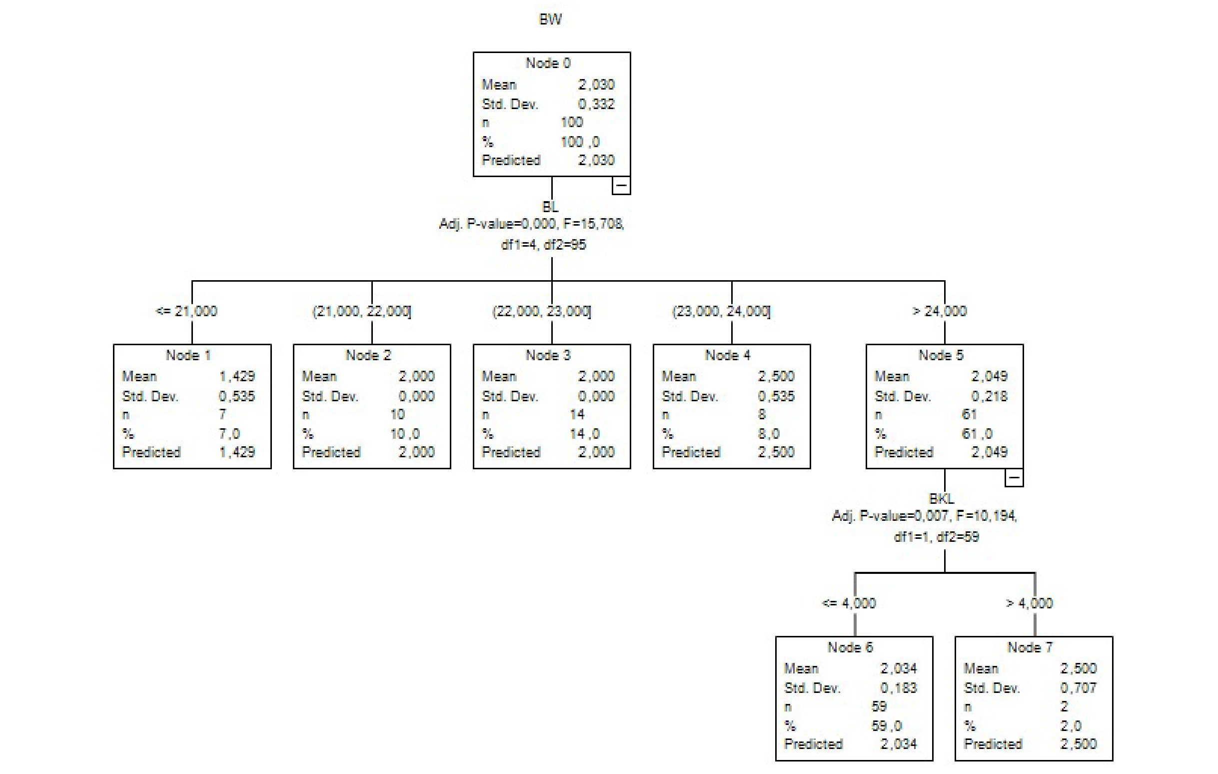

Figure 2 CHAID model

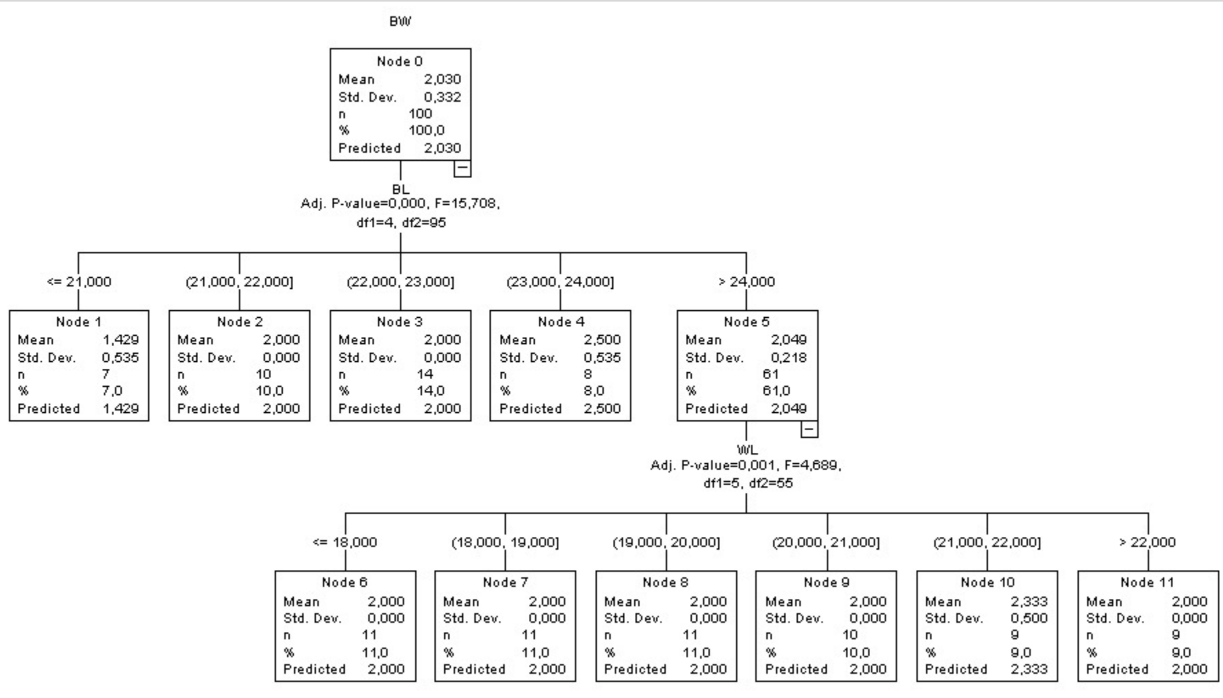

Figure 3 Exhaustive-CHAID model

BW of 2.00 kg. Among all these terminal nodes, node 4 shared a maximum predicted value than other nodes.

Table 3 below shows the predictive performance of CART, CHAID and EX-CHAID models. The results from the Table showed that CART model had a high value of r, R^2, AdjR^2, and a lower RMSE, AIC and BIC. The findings further displayed that the CHAID model had the lowest r, R^2, AdjR^2, and the highest RMSE, AIC and BIC. In this case, CART was the best predictive model for the prediction of body weight of the White Leghorn chicken breed.

Table 3 Predictive performance of CART, CHAID and Ex-CHAID

Criteria	CART	CHAID	EXHAUSTIVE-CHAID	Decision
r	0.84	0.66	0.72	Greater is better
R^2	0.71	0.44	0.52	Greater is better
Adj R^2	0.70	0.43	0.52	Greater is better
RMSE	0.18	0.25	0.23	Smaller is better
AIC	-341.77	-276.94	-293.38	Smaller is better
BIC	-339.16	-274.33	-290.78	Smaller is better

r: correlation coefficient, R^2: coefficient of determination, Adj R^2: adjusted coefficient of determination, RMSE: root mean square error, AIC: Akaike information criterion, BIC: Bayesian information criterion.

DISCUSSION AND CONCLUSION

Due to the lack of weighing scales, rural farmers with limited resources can predict livestock body weight using morphometric traits (Adhianto and Harris, 2020). Firstly, the present study examined the relationship between body weight and morphometric traits of White Leghorn using Pearson's correlation. The results on morphometric traits revealed that body length had a positive highly significant correlation with body weight, with chest girth, wing length and chicken height having a positive significant correlation with body weight. The findings of the current study were in accordance with those of Tyasi et al. (2020), who found that beak length, wing length and back length played an important role in the body weight of Potchefstroom Koekoek laying hens. Similarly, Dzungwe et al. (2018) found that morphological traits such as wing length had a significant role in the body weight of