



Journal of Frontiers in Multidisciplinary Research

Integrating Machine Learning with Morphometric Indicators to Predict Body Weight in Female Indigenous Matebele Goats Across Tooth-Age Classes in Zimbabwe

Never Assan^{1*}, Michael Musasira², Abbegal Dube³, Edward Manda Mkokora⁴

¹ Zimbabwe Open University, Faculty of Agriculture, Department of Agriculture Management, Bulawayo Regional Campus, Bulawayo, Zimbabwe

² Matopos Research Station, Ministry of Agriculture, Mechanization and Irrigation Development, Department of Research and Extension, Private Bag 5137, Bulawayo, Zimbabwe

³ Esigodini Agricultural College, Ministry of Agriculture, Mechanization, and Irrigation Development, P.O. Box 5808, Bulawayo, Zimbabwe

⁴ Masotsha High School, 3711 Sigwadi Road, Magwegwe North, Bulawayo, Zimbabwe

* Corresponding Author: **Never Assan**

Article Info

E-ISSN: 3050-9726

P-ISSN: 3050-9718

Volume: 07

Issue: 01

Received: 10-11-2025

Accepted: 12-12-2025

Published: 14-01-2026

Page No: 45-50

Abstract

Accurate estimation of body weight (BWT) using linear body measurements (LBMs) is critical in smallholder goat systems, where scales are often unavailable. This study evaluated four machine learning (ML) models—Random Forest (RF), Decision Tree (CART), Support Vector Regression (SVR), and Multi-Layer Perceptron (MLP)—to predict BWT from heart girth (HG), withers height (WTH), body length (BL), and rump height (RH). Using 132 female goat observations, RF achieved the highest predictive performance ($R^2 = 0.759$; RMSE = 2.425 kg), with HG identified as the dominant predictor, corroborating previous findings of its strong correlation with BWT. CART models offer interpretable decision rules suitable for field applications, highlighting their practical utility for smallholder farmers. These results demonstrate that ML approaches can provide accurate and accessible tools for predicting BWT and informing feeding strategies, health interventions, selection, and market decisions in resource-limited goat production systems.

DOI: <https://doi.org/10.54660/JFMR.2026.7.1.45-50>

Keywords: Indigenous Matebele goat females, body weight prediction, heart girth, machine learning, Random Forest, CART, tooth-age

1. Introduction

Accurate estimation of body weight (BWT) in goats is a cornerstone of effective livestock management, informing feed allocation, medication dosing, and market valuation (Assan *et al.*, 2024; Pesmen & Yardimci, 2025)^[4]. In smallholder systems, access to weighing scales is often limited, making linear body measurements (LBMs), such as heart girth (HG), withers height (WTH), body length (BL), and rump height (RH), practical and accessible proxies for estimating BWT (Mokoena *et al.*, 2022; Mathapo *et al.*, 2025)^[20, 16]. These LBMs have been shown to correlate strongly with body weight across different goat breeds and management systems, highlighting their utility for farmers in low-resource settings (Gonçalves *et al.*, 2025; Iqbal *et al.*, 2021)^[9, 13].

Traditional linear regression models are commonly used to predict BWT from LBMs; however, they often fail to capture the complex, nonlinear relationships between measurements and weight, particularly across different maturity or tooth-age classes (Abbas *et al.*, 2021; Kozaklı *et al.*, 2024)^[1, 14]. Recent advances in machine learning (ML) offer promising alternatives for modeling these complex systems. Algorithms such as Random Forest (RF), Support Vector Regression (SVR), multilayer perceptron (MLP), and Classification and Regression Trees (CART) can effectively handle nonlinear relationships and

interactions among variables, providing more accurate and robust predictions (Haldar *et al.*, 2023; Tırnk *et al.*, 2023)^[11, 23]. Several studies have demonstrated the superiority of ML methods over traditional regression methods for BWT prediction in goats. For example, Gonçalves *et al.* (2025)^[9] showed that RF models could achieve R² values exceeding 0.90, whereas Mokoena *et al.* (2022)^[20] found that data mining algorithms outperformed linear models in Kalahari Red goats. Furthermore, CART models offer interpretable decision rules that can be directly applied in the field, making them practical for smallholder farmers (Haldar *et al.*, 2023; Abbas *et al.*, 2021)^[11, 1].

