INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

Decision Tree and Automatic Linear Modeling Approaches to Predict Body Weight in Indigenous Sabi Sheep and Matebele Goat Females of Zimbabwe

Never Assan[1], Michael Musasira[2], Edward Manda Mkokora[3], Abbegal Dube[4]

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

[2]Matopos Research Station, Ministry of Lands and Agriculture, Department of Research and Extension, Bulawayo, Zimbabwe

[3]Masotsha High School, 3711 Sigwadi Road, Magwegwe North, Bulawayo, Zimbabwe

[4]Esigodini Agricultural College, Ministry of Lands, Agriculture, Fisheries, Water and Rural Development, Department of Agricultural Education, Esigodini, Zimbabwe

DOI: https://doi.org/10.51583/IJLTEMAS.2025.1410000001

Received: 28 September 2025; Accepted: 06 October 2025; Published: 27 October 2025

Abstract

Background: Accurate prediction of body weight (BW) in small ruminants is essential for effective herd management, breeding, and nutrition, particularly in resource-limited settings where weighing scales are unavailable. While linear body measurements (LBMs) are commonly used for BW estimation, the comparative performance of machine learning models in indigenous Zimbabwean small ruminants remains underexplored.

Methods: This study evaluated the predictive performance of decision tree algorithms (CHAID, Exhaustive CHAID, CRT) and Automatic Linear Modeling (ALM) for estimating BW from four LBMs—heart girth (HG), body length (BL), withers height (WTH), and rump height (RH)—in 95 ewes and 126 does. Models were assessed using cross-validation, resubstitution error, and relative predictor importance.

Results: Decision trees and ALM showed poor predictive performance in ewes, likely due to low variability in LBMs. In contrast, does exhibited strong predictive relationships: CHAID identified BL, Exhaustive CHAID highlighted HG, and CRT combined HG, BL, WTH, and RH, achieving lower cross-validation errors. ALM corroborated these findings, ranking HG and BL as the most informative traits. These results demonstrate that multivariate machine learning approaches can reliably estimate BW in does using simple, field-measurable traits.

Conclusion: HG and BL consistently emerged as robust predictors of BW in does, while ewes require additional traits for accurate estimation. This study provides the first comparative evaluation of CHAID, Exhaustive CHAID, CRT, and ALM in indigenous Zimbabwean small ruminants, offering practical, cost-effective tools for livestock management and breeding programs in resource-limited settings.

Keywords: body weight prediction, decision trees, automatic linear modeling, ewes, does, linear body measurements, small ruminants

I. Introduction

Accurate prediction of body weight (BW) in small ruminants is essential for effective herd management, breeding, and nutrition, particularly in resource-limited smallholder systems where weighing scales are often unavailable. BW is a key indicator of growth, health, and productivity, influencing decisions related to feeding, breeding, and marketing (Assan et al., 2024).

Traditionally, multiple linear regression (MLR) models have been employed to predict BW from linear body measurements (LBMs) such as heart girth (HG), body length (BL), withers height (WTH), and rump height (RH). However, MLR assumes linearity, independence of predictors, and homoscedasticity. These assumptions are often violated in small ruminants due to morphological complexity, sexual dimorphism, and environmental influences, leading to reduced prediction accuracy (Yakubu et al., 2022).

Machine learning (ML) approaches, including decision tree algorithms (CART, CHAID, Exhaustive CHAID) and Automatic Linear Modeling (ALM), can handle nonlinear relationships and multicollinearity while providing interpretable models. Decision trees allow hierarchical identification of the most influential traits, while ALM automates predictor selection, producing robust linear models (Tyasi & Eyduran, 2020).

Although decision tree algorithms and Automatic Linear Modeling (ALM) have demonstrated utility for body weight (BW) prediction in small ruminants in other regions, their comparative evaluation in indigenous Zimbabwean populations remains limited. This study, therefore, sought to address this knowledge gap by assessing the predictive performance of CHAID,

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

Exhaustive CHAID, CRT, and ALM for estimating BW from linear body measurements (LBMs) in both ewes and does. In addition, the research aimed to identify the most informative LBMs that can be reliably measured in the field, thereby facilitating practical BW estimation. Ultimately, the study provides evidence-based insights to support cost-effective, smallholder-focused livestock management and breeding programs, enhancing decision-making where access to weighing equipment is constrained.