French Broiler Guinea Fowl. Ojo et al. (2014) also indicated that wing length had a significant correlation to the body weight of Japanese Quail. However, the findings of the current study contradict with the findings of Tyasi et al. (2017) who reported non-significant correlation between body weight and studied morphometric traits of Chinese Dagu chickens. The differences observed might be due to the genetic variations of the breed. Yunusa and Adeoti (2014) reported similar results to the current study, where body length and breast length were crucial morphometric traits, which helped to establish body weight of Yoruba and Fulani ecotype chickens. The correlation results of the current study imply that increasing body length, wing length, chest girth and chicken height might improve the body weight of White Leghorn chicken breed. The study further developed the best model for prediction of body weight from morphometric traits of White Leghorn chicken breed using data mining algorithms. The current study revealed that CART model explained 71% of the variance in White Leghorn body weight and demonstrated superior predictive accuracy. The CART model outperformed the CHAID and Ex-CHAID models. CART model showed chest girth as the best predictor trait of body weight of White Leghorn chicken breed. The findings of the current study were in accordance with those of Tyasi et al. (2024), who found that CART is the most effective model in the prediction of body weight in chickens as compared to CHAID. The findings of the current study were in disagreement with the findings of Gevrekçi and Takma (2018) who showed that CHAID model was the best in the prediction of body weight in poultry species. However, the variations in the findings might be due to breed and environment differences, and data characteristics. More studies need to be conducted on the prediction of body weight using data mining algorithms in chickens. CHAID performed poorly, and this might be caused by small sample size within the nodes, which resulted in overfitting of the model. The findings of the current study on CHAID performance were supported by Lemke et al. (2009).