This study aimed to evaluate and compare the performances of RF, CART, SVR, and MLP models for predicting BWT from LBMs in female goats, incorporating tooth age as a categorical predictor. By leveraging these ML approaches, we aim to provide smallholder farmers with accessible, accurate, and practical tools for body weight estimation, thereby improving livestock management, health interventions, and market decision-making (Assan *et al.*, 2024; Pesmen & Yardimci, 2025)^[4].

2. Materials and Methods

2.1. Study Site

This study was conducted at the Matopos Research Station, Bulawayo, Zimbabwe (22.23°S, 31.30°E). The area experiences a dry season from April to October and a rainy season from November to March, with a mean annual rainfall below 446.8 mm (Assan, 2023)^[3]. Temperatures range from 11.4°C to 21.6°C during the hottest months, with low overall rainfall (<450 mm) (Hagreaves *et al.*, 2004; Homann *et al.*, 2007)^[10, 12]. The landscape is predominantly rangeland with sweet veld vegetation, providing high-quality forage suitable for sustaining small ruminants (Ward *et al.*, 1979; Ncube, 2005; Van Rooyen *et al.*, 2007)^[25, 21].

2.2. Flock Management and Data Collection

Indigenous Matebele goat flock management has been described previously (Assan *et al.*, 2024)^[4]. For this study, body weight (BWT) was measured using a balance weighing scale, and linear body measurements (LBMs) were obtained using a calibrated measuring tape, following FAO (2012)^[7] guidelines. Data were collected from 132 female goats, including four LBMs: heart girth (HG), withers height (WTH), body length (BL), and rump height (RH). Observations with incomplete data were excluded, resulting in 132 complete records. LBMs were measured with animals standing in a natural position and head raised, performed by a single trained technician to minimize intra-observer variation (Yilmaz *et al.*, 2013)^[27]. Circumferences were measured using a flexible tape, while calipers were used for length and height measurements. All procedures adhered to

the ethical standards approved by the Ethical Committee of the Matopos Research Institute, with efforts made to minimize animal discomfort.

Definitions of Linear Body Measurements

Heart Girth (HG): Circumference around the chest, compressing excess hair for mature animals. HG is highly repeatable and strongly correlated with BWT.

Withers Height (WTH): Distance from a platform to the withers, measured using a vertical stick while the animal stood with equal weight distribution.

Body Length (BL): Distance from the base of the neck to the tail or nose while maintaining a straight backbone.

Rump Height (RH): Distance from the platform surface to the rump, measured similarly to the WTH.

2.3 Model Development and Evaluation

Prior to model training, numeric features were standardized for Support Vector Regression (SVR) and Multi-Layer Perceptron (MLP) models to ensure uniform feature scaling, while Random Forest (RF) and Decision Tree (CART) regressors were trained on raw data to preserve their inherent ability to handle unscaled variables. The dataset was partitioned into training (70%) and test (30%) subsets using a fixed random seed (random_state = 42) to maintain reproducibility.

Four regression models were developed and compared: a random forest (RF) regressor comprising 300 trees, a decision tree (CART) regressor, a Support Vector Regression (SVR) model employing a radial basis function (RBF) kernel, and a Multi-Layer Perceptron (MLP) configured with two hidden layers containing 64 and 32 neurons, respectively, with a maximum of 3000 training iterations. The model performance was assessed using the coefficient of determination (R²), root mean squared error (RMSE), and mean absolute error (MAE). Additionally, feature importance values were derived from the RF model to identify key predictors, and decision rules from the CART model were examined to enhance field-level interpretability and practical understanding of predictive relationships.

3. Results

The descriptive statistics for body weight (BWT) and morphometric traits of indigenous Matebele goat females across tooth age classes are presented in Table 1. Older goats had higher mean BWT and heart girth (HG), indicating progressive growth and increased body volume with maturation. These results highlight the consistent proportionality between body dimensions and weight across tooth age classes, demonstrating the reliability of linear body measurements as growth indicators.