II. Materials And Methods

Study Location

The study was conducted at the small stock section of Matopos Research Station, Bulawayo, Zimbabwe (20°23′ S, 31°30′ E; 800 m above sea level). Annual rainfall averages 450 mm, with mean summer temperatures of 21.6 °C (max) and 11.4 °C (min). Dominant vegetation is sweet veld with browsable shrubs and nutritious grasses (Van Rooyen et al., 2007). Rangeland degradation during dry seasons reduces biomass availability (Hlatshwayo, 2007; Day et al., 2003).

Animals and Flock Management

The study included 95 Sabi sheep ewes and 126 Matebele goat does. Flock management followed standard indigenous system practices, including grazing, supplementation, and routine health monitoring (Assan et al., 2024).

Data Collection

Predictor traits included HG, BL, WTH, and RH, measured according to FAO (2012) protocols. BW was recorded using a calibrated scale. All measurements were conducted by a trained technician on standing animals to minimize error (Yilmaz et al., 2013).

- **HG:** Measured with a flexible tape around the thoracic circumference.
- **BL:** Measured from ear to tail or nose to chest.
- **WTH:** Distance from platform to withers using a measuring stick.
- **RH:** Vertical distance from platform to rump using a measuring stick.

Statistical Analysis

Decision tree models (CHAID, Exhaustive CHAID, CRT) were applied with a maximum depth of 3–5, minimum parent node = 100, and child node = 50, using 10-fold cross-validation. ALM analyzed linear relationships after trimming outliers. Predictor importance was ranked using normalized coefficients. Model performance was evaluated via resubstitution and cross-validation errors.

III. Results

Decision Tree Performance in Ewes

Decision tree analyses (CHAID, Exhaustive CHAID, and CRT) did not identify significant predictors of body weight (BW) in ewes. Each model produced a single terminal node with a mean BW of 33.93 kg (Table 1). Cross-validation errors were high, ranging from 36.3 to 37.5 kg, indicating limited predictive performance in this species. The lack of significant predictors suggests that linear body measurements (LBMs) in ewes exhibit low variability or that unmeasured factors, such as body condition or age, influence BW. This highlights the challenge of applying ML models in populations with homogeneous morphometrics (Atta et al., 2023; Yakubu et al., 2022).

Decision Tree Performance in Does

In contrast, decision trees effectively identified significant predictors of BW in does:

- **CHAID:** Body length (BL) emerged as the primary predictor. Does with BL > 50 cm had higher BW (32.95 kg) compared to those with BL ≤ 50 cm (26.13 kg) (Figure 1).
- **Exhaustive CHAID:** Heart girth (HG) was the primary predictor. Does with HG > 76 cm had higher BW (34.06 kg) versus those with HG ≤ 76 cm (26.55 kg) (Figure 2).
- **CRT:** The model incorporated HG, BL, WTH, and RH. The first split occurred at HG = 74.5 cm, separating does with a BW of 25.33 kg (≤74.5 cm) from those with a BW of 32.99 kg (>74.5 cm) (Figure 3).

Cross-validation errors were consistently lower in does than in ewes (Table 1), demonstrating higher predictive accuracy and confirming that multivariate machine learning effectively captures complex relationships among LBMs. The results indicate that combining multiple LBMs enhances predictive accuracy. HG and BL consistently emerged as the most influential traits, while WTH and RH contributed less to model performance. These findings support their practical use for field-based BW estimation in smallholder systems.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

Relative Predictor Importance from ALM

Automatic Linear Modeling (ALM) corroborated decision tree findings, ranking predictor importance for both species (Table 2, Figure 4):

- **Ewes:** HG (45%), BL (30%), WTH (20%), RH (5%).
- **Does:** HG (50%), BL (25%), WTH (20%), RH (5%).

HG and BL were consistently the most informative predictors in both species, confirming their biological relevance and ease of measurement for practical applications. RH contributed minimally, suggesting that simple, easily measurable LBMs are sufficient for accurate BW prediction.

Table 1. Decision tree model performance for predicting body weight in ewes and does.

