Proceedings of 30th Annual Conference of ASAN 2025, Abuja, Nigeria (30-34)

Predicting Body Weight in Pullets Using Machine Learning Techniques

Aliyu, A.M.[1], Yahuza, S.[1], Abdulhamid, S.U[2]. Rotimi, E. A.[1], Bature, I.[1], Bichi, A.A[1]., Z. S. Jibia[1], and Mustapha, Y[3].
[1] Department of Animal Science, Faculty of Agriculture, Federal University Dutsin-Ma.
[2] Department of Animal Production Technology, Audu Bako College of Agriculture, Dambatta
[3] Center for Innovation and Training in Animal Husbandry, Nigerian Institute of Animal Science, Kano State
* Corresponding Author: Aliyu, A.M., amaliyu@fudutsinma.edu.ng. +2348064209979

Abstract
This study explores the use of machine learning to predict the body weight of pullets based on non-invasive morphometric traits. Data from 100 Isah Brown birds, including body weight, length, girth, and limb dimensions, were analysed using four models: Gaussian Process Regression (GPR), Linear Regression (LR), Regression Tree (RT), and Kernel Regression (KR) and evaluated using key metrics: R-Squared (R^2), Determination Coefficient (R), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Error (MAE). GPR delivered the highest accuracy (R^2 = 0.99 testing, 0.95 validation) with the lowest RMSE values, outperforming other models. To determine the most effective model, performance outcomes were compared across all algorithms. GPR ultimately demonstrated the highest predictive accuracy, indicating its suitability for real-world poultry applications.

Description of Problem

Accurate estimation of body weight in pullets is essential for effective poultry management, influencing feed optimisation, health monitoring, and market readiness [1]. With recent advances in artificial intelligence, machine learning (ML) has emerged as a promising alternative for non-invasive weight estimation in poultry farming. ML algorithms are capable of identifying complex, non-linear relationships within morphometric data such as body length, girth, and limb dimensions, enabling accurate prediction of body weight [2].

Gaussian Process Regression (GPR) is a non-parametric, probabilistic modeling approach capable of capturing complex non-linear patterns while providing both predictive estimates and associated uncertainty measures [3]. Its ability to model intricate relationships with limited data makes it particularly effective for high-precision poultry body weight estimation with minimal error. Linear Regression (LR) is a parametric model based on the assumption of a linear relationship between independent and dependent variables. Although it offers reasonable accuracy and interpretability, its inability to account for non-linear biological processes inherent in poultry growth dynamics reduces its suitability for this application. Regression Tree (RT) is a decision-based, non-parametric algorithm that recursively splits data into subsets using hierarchical decision rules. It demonstrated predictive accuracy comparable to GPR, efficiently capturing complex trait interactions, which enhances its utility for body weight prediction across heterogeneous poultry populations [4]. Kernel Regression (KR) utilizes kernel functions to approximate non-linear relationships. However, empirical results in this study indicate subpar performance, potentially due to model overfitting or inadequate kernel tuning, thus diminishing its practical reliability for poultry weight estimation

Previous studies have demonstrated the success of ML models in precision livestock applications, particularly in poultry, where image and sensor-based techniques can yield real-time insights with high accuracy [5], [6]. However, comparative evaluations of different ML models using conventional body measurements remain limited. This study aims to evaluate the predictive performance of four ML techniques, such as Gaussian Process Regression (GPR), Linear Regression (LR), Regression Tree (RT), and Kernel Regression (KR), using morphometric data collected from pullets. The objective is to identify the most effective and practical model for on-farm deployment, particularly in resource-constrained settings. By integrating ML into poultry monitoring systems, this research contributes to the ongoing shift toward precision agriculture in Nigeria and similar developing economies. It provides an evidence-based pathway for enhancing productivity, improving animal welfare, and reducing reliance on invasive weighing practices.

Materials and Methods

This study was conducted at the Professor Lawal Abdu Saulawa Livestock Research and Teaching Farm of the Department of Animal Science, Federal University Dutsin-Ma, Katsina State, Nigeria (latitude 12°29'E). A total of 100 eight-week-old pullets reared under conventional floor housing and fed a standard commercial diet were used.