Table 1: Descriptive statistics of body weight (bwt) and morphometric traits of indigenous Matebele female goats across tooth-age classes

Tooth-age class	n	BWT (kg)±SE	HG (cm)±SE	WTH (cm)±SE	BL (cm)±SE	RH (cm)±SE
2-tooth	20	23.30±0.52	69.30±1.20	47.00±0.95	56.40±0.88	49.20±0.75
6-tooth	29	31.28±0.72	77.55±0.72	51.23±0.81	60.78±0.64	51.89±0.82
Full mouth	8	31.13±1.60	77.75±1.33	51.88±1.20	61.25±1.05	52.00±1.18
Broken mouth	35	33.86±0.65	77.91±0.68	52.14±1.13	62.69±0.56	53.12±1.10

A strong positive linear relationship between HG and BWT was observed in all age classes (Figure 1). This relationship demonstrates that HG serves as an effective proxy for

estimating body weight in field conditions where weighing scales are unavailable.

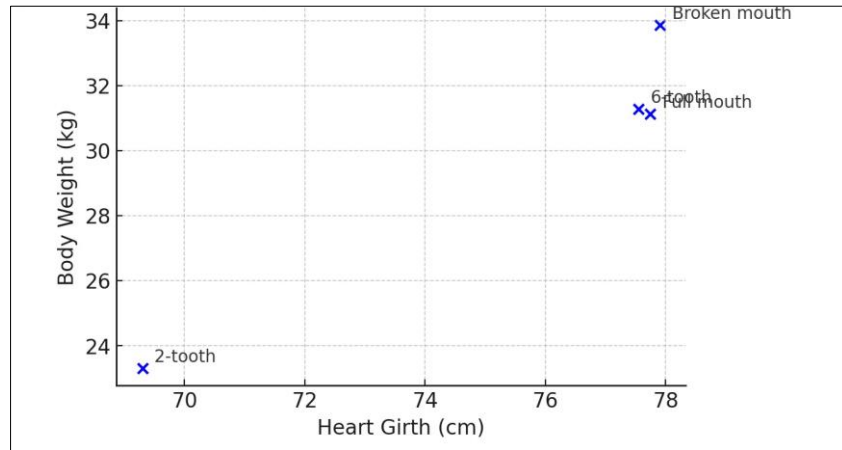


Fig 1: Relationship between Heart Girth and Body weight

The predictive performance metrics of the four machine learning models are summarized in Table 2. Random Forest (RF) achieved the highest predictive accuracy ($R^2 = 0.759$; RMSE = 2.425 kg), followed by CART ($R^2 = 0.697$). The

SVR and MLP models performed less effectively owing to the limited sample size and sensitivity to hyperparameters, emphasizing the robustness of the ensemble-based algorithms.

Table 2: Predictive performance of machine learning models for indigenous Matebele goat female body weight estimation

Model	R^2	RMSE (kg)	MAE (kg)
Random Forest	0.759	2.425	1.329
Decision Tree (CART)	0.697	2.72	1.35
Support Vector Regression (SVR)	0.537	3.363	2.416
Neural Network (MLP)	0.209	4.394	2.142

Figure 2 compares the predictive performance of the models. The superior accuracy of RF reflects its ability to capture nonlinear interactions among body measurements, whereas

CART offers interpretable rules suitable for smallholder decision-making.

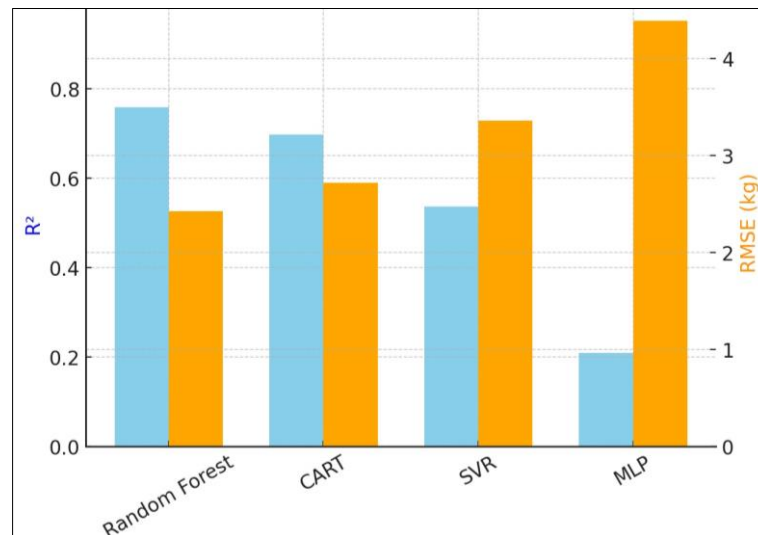


Fig 2: Comparative Performance of ML Models

Feature importance analysis from the RF model (Table 3) identified heart girth (HG) as the dominant predictor (importance = 0.737), followed by body length (BL), rump