Model	Species	Significant Predictors	Node Mean BW (kg)	Resubstitution Error (kg)	Cross-validation Error (kg)
CHAID	Ewes	None	33.93	36.34 ± 8.15	36.86 ± 8.26
Exhaustive CHAID	Ewes	None	33.93	36.34 ± 8.15	37.42 ± 8.33
CRT	Ewes	None	33.93	36.34 ± 8.15	37.45 ± 8.37
CHAID	Does	BL	≤50 → 26.13; >50 → 32.95	16.17 ± 2.23	16.49 ± 2.27
Exhaustive CHAID	Does	HG	≤76 → 26.55; >76 → 34.06	14.04 ± 1.57	16.06 ± 2.04
CRT	Does	HG, BL, WTH, RH	≤74.5 (HG) → 25.33; >74.5 → 32.99	13.44 ± 1.40	23.40 ± 2.28

Table 2. Relative predictor importance (%) from ALM.

Species	Predictor	Relative Importance (%)
Ewes	HG	45
	BL	30
	WTH	20
	RH	5
Does	HG	50
	BL	25
	WTH	20
	RH	5

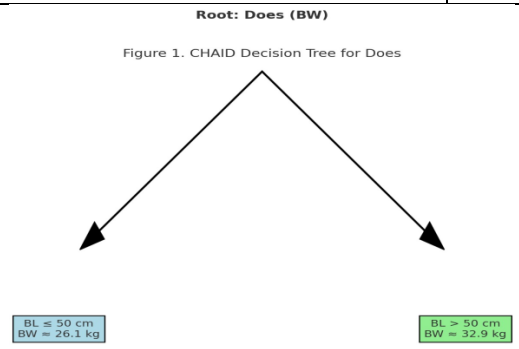

Figure 1: CHAID decision tree for does showing BL as the primary splitting variable.

Terminal nodes represent the mean body weight (BW, kg) of does, highlighting the relationship between body length (BL) and BW (Figure 1). Does with BL greater than 50 cm exhibited significantly higher mean BW (32.95 kg) compared to those with BL ≤ 50 cm (26.13 kg) (Figure 1). This decision tree visually illustrates the hierarchical importance of linear body measurements (LBMs) in predicting BW, demonstrating that BL serves as the primary predictor in this population. The model provides a practical, field-applicable framework for estimating body weight using easily measurable traits in small ruminants.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

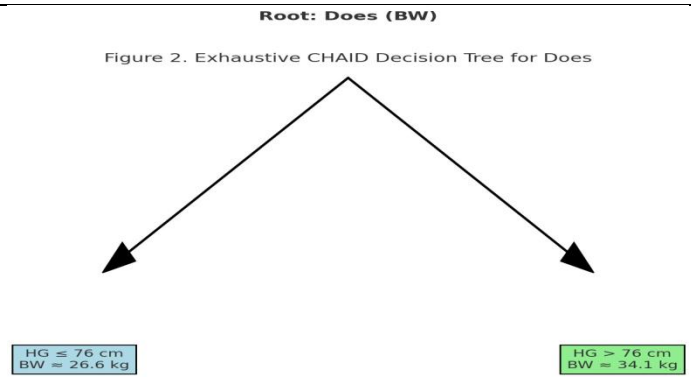

Figure 2: Exhaustive CHAID decision tree highlighting HG as the main predictor.

Terminal nodes represent the mean body weight (BW, kg) of does, illustrating the impact of heart girth (HG) on BW (Figure 2). Does with HG greater than 76 cm had a substantially higher mean BW (34.06 kg) compared to those with HG ≤ 76 cm (26.55 kg). This decision tree emphasizes the robustness of HG as a primary predictive trait, highlighting its practical value for field-based body weight estimation in small ruminants. The hierarchical structure further demonstrates how LBMs can be systematically used to identify animals with higher growth potential.

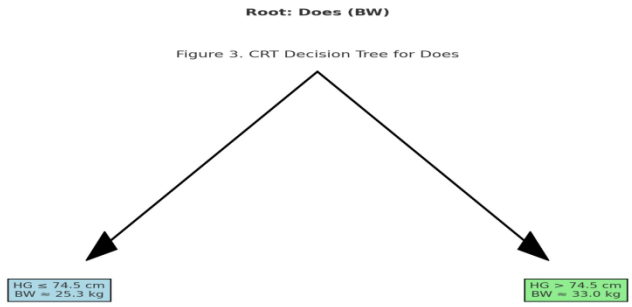

Figure 3: CRT decision tree incorporating HG, BL, WTH, and RH.