The current study concludes that there is a positive relationship that exists between body weight and chest girth, body length and wing length. These traits can be used to enhance body weight of White Leghorn chickens. CART model proved to be the best-fitting model for accurately predicting body weight, indicating that chest girth is a key predictor of body weight in the White Leghorn chicken breed. The findings from the current study might help researchers, chicken breeders or farmers on which morphometric traits to use when they want to predict the BW of their chickens. It is recommended that further studies can be conducted on prediction of the body weight of chickens using different data mining algorithms.

ACKNOWLEDGEMENTS

The authors deeply thank the University of Limpopo for providing the necessary resources and facilities needed to carry out the study, and National Student financial Aid Scheme (NSFAS).

CONFLICT OF INTEREST

The authors declared that there is no conflict of interest.

CONTRIBUTION

Conception: TLT; Design: TLT; Supervision: TLT, MCM; Materials: HLM, RLM, VRH, HJP, LTR, PXR; Collection/or Processing: HLM, RLM, VRH, HJP, LTR, PXR; Analysis and/or Interpretation: MCM, HLM, RLM, VRH, HJP, LTR, PXR; Literature Search: HLM, RLM, VRH, HJP, LTR, PXR; Writing – Original Draft: HLM, RLM, VRH, HJP, LTR, PXR; Critical Review: TLT, MCM

REFERENCES

Adhianto K, Harris I. 2020. Prediction of body weight through body measurements in Boerawa bucks at Tanggamus Regency of Indonesia. Bulgarian J Agricult Sci, 20(6), 1273-9.

Alabi OJ, Ng'ambi JW, Norris D, Egena SSA. 2012. Comparative study of three indigenous chicken breeds of South Africa: Body weight and linear body measurements. Agricult J, 7(3), 220-5.

Assan N. 2013. Bioprediction of body weight and carcass parameters from morphometric measurements in livestock and poultry. Sci J Rev, 2, 140-50.

Dalal DS, Ratwan P, Malik BS, Patil CS, Kumar M. 2020. Principal component analysis of morphological traits of synthetic White Leghorn chicken. Indian J Anim Sci, 90(11), 1551–5.

Dzungwe JT, Gwaza DS, Egahi JO. 2018. Statistical modeling of body weight and body linear measurements of the French broiler Guinea fowl in the humid tropics of Nigeria. J Poult, Fisheries Wildlife Sci, 6(2), 1-4.

Ebong UN, Sam IM, Essien C. 2023. Prediction of body weight and carcass yield from morphometric traits of three strains of broiler chicken. Anim Res Int, 20(3), 5047-58.

Ewonetu KS. 2017. Growth performance and rearing costs of Fayoumi and White Leghorn chicken breeds. East African J Sci, 11(1), 37-42.

Gevrekçi Y, Takma Ç. 2018. A comparative study for egg production in layers by decision tree analysis. Pak J Zoolo, 50(2), 437-44.

Lemke C, Gabrys B, Buhmann JM. 2009. Automatic selection of spectral channels using decision trees. Patterns Recognition Letters, 30(9), 865-72.