height (RH), and withers height (WTH). The strong influence of HG is consistent with its biological role in reflecting body volume and muscle mass.

Table 3: Random Forest feature importance for predicting indigenous Matebele goat female body weight.

Predictor	Importance score
Heart girth (HG)	0.737
Body length (BL)	0.109
Rump height (RH)	0.09
Withers height (WTH)	0.064

The simplified CART decision thresholds are listed in Table 4. Goats with HG ≤ 71 cm were classified as light (~21 kg), those between 72 and 74 cm as medium (~23 kg), and those

with HG ≥ 75 cm and RH ≥ 50 cm as heavy (~29 kg). These rules (Figure 3) are intuitive and easily applicable by farmers using simple measurements with a tape measure.

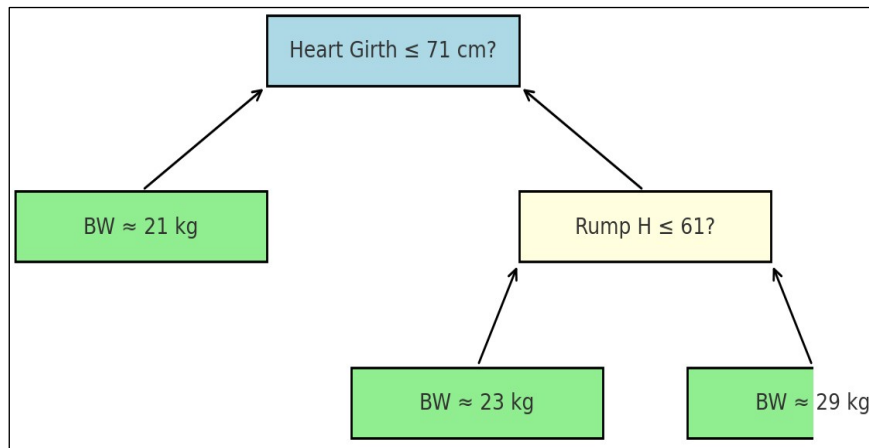


Fig 3: Simplified Cart Decision Tree for Goat Body Weight

Table 4: Simplified CART-derived decision rules for estimating indigenous Matebele goat female body weight.

Heart Girth (cm)	Rump Height (cm)	Body Length (cm)	Predicted BW (kg)
≤ 71	—	≤ 44	~21
72–74	≤ 61	≤ 50	~23
≥ 75	≥ 50	—	~29

The residual error distributions (Figure 4) show that RF exhibited the narrowest and most symmetric residual spread centered around zero, indicating high predictive precision

and minimal bias. CART maintained moderate precision, whereas SVR and MLP displayed wider residual dispersion, suggesting under-fitting (Figure 4).