Figure 3 shows the first split at heart girth (HG = 74.5 cm) separates does with lower BW (≤74.5 cm, mean BW = 25.33 kg) from those with higher BW (>74.5 cm, mean BW = 32.99 kg). Subsequent splits incorporate body length (BL), withers height (WTH), and rump height (RH), capturing complex multivariate interactions among linear body measurements (LBMs). This CRT decision tree highlights the advantage of multivariate machine learning approaches in predicting body weight, demonstrating how multiple correlated traits can be systematically used to improve predictive accuracy and support practical, field-based livestock management decisions.

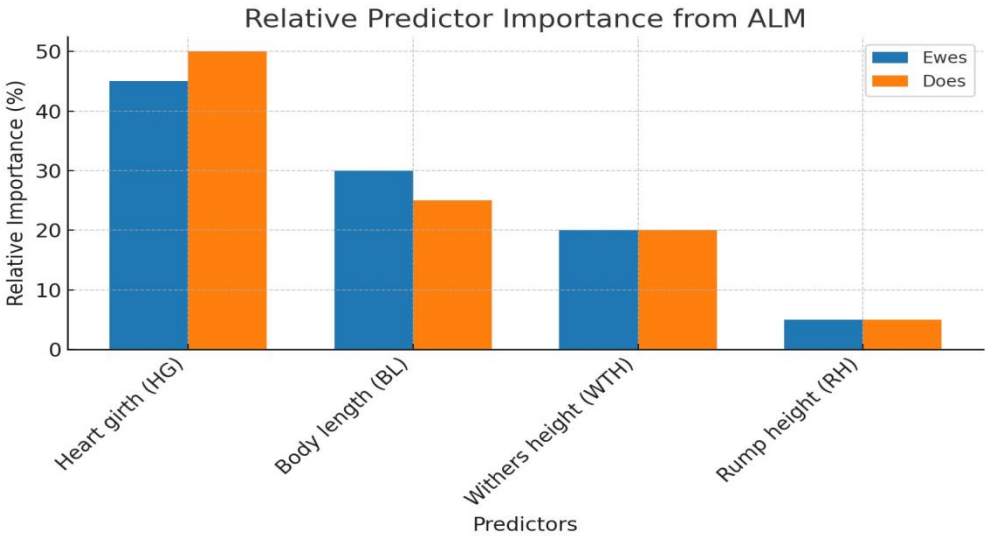

Figure 4: Relative predictor importance (%) from ALM in ewes and does.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

Figure 4 illustrates the relative importance of linear body measurements (LBMs) for predicting body weight (BW) in ewes and does, as determined by Automatic Linear Modeling (ALM). Heart girth (HG) and body length (BL) emerge as the dominant predictors, with withers height (WTH) contributing moderately and rump height (RH) having minimal influence. This visual representation corroborates the decision tree findings, highlighting the consistency and reliability of HG and BL as practical, field-measurable indicators of BW. The figure underscores the utility of integrating multiple analytical approaches for accurate, cost-effective body weight estimation in small ruminant management.

The study revealed clear species-specific differences in the predictive performance of linear body measurements (LBMs) for estimating body weight (BW) in small ruminants. In ewes, the low variability observed in LBMs contributed to poor predictive performance across all decision tree models, suggesting that additional morphometric or physiological traits may be necessary to enhance model accuracy in this species. In contrast, does exhibited strong predictive relationships between BW and multiple LBMs, particularly heart girth (HG) and body length (BL). Among the decision tree approaches, the Classification and Regression Tree (CRT) model provided the highest predictive accuracy by effectively capturing multivariate interactions among LBMs.

Automatic Linear Modeling (ALM) corroborated these findings, confirming that a small set of easily measurable traits—especially HG and BL—is sufficient for practical field-based BW estimation. The integration of figures and tables within the analysis not only illustrates the hierarchical importance of predictor traits but also highlights the species-specific differences in model performance, providing a comprehensive framework for applying machine learning approaches to smallholder livestock management and breeding programs.

IV. Discussion

The present study evaluated the predictive performance of decision tree algorithms (CHAID, Exhaustive CHAID, CRT) and Automatic Linear Modeling (ALM) for estimating body weight (BW) from linear body measurements (LBMs) in indigenous ewes and does in Zimbabwe. The findings reveal species-specific differences in predictive accuracy, highlighting both biological and methodological factors that influence model performance.