Morphometric measurements were collected weekly for six weeks from each bird. The features recorded included body weight using a CAS weighing scale (CAS Corporation, South Korea), Body length (BL), body girth (BG), drumstick length (DL), thigh length (TL), shank length (SL), and wing length (WL). Body weights were measured using a Camry mechanical scale (20 kg capacity), while linear body traits were measured using an electronic digital calliper. Body length (BL) was measured from the back of the neck to the tail. Breast girth (BG) was the circumference of the chest, from the right-wing joint across the breast to the left-wing joint. Thigh length (TL) was the distance from the hock joint to the pelvic joint. Shank length (SL) was measured from the shank joint to the end of the foot. Wing length (WL) was the distance between the tips of the longest primary feathers when the wings were stretched [7].

Prior to model development, all data underwent preprocessing to address missing values and ensure uniform scaling. The dataset was split, allocating 70% for training and 30% for validation. The algorithms, once trained, were assessed with validation datasets to confirm their predictive accuracy. Four machine learning algorithms-Gaussian Process Regression (GPR), Linear Regression (LR), Regression Tree (RT), and Kernel Regression (KR) were trained to predict body weight using the morphometric features.

Model performance was evaluated using key regression metrics: R-Squared (R^2), Determination Coefficient (R), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Error (MAE). A Complete Randomised Design (CRD) guided the experimental setup. Statistical analysis and model development were performed using MATLAB (R2023a) with cross-validation employed to enhance the reliability of results.

Results and Discussion

Table 1 provides a rigorous quantitative assessment of four machine learning algorithms. The GPR algorithm demonstrated exceptional performance, achieving near-perfect training accuracy (R^2 = 0.990) and maintaining strong validation results (R^2 = 0.956). This model exhibited minimal prediction errors across all metrics, including the lowest RMSE (40.04 training; 80.83 validation) and MAE (25.45 training; 46.71 validation) values. The remarkably low MAPE scores (0.112% training; 0.215% validation) further confirm its precision. These findings corroborate the theoretical advantages of GPR described by [8], particularly its Bayesian probabilistic framework that effectively handles non-linear relationships in limited datasets while providing uncertainty estimates. The model's non-parametric architecture proved ideal for capturing the complex biological variability inherent in poultry morphometrics.

Table 1: Body Weight Prediction using Different Models and Key Performance Indicators

Model Type	Status	R^2	R	MAPE	RMSE	MSE	MAE
GPR	Training	0.990	1.000	0.112	40.039	1603.167	25.450
	Validation	0.956	0.999	0.215	80.834	6534.169	46.715
LR	Training	0.983	0.999	0.024	5.0339	50.792	2579.911
	Validation	0.940	0.999	0.1604	31.9180	94.358	8903.440
RT	Training	0.990	0.999	2.476	39.894	1591.549	28.751

	Validation	0.934	0.998	3.472	99.269	9854.431	56.821
KR	Training	0.967	0.989	11.979	71.872	2512.298	5165.589
	Validation	0.928	0.973	62.963	103.302	975.609	10670.930

NOTE: R^2 (R-Squared), R (Determination Coefficient), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Squared Error), MSE (Mean Squared Error), MAE (Mean Absolute Error), (Gaussian Process Regression), LR (Linear Regression), RT (Regression Tree), and (Kernel Regression).

The LR and RT, the parametric LR model showed competent performance (R^2 = 0.983 training; 0.940 validation), though with marginally higher error metrics (RMSE = 50.79; MAE = 42.50 training) compared to GPR. As noted by [8] Its constrained linear formulation limits its ability to model the non-linear biological relationships present in the data. The RT algorithm, a non-parametric decision-based approach, achieved training accuracy comparable to GPR (R^2 = 0.990; RMSE = 39.89; MAE = 28.75), but showed reduced generalization capability during validation (R^2=0.934; RMSE=99.27; MAE=56.82). This performance pattern, consistent with findings by [9], suggests some degree of overfitting while still effectively capturing complex feature interactions. The KR model exhibited the poorest predictive performance, with the lowest validation R^2 (0.928) and substantially elevated error metrics (RMSE=103.30; MAE=71.86; MAPE=62.96%). Despite its theoretical capacity for non-linear modeling, the observed performance deficiencies align with [10]'s findings regarding kernel method limitations. The model's inadequate performance likely stems from suboptimal kernel parameterization and pronounced overfitting, rendering it unsuitable for practical poultry weight prediction applications where high precision is required.