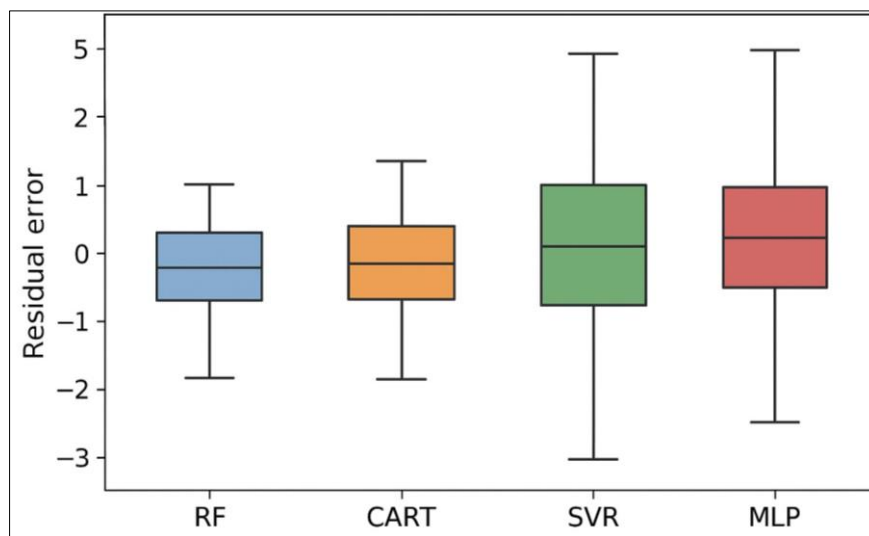


Fig 3: Residual error distributions across machine learning models (RF, CART, SVR, and MLP)

The key model hyperparameters and performance notes are summarized in Table 5. RF demonstrated the best trade-off between accuracy and generalization, whereas CART offered

transparency suitable for practical applications. The SVR and MLP models underperformed owing to the small data size and high complexity.

Table 5: Key hyperparameters and performance notes for machine learning models predicting indigenous Matebele goat female body weight.

Model	Key	Description / Notes
Random Forest (RF)	n_estimators	It achieved
Decision Tree (CART)	max_depth	It provides
Support Vector Regression (SVR, RBF kernel)	C	Performance was sensitive to small sample
Multi-Layer Perceptron (MLP)	hidden_layer_sizes	The model underperformed

Finally, Figure 5 presents a conceptual framework that integrates three key pathways—data enhancement, model optimization, and field application—to improve predictive modeling in smallholder goat systems. This approach

emphasizes the combination of robust ensemble learning with practical decision-support tools to strengthen data-driven livestock management strategies.

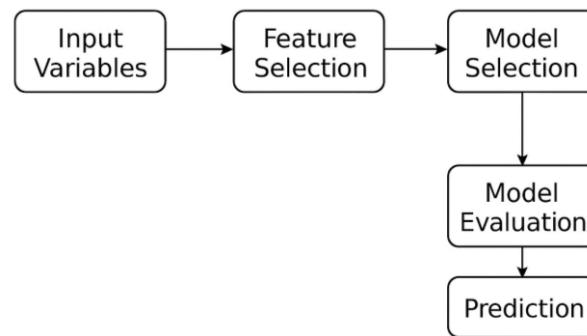


Fig 5: Conceptual framework for improving predictive modeling in indigenous Matebele goat female body weight estimation

4. Discussion

Heart girth (HG) emerged as the strongest single predictor of body weight, consistent with the extensive literature demonstrating its reliability across livestock species. As shown in Figure 1, the scatterplot reveals a strong linear trend between HG and body weight, reinforcing the conclusion that chest circumference directly reflects body mass (Melesse *et al.*, 2022; Yakubu *et al.*, 2021). This makes HG a valuable proxy for weighing scales.

Random Forest (RF) demonstrated the best overall predictive accuracy, balancing robustness and performance, even with a modest sample size ($n = 132$). The model's superior performance is evident in the comparative metrics displayed in Figure 2, which highlights its lower RMSE and higher R^2 relative to other approaches. These findings are consistent with those of studies confirming the effectiveness of ensemble models in animal breeding and morphometric prediction (Getachew *et al.*, 2023; Sharma *et al.*, 2022).

Although less accurate than RF, CART provides interpretable decision rules that can be highly valuable for extension services and field-based applications. An example rule path is illustrated in Figure 3, which shows how simple thresholds can be used to estimate body weight. This interpretability is especially useful for training farmers and extension officers in resource-limited settings (Okoro *et al.* 2021).

In contrast, Support Vector Regression (SVR) and Multilayer Perceptron (MLP) underperformed in this study, likely because of the relatively small dataset and the sensitivity of the models to hyper parameter tuning. As illustrated in the residual error plots (Figure 4), SVR and MLP had wider and less symmetric error distributions than RF. This outcome is consistent with evidence that neural networks and kernel-based methods require larger, more complex datasets to surpass ensemble tree models in terms of predictive accuracy (Chen *et al.*, 2020; Tesema *et al.*, 2024).