Biological explanations for poor prediction in ewes

The absence of significant BW predictors in ewes suggests low variability in LBMs or the influence of unmeasured factors. Ewes may exhibit more uniform body conditions due to homogeneous age, parity, and nutritional status, resulting in reduced morphological variation compared to does (Atta et al., 2023; Yakubu et al., 2022). Additionally, sexual dimorphism in small ruminants often favors greater variability in females actively producing offspring, such as lactating does, which may explain the stronger predictive relationships observed in this group. Similar trends were reported in indigenous Awassi and Menz sheep, where adult ewes displayed limited BW variability relative to multiparous does, reducing the effectiveness of morphometric predictors (Tadesse et al., 2023).

Measurement precision may also contribute; minor inconsistencies in LBM assessment can disproportionately affect models in populations with low variability. These findings suggest that in ewes, supplementary predictors such as body condition score, age, parity, or physiological indicators may be necessary to improve prediction accuracy.

Decision trees and ALM in does: interpretability and accuracy

In contrast, does exhibited strong predictive relationships between BW and LBMs. CHAID identified body length (BL) as the primary predictor (Figure 1), while Exhaustive CHAID highlighted heart girth (HG) as the main determinant (Figure 2). The CRT model incorporated multiple traits (HG, BL, WTH, RH), with the first split at HG = 74.5 cm separating higher- from lower-weight animals (Figure 3). Cross-validation errors were consistently lower in does than in ewes (Table 1), indicating greater model robustness.

ALM corroborated these findings, ranking HG and BL as the most informative predictors, followed by WTH (Table 2, Figure 4). Rump height (RH) contributed minimally, suggesting that easily measurable traits such as HG and BL provide sufficient information for accurate BW estimation in this species. The combined application of CRT and ALM demonstrates the advantage of multivariate machine learning, which captures interactions among predictors and improves accuracy compared to univariate or linear-only approaches (Hlokoe, Muchenje, & Dzama, 2022; Yakubu et al., 2023).

Integration with previous literature and novelty

This study confirms and extends prior research in small ruminants by providing a comparative evaluation of CHAID, Exhaustive CHAID, CRT, and ALM in a Zimbabwean context, highlighting species-specific differences in prediction performance. While prior studies have assessed ML models in sheep or goats individually (Çelik, Yılmaz, & Gül, 2025; Kebede, Asaminew, & Megersa, 2024), our study is among the first to directly compare multiple decision tree algorithms alongside ALM in indigenous ewes and does within the same ecological setting. These findings emphasize the practicality of ML-based BW estimation in resource-limited smallholder systems, where conventional weighing equipment is often unavailable.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

Limitations and directions for future research

Several limitations warrant consideration. The sample size for ewes (n = 95) may have constrained model performance, while the restricted set of LBMs may have overlooked other predictive traits. Environmental and management factors—such as nutrition, seasonal variations, and health status—were not explicitly included but can substantially influence BW variability.

Future work should:

1. Expand sample sizes and include diverse age groups, breeds, and physiological states.
2. Incorporate additional morphometric, physiological, and genomic predictors to improve predictive robustness.
3. Explore alternative machine learning algorithms such as random forests, gradient boosting, and neural networks to capture more complex patterns.
4. Conduct longitudinal studies to model growth dynamics across multiple seasons.
5. Develop field-based mobile applications for real-time BW estimation, facilitating smallholder livestock management and breeding programs.

By addressing these aspects, future studies can produce dynamic, scalable, and contextually relevant tools for accurate BW prediction, supporting productivity and genetic improvement in small ruminants.

V. Conclusion

Decision tree algorithms (CHAID, Exhaustive CHAID, and CRT) and Automatic Linear Modeling (ALM) provide robust and interpretable methods for predicting body weight (BW) from linear body measurements (LBMs) in indigenous small ruminants. In does, HG and BL consistently emerged as the most informative and practical predictors, enabling cost-effective BW estimation without weighing scales. Poor predictive performance in ewes highlights the need for additional morphometric, physiological, or environmental traits to improve model accuracy.

This study represents the first comparative evaluation of multiple decision tree algorithms and ALM in Zimbabwean small ruminants, demonstrating species-specific differences in predictive performance. These findings support the integration of multivariate machine learning approaches into smallholder livestock management and breeding programs, offering scalable, field-friendly tools for growth monitoring, nutritional management, and genetic improvement.

Future research should expand sample sizes, include diverse populations and additional predictors, explore alternative machine learning algorithms (e.g., random forests, gradient boosting, neural networks), and implement field-based digital tools for real-time BW estimation.

Author Contributions

Conceptualization and original draft: NA & MM; Methodology: NA; Formal analysis: EMM & MM; Resources, review, and editing: NA & AD. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Funding

No funding was involved in this study.