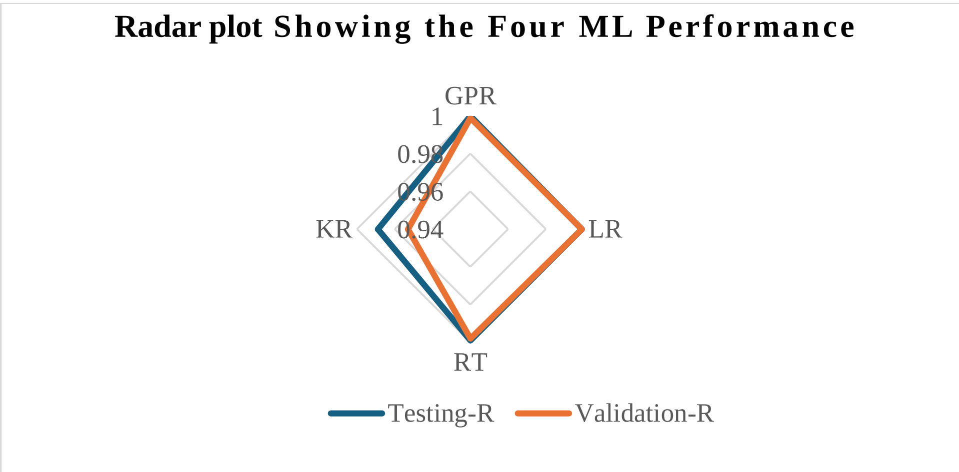

Figure 1: Relationship Between Observed and Predicted Body Weights of Pullets Across Machine Learning Models

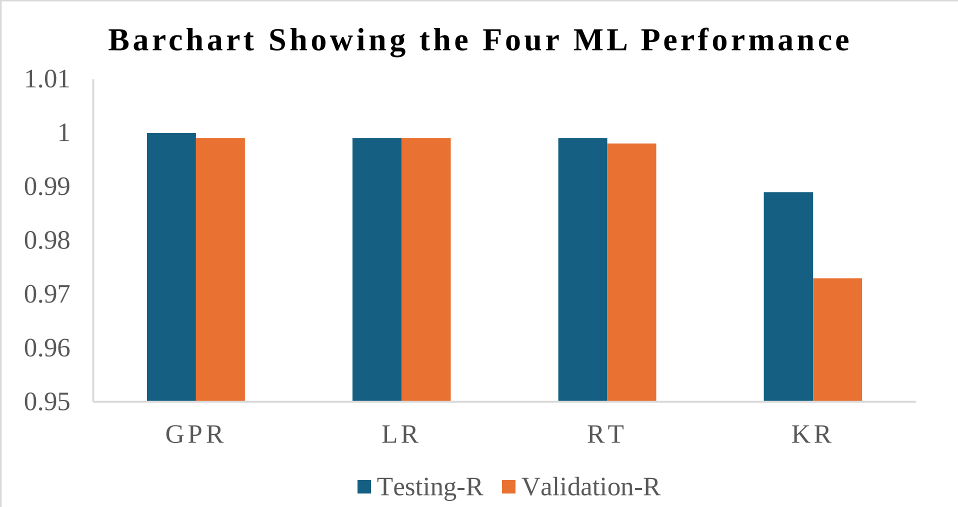

Figure 2: Comparative Error Metrics (RMSE, MAE, MAPE) of Gaussian Process Regression, Linear Regression, Regression Tree, and Kernel Regression Models

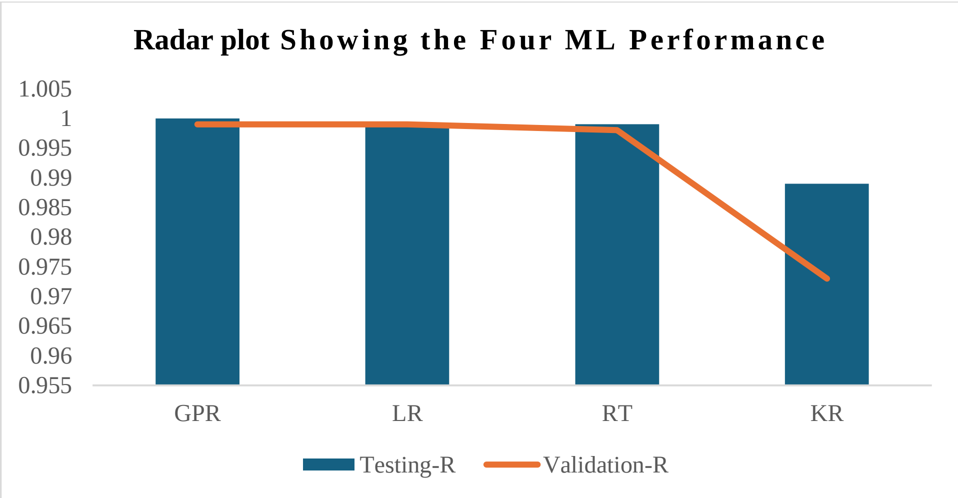

Figure 3: Validation Performance of Machine Learning Models: R^2 Values and Error Distributions for Pullet Body Weight Prediction