Future improvements could include expanding the sample size, incorporating continuous predictors such as age and physiological status, and employing mixed-effects models where repeated measures are available. Additionally, including environmental and management covariates may refine the predictive accuracy by capturing external influences on growth. A conceptual framework summarizing the proposed refinements is presented in Figure 5. Such multidimensional modeling approaches align with recent calls for the integration of genetics, environment, and

management in livestock improvement strategies (Wang *et al.*, 2023; Vanvanhossou *et al.*, 2025).

5. Conclusion

This study confirmed that heart girth is the most reliable linear predictor of body weight, consistent with recent findings across livestock species. Among the models tested, Random Forest offered the best balance of predictive accuracy and robustness, whereas CART provided simple, interpretable rules suitable for extension and field use. In contrast, the SVR and MLP models underperformed, likely because of the modest sample size and tuning requirements, highlighting the need to align model complexity with the dataset scale.

Future improvements should focus on expanding the sample size, incorporating age, sex, and physiological state as continuous predictors, and integrating environmental and management variables to enhance model transferability across diverse production systems. These insights underscore the dual need for scientific accuracy and field usability, pointing toward hybrid approaches that combine interpretable rules with ensemble methods to strengthen decision-making in smallholder and experimental livestock systems.

Author Contributions: NA and MM conceptualized the study and wrote the original draft. NA developed the methodology. EMM conducted a formal analysis. NA and MM reviewed and edited the manuscript text. All the authors have reviewed and approved the final version of the manuscript.

Conflict of interest: The authors declare no conflicts of interest.

Funding: This study was conducted without financial support or funding.

Statement of animal rights: The Ethical and Academic Research Committee of Experimental Animals at the Matopos Research Station, Zimbabwe, approved this study.

Data availability statement: The data supporting the findings of this study can be obtained from the corresponding author upon reasonable request.

Acknowledgment: The authors wish to thank the Matopos Research Station for kindly allowing us to utilize their experimental flock of indigenous Matebele goats for this study.