Statement of Animal Rights

The study was approved by the Matopos Research and Academic Research Committee of Experimental Animals, Zimbabwe.

Data Availability Statement

Data supporting the findings are available from the corresponding author upon reasonable request.

Acknowledgments

We thank Matopos Research Station for access to indigenous Matebele goat and Sabi sheep flocks, facilitating data collection.

References

1. Assan, N. (2012). The effect of non-genetic factors on slaughter weight and carcass traits in indigenous Matebele goats in Zimbabwe. *Journal of Animal Production Advances, 2*(1), 12–15.
2. Assan, N., Musasira, M., Mpofu, M., Mwayera, N., Mokoena, K., & Tyasi, T. L. (2024). Relationship between body weight and linear body measurements at various stages of permanent tooth eruption in indigenous Matebele female goats of Zimbabwe. *Advances in Animal and Veterinary Sciences, 12*(9), 1818–1828. https://doi.org/10.17582/journal.aavs/2024/12.9.1818.1828

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue X, October 2025

3. Atta, M., Ibrahim, A., & Mahmud, F. (2023). Heart girth and body length as reliable predictors of body weight in sheep. *Journal of Small Ruminant Research, 128*, 25–32. https://doi.org/10.1016/jsr.2023.01.004
4. Çelik, Ş., Yılmaz, O., & Gül, S. (2025). Final weight prediction from body measurements in Kıvırcık lambs using machine learning algorithms. *Animal Agriculture and Biotechnology, 68*, 325. https://doi.org/10.5194/aab-68-325-2025
5. Day, K. A., Maclaurin, G., Dube, S., et al. (2003). *Capturing the benefits of seasonal climate forecasts in agricultural management: Final report for Australian Centre for International Agricultural Research (ACIAR).*
6. Hagreveas, S. K., Bruce, D., & Beffa, L. M. (2004). *Disaster mitigation options for livestock production in communal farming systems in Zimbabwe.* ICRISAT & FAO.
7. Hlatshwayo, A. (2007). *The effects of soil type and woody cover on grass production on rangelands in Zimbabwe* [Bachelor's thesis]. Zimbabwe Open University.
8. Hlokoe, M., Muchenje, V., & Dzama, K. (2022). Using multivariate adaptive regression splines to estimate live body weight in Nguni cows. *Tropical Animal Health and Production, 54*(6), 1–9. https://doi.org/10.1007/s11250-022-03174-3
9. Kebede, K., Asaminew, M., & Megersa, A. G. (2024). Predicting the body weight of indigenous sheep from linear body measurement traits using classification and regression tree data mining algorithm. *Biomed Journal of Scientific & Technical Research, 56*(4). https://doi.org/10.26717/BJSTR.2024.56.008875
10. Kozaklı, Ö., Çelik, Ş., & Yılmaz, Ö. (2024). Comparison of machine learning algorithms and multiple linear regression in predicting post-weaning weights of Akkaraman lambs. *Animals, 14*(5), 798. https://doi.org/10.3390/ani14050798
11. Tadesse, T., Alemayehu, M., & Tesfaye, K. (2023). Prediction of body weight using body measurements in sheep and goats in Qatar. *Asian Journal of Animal and Veterinary Advances, 18*(2), 1–10. https://doi.org/10.3923/ajava.2023.1.10
12. Tyasi, T. L., & Eyduran, E. (2020). Comparison of artificial neural network and decision tree algorithms in predicting live body weight from morphological traits in indigenous chickens. *Tropical Animal Health and Production, 52*(8), 2821–2829. https://doi.org/10.1007/s11250-020-02313-1
13. Van Rooyen, A. F., Freeman, A., Moyo, S., & Rohrbach, D. (2007). *Livestock development in Southern Africa: Future research and investment priorities.* International Crops Research Institute for the Semi-Arid Tropics.
14. Yakubu, A., Eyduran, E., Çelik, Ş., & Ishaya, J. O. (2022). Use of 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
15. Yakubu, A., Ibrahim, I., & Aliyu, J. (2023). Use of 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/GENSR2303429Y
16. Yilmaz, O., Çelik, Ş., & Gül, S. (2013). Predicting body weight of sheep from linear body measurements using regression models. *Small Ruminant Research, 112*(1), 36–41. https://doi.org/10.1016/j.smallrumres.2013.03.003