The figures reinforce the findings summarized in Table 1. As the table shows, Gaussian Process Regression (GPR) consistently achieved the highest predictive accuracy (training $R^2 = 0.990$; validation $R^2 = 0.956$) with the lowest error margins (RMSE and MAE). This superiority is visually evident in Figure 1, where GPR's predicted values align closely with observed weights, indicating strong generalization capability. Figure 2 further highlights that GPR maintained minimal error values compared to Linear Regression (LR), Regression Tree (RT), and Kernel Regression (KR). While LR demonstrated moderate performance with slightly higher error levels, RT performed competitively during training but showed overfitting in validation, which is visible in its broader error spread in Figure 3. KR clearly underperformed, exhibiting both the lowest R^2 (0.928 validation) and the highest error metrics (MAPE = 62.96%), a trend that the figures emphasize through its inconsistent predictive pattern. Overall, the visual evidence provided by the figures corroborates the statistical evidence in Table 1, confirming GPR as the most robust and reliable model for pullet body weight prediction.

Conclusion and Application

GPR demonstrated superior predictive performance, achieving exceptional training accuracy (R^2=0.990; RMSE =40.04; MAE

=25.45). Validation confirmed its robustness, maintaining high explanatory power (R^2=0.956) with acceptable error margins (RMSE =80.83; MAE =46.72). GPR's effectiveness is attributed to its non-parametric Bayesian framework, which models complex nonlinear patterns and quantifies prediction uncertainty, establishing it as the most reliable algorithm for pullet weight estimation in this study. Suboptimal Model-KR exhibited the poorest predictive outcomes, particularly in validation, showing marginal explanatory power (R^2=0.928) and substantially elevated error metrics (RMSE =103.30; MAE =71.86). The high mean absolute percentage error (MAPE =62.96%) indicated systematic inaccuracies.

Future Recommendations

Future research should focus on integrating image-based morphometrics and real-time sensor data to improve model generalizability and automation in commercial poultry operations. Furthermore, validating these models with larger, more diverse datasets and various chicken breeds will bolster their robustness and applicability. Exploring kernel parameter fine-tuning and ensemble learning methods may also further optimize performance, leading to more scalable solutions.

References

[1] A. S. George and A. S. H. George, "Optimizing poultry production through advanced monitoring and control systems," *Partners Univers. Int. Innov. J.*, vol. 1, no. 5, pp. 77–97, 2023.

[2] K. Kebede, M. Asaminew, and A. G. Megersa, "Predicting the body weight of indigenous sheep from linear body measurement traits using classification and regression tree data mining algorithm," *Biomed. J. Sci. Tech. Res*, vol. 56, no. 4, pp. 48232–48239, 2024.

[3] J. Liu, "Gaussian process regression for non-Gaussian data." Newcastle University, 2022.

[4] M. M. Pichler, "Machine Learning and Deep Learning in Ecology–from predictions to mechanistic inference." 2024.

[5] N. Li, Z. Ren, D. Li, and L. Zeng, "Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming," *animal*, vol. 14, no. 3, pp. 617–625, 2020.

[6] L. O. Tedeschi, P. L. Greenwood, and I. Halachmi, "Advancements in sensor technology and decision support intelligent tools to assist smart livestock farming," *J. Anim. Sci.*, vol. 99, no. 2, p. skab038, 2021.

[7] 2018 FAO, "Guidelines on methods for estimating livestock production and productivity," 2018. [Online]. Available: https://openknowledge.fao.org/server/api/core/bitstreams/a7548126-efdc-459f-9742-e86ac3b781bc/content

[8] E. Chan, "Prediction of Beef Cattle Body Weight Based on Morphometric Measurements Extracted from 2D Images." University of Guelph, 2023.

[9] R. Portillo-Salgado *et al.*, "Prediction of egg weight from external egg traits of Guinea Fowl using multiple linear regression and regression tree methods," *Brazilian J. Poult. Sci.*, vol. 23, no. 03, p. eRBCA-2020, 2021.

[10] T. L. Tyasi *et al.*, "Classification and regression tree (CRT) analysis to predict body weight of Potchefstroom koekoek laying hens," *Adv. Anim. Vet. Sci*, vol. 8, no. 4, pp. 354–359, 2020.