6. References

1. Abbas A, *et al.* Body weight prediction using different data mining algorithms in goats. *Small Ruminant Res.* 2021;199:106307. doi:10.1016/j.smallrumres.2021.106307
2. Ampomah EK, Qin Z, Nyame G. Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement. *Information.* 2020;11(6):332. doi:10.3390/info11060332
3. Assan N, Musasira M, Mpfu M, Mwareya N. Species-dependent correlation analysis and regression models of body weight on linear body measures in indigenous sheep and goats of Zimbabwe. *Adv Mod Agric.* 2023;4(2):2388. doi:10.54517/ama.v4i2.2388
4. Assan N, Musasira M, Mpfu M, Mwayera N, Mokoena K, Tyasi TL. Relationship between body weight and linear body measurements at various stages of permanent tooth eruption in indigenous Matebele female goats in Zimbabwe. *Adv Anim Vet Sci.* 2024;12(9):1818-1828. doi:10.17582/journal.aavs/2024/12.9.1818.1828
5. Cao S, Xu Q, Li X, Qi F. The predictive capability of a novel ensemble tree-based algorithm for assessing groundwater potential. *Sustainability.* 2021;13(5):2459. doi:10.3390/su13052459
6. Chen CH, Tanaka K, Kotera M, Funatsu K. Comparison and improvement of the predictability and interpretability with ensemble learning models in QSPR applications. *J Cheminform.* 2020;12(1):19. doi:10.1186/s13321-020-0417-9
7. Food and Agriculture Organization of the United Nations. *FAO animal production and health guidelines.* Rome: FAO; 2012.
8. Getachew T, Haile A, Tessema T, Dea D, Edea Z, Rischkowsky B. Participatory identification of breeding objective traits and selection criteria for indigenous goats of pastoral communities in Ethiopia. *Trop Anim Health Prod.* 2020;52:2145-2155. doi:10.1007/s11250-020-02243-4
9. Gonçalves MA, *et al.* Prediction of weight and body condition score of dairy goats using Random Forest algorithm and digital imaging data. *Animals.* 2025;15(10):1449. doi:10.3390/ani15101449
10. Hagreaves SK, Bruce D, Beffa LM. Disaster mitigation options for livestock production in communal farming systems in Zimbabwe. [Place unknown]: ICRISAT & FAO; 2004.
11. Haldar A, *et al.* Body weight prediction using recursive partitioning and regression trees (RPART) model in Indian Black Bengal goat breed: a machine learning approach. *Indian J Anim Res.* 2023;57(9):1251-1257. doi:10.18805/IJAR.B-4894
12. Homann S, Van Rooyen AF, Moyo T, Nengomasha Z. Goat production and marketing: baseline information for semi-arid Zimbabwe. [Place unknown]: ICRISAT; 2007.
13. Iqbal F, *et al.* Comparing the predictive ability of machine learning algorithms in predicting the live body weight of Beetal goats in Pakistan. *Pak J Zool.* 2021;53(1):231-238. doi:10.17582/journal.pjz/20191003081007
14. Kozaklı Ö, *et al.* Comparison of machine learning algorithms and multiple linear regression for predicting the live body weight of Beetal goats. *Pak J Zool.* 2024;56(1):231-238. doi:10.17582/journal.pjz/2022.56.1.231.238
15. Maseko SK, Van der Laan M, Tesfamariam EH, Delport M, Otterman H. Evaluating machine learning models and identifying key factors influencing spatial maize yield predictions in data-intensive farm management. *Eur J Agron.* 2024;157:127193. doi:10.1016/j.eja.2024.127193
16. Mathapo MC, *et al.* Prediction of body weight from linear body measurements in Nguni goats. *J Anim Sci Technol.* 2025;67(1):1-10. doi:10.1186/s40781-025-00452-6
17. Mathapo MC, Mugwabana TJ, Tyasi TL. Prediction of body weight from morphological traits of South African non-descript indigenous goats using different data mining algorithms. *Trop Anim Health Prod.* 2022;54:102. doi:10.1007/s11250-021-03018-5
18. Melesse A, Banerjee S, Lakew A, Mersha F, Hailemariam F, Tsegaye S. Morphological characterization of indigenous sheep in the Southern Regional State of Ethiopia. *Anim Genet Resour.* 2013;52:39-50.
19. Mokoena K, Tyasi TL. Morphological structure of South African Boer goats explained by principal component analysis. *Veterinaria.* 2021;70(3):325-334.
20. Mokoena K, *et al.* Predicting body weight of Kalahari Red goats from linear body measurements using data mining algorithms. *Vet World.* 2022;15:1719-1726. doi:10.14202/vetworld.2022.1719-1726
21. Ncube S. The role of proanthocyanidins and related flavonoids in the utilization of sorghum grain varieties as feed for ruminants in the semi-arid areas of Zimbabwe [doctoral dissertation]. Harare: University of Zimbabwe; 2005.
22. Okoro VMO, Chikereuba U, Iheanacho G. West African Dwarf goat populations of the tropical rainforest zone: characterization using morpho-biometric traits. *Trop Anim Health Prod.* 2023;55(3):222. doi:10.1007/s11250-023-03641-0
23. Tirink C, *et al.* Estimation of body weight based on biometric measurements in Polish Merino sheep using machine learning algorithms. *Animals.* 2023;13(7):1146. doi:10.3390/ani13071146
24. Wang Z, Doekes H, Bijma P. Towards the genetic improvement of social behaviors in livestock using large-scale sensor data. *Genet Sel Evol.* 2023;55:67. doi:10.1186/s12711-023-00840-z
25. Ward HK, Richardson FD, Denny RP, Dye PJ. Matopos Research Station: a perspective. *Rhod Agric J.* 1979;76(1):5-18.
26. Yakubu A. Path coefficient and path analysis of body weight and biometric traits in Yankasa lambs. *Slovak J Anim Sci.* 2010;43(1):17-25.
27. Yilmaz O, Cemal I, Karaca O. Estimation of mature live weight using some body measurements in Karya sheep. *Trop Anim Health Prod.* 2013;45:397-403. doi:10.1007/s11250-012-0229-7.

How to Cite This Article

Assan N, Musasira M, Dube A, Mkokora EM. Integrating machine learning with morphometric indicators to predict body weight in female indigenous Matebele goats across tooth-age classes in Zimbabwe. *J Front Multidiscip Res.* 2026;7(1):45-50. doi:10.54660/JFMR.2026.7.1.45-50.

Creative Commons (CC) License

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed

under the identical terms.