Molabe KM, Tyasi TL, Mbazima V, Gunya B. 2024. Genetic polymorphism of growth hormone gene in Dorper sheep using PCR-RFLP. Indian J Anim Res, 58(3), 371-5.

Negash F. 2021. Predicting body weight of Ethiopian indigenous chicken populations from morphometric measurements. Turkish J Agricult - Food Sci Technolo, 9(6), 1138-43.

Nosike RJ, Onunkwo DN, Obasi EN, Amaduruonye W, Ukwu HO, Nwakpu OF, et al. 2017. Prediction of body weight with morphometric traits in some broiler chicken strains. Nigerian J Anim Product,, 44(3), 15-22.

Ogunshola OJ, Daramola SA, Omotoso OB, Baki OI, Chineke CA. 2017. Body weight and morphometric traits of Fulani ecotype chickens in South Western Nigeria. Nigerian J Anim Product, 44(5), 19-23.

Ojo V, Fayeye TR, Ayorinde KL, Olojede H. 2014. Relationship between Body Weight and Linear Body Measurements in Japanese Quail (Coturnix coturnix japonica). J Sci Res, 6(1).

Tyasi TL, Qin N, Jing Y, Mu F, Zhu HY, Liu D, et al. 2017. Assessment of relationship between body weight and body measurementtraits of indigenous Chinese Dagu chickens using path analysis. Indian J Anim Res, 51(3), 588-93.

Tyasi TL, Makgowo KM, Mokoena K, Rashijane LT, Mathapo MC, Danguru LW, et al. 2020. Classification and regression tree (CRT) analysis to predict body weight of Potchefstroom koekoek laying hens. Advances in Animal and Veterinary Sciences, 8(4), 354-59.

Tyasi TL, Eyduran E, Celik S. 2021. Comparison of tree-based regression tree methods for predicting live body weight from morphological traits in Hy-line silver brown commercial layer and indigenous Potchefstroom Koekoek breeds raised in South Africa. Trop Anim Health Product, 53(1), 1-8.

Tyasi TL, Mathapo MC, Mokoena K, Hlokoe VR, Molabe KM. 2024. Use of data mining algorithms in chicken breeding: A systematic review. Advant Poult Sci Res, 12(3), 45-56.

Yunusa AJ, Adeoti TM. 2014. Multivariate analysis for body weight and some linear body measurements of Nigerian indigenous chickens. Slovak J Anim Sci, 47(3), 142-8.

PREDVIĐANJE TJELESNE TEŽINE NA OSNOVI MORFOMETRIJSKIH KARAKTERISTIKA BIJELOG LEGHORNA KORIŠTENJEM ALGORITMA ZA RUDARENJE PODATAKA

SAŽETAK

Cilj istraživanja je kreirati prediktivni model za određivanje tjelesne težine Bijelog leghorna na osnovi morfometrijskih karakteristika korištenjem algoritama za rudarenje podataka. Podaci su prikupljeni na 100 pilića, a odnose se na tjelesnu težinu (BW), dužinu kljuna (BKL), dužinu grebena (KL), obim prsa (CG), obim tijela (BG), dužinu goljenice (SL), dužinu leđa (BCL), obim golenjače (SC) i dužinu krila (WL). Za analizu su korišteni: *Chi-Squared Automatic Interaction Detection* (CHAID), *Classification and Regression Trees* (CART) i *Exhaustive chi-squared Automatic Interaction detection* (EX-CHAID). Na osnovu usklađenosti, CART model je bio najbolji model za predviđanje tjelesne težine kod Bijelog leghorna sa višim vrijednostima koeficijenta korelacije (r = 0.84) i koeficijenta determinacije (R^2 = 0.71) i nižim korijenom srednje kvadratne pogreške (RMSE = 0.18), *Akaike* informacijskim kriterijem (AIC = -341.77) i Bayesovim informacijskim kriterijem (BIC = -339.16). CART model je identificirao CG, BL i WL kao ključne faktore varijacije BW, ukazujući kako fokusiranje na ove karakteristike može olakšati predviđanje tjelesne težine i biti od pomoći uzgajivačima u poboljšanju karakteristika pilića.

Ključne riječi: CART, CHAID, Exhaustive CHAID, usklađenost