COMPARISON OF FIVE MACHINE LEARNING ALGORITHMS FOR PREDICTION OF BODY WEIGHTS OF CATTLE

[*1,4]Iyiola-Tunji, A.O., [2]Muniru, I.O., [3,5]John, P.A. and [4]Salihu, S.S.

[1]National Agricultural Extension and Research Liaison Services, Ahmadu Bello University, Zaria, Kaduna State

[2]Department of Biomedical Engineering, Faculty of Engineering, University of Ilorin, Ilorin, Kwara State

[3]Biotechnology Advance Research Centre, Sheda, FCT, Abuja

[4]Department of Animal Science and Range Management, Modibbo Adama University, Yola, Adamawa State

[5]Department of Animal Science, Ahmadu Bello University, Zaria, Kaduna State

Corresponding author: tunjiyiola@yahoo.com; +2347034234136

ABSTRACT

This study compares five machine-learning algorithms for estimating cattle body weights using the linear body measurements of the animals. Bunaji, Rahaji, and Sokoto Gudali breeds of cattle were used to gather information on body weight and linear body measurements. One hundred of each breed were used to sample the animals in Adamawa State's Yola North Local Government Area. Approved reference marks for cattle by FAO were used for the linear body measurements. All the variables of body weight and linear body measurements were entered into the Python programming language. Five machine learning algorithms were compared to the model's performance. The techniques are K-Nearest Neighbor Regressor, Support Vector Regressor, Extreme Gradient Boosting (XGB), Random Forest (RF), Neural Network (NN), and Support Vector Regressor (SVR) (KNN). The proportion of variation in the data set was explained by the predictor variables up to 95.7% and 95.4%, respectively, according to the coefficient of determination (R^2) for the two algorithms. The Support Vector Regressor's (SVR) performance was the least effective, with the lowest R^2 and greatest MSE, RMSE, and MAE values. Extreme Gradient Boosting (XGB), a machine learning algorithm, is the most effective at predicting the body weight of the three Nigerian cattle breeds studied. This research should lead to the creation of computer applications that will make determining cattle body weight easier. More data from Nigerian cattle breeds should be added to the model.

Keywords: Cattle, body weight, linear body measurements, Machine Learning, Artificial Intelligence

INTRODUCTION

With the development of artificial intelligence (AI), many scientific and technological procedures could now be automated. It is a versatile tool that allows users to reconsider how they combine information, evaluate data, and use the ensuing insights to enhance decision-making. The thoughts created by AI have also enhanced the realm of agriculture. Helping farmers and stakeholders make better decisions by implementing sustainable agricultural practices is vital for anticipating effective solutions, especially when using digital technologies like the Internet of Things (IoT), Artificial Intelligence (AI), and cloud computing (Ben Ayed and Hanana, 2021). Machine and deep learning, which are subsets of AI, are also widely employed to enhance data processing and information collection for agriculture. The productivity of the livestock subsector of agriculture would increase with automated measurements of specific data, especially from developing countries like Nigeria. This study, on the other hand, compares five machine-learning techniques for estimating cattle body weights using the linear body measurements of animals.

MATERIALS AND METHODS

Bunaji, Rahaji, and Sokoto Gudali breeds of cattle were used to gather information on body weight and linear body measurements. One hundred of each breed were used to sample the animals in Adamawa State's Yola North Local Government Area . Using the random sample method, the cattle breeds of Rahaji, Bunaji, and Sokoto Gudali were selected because a sizable proportion of agropastoralists rear them. Approved body reference marks for cattle by FAO (2012) were used. The following are descriptions of the body measurements: BWT: Body weight using a weighing band with a kilogram marking, body weight was calculated (kg). Horn length (HLT) is the measurement of the horn's length from tip to base (cm). Face length (FLT) is the measurement from the horny part of the forehead to the upper lip (cm). Ear length (ELT) is the distance between the zygomatic arch's base and the ear's tip (cm). The distance between the outside edges of both eyes is used to calculate head width (HWD) (cm). Neck length (NLT): Defined as the distance between the base of the top shoulder and the cervical vertebra (cm). Neck circumference (NCR): The measurement of the neck's circumference at its middle (cm).

Foreleg length (FLL) is the measurement from the olecranon process's extremities to the coronet's mid-lateral point (cm). Height at withers (HTW): This measurement refers to the height of the animal measured from the ground to the point of its withers, which is the highest point of the *processus spinalis* of the thoracic vertebra (cm). Chest girth (CGT): This is the measurement of the circumference of the torso just behind the front legs (cm). Body length (BLT) is the measurement from the pin bone to the tip of the shoulder (lateral tuberosity of the humerus) (cm). The length around the cannon is used to calculate the loin girth (LGT). RHT: Rump height is the height of the rump from the ground (cm). measured as the length of the hind limb, the rear leg length (RLL) (cm). Measured as the distance between the tip of the tail and the base end of the tail touching the body, tail length (TLT) (cm).

For the purpose of creating predictive models for cattle body weight, all the variables were entered into the Python programming Language. Five machine learning algorithms were compared to the model's performance. The techniques are K-Nearest Neighbor Regressor (KNN), Support Vector Regressor (SVR), Extreme Gradient Boosting (XGB), Random Forest (RF) and Neural Network (NN).

RESULTS AND DISCUSSION

Table 1 and Figures 1 to 5 display the model's performance in forecasting the body weight of the three breeds of cattle (Bunaji, Rahaji, and Sokoto Gudali). The predictive model performed best with the XGB and RF algorithms. The proportion of variation in the data set was explained by the predictor variables up to 95.7% and 95.4%, respectively, according to the coefficient of determination (R^2) for the two algorithms. The Mean Square Error (MSE), which measures how near an estimate or forecast is to reality, is used to compare estimates and projections to actual data. The XGB body weight prediction model got the lowest MSE when compared to the other ML algorithms that were used.

In terms of magnitude, MAE is often comparable to RMSE but significantly smaller. The scale used by the MAE is the same as that of the measured data. Since this is a scale-dependent accuracy metric, it cannot be used to compare series that have been scaled differently. In this investigation, Random Forest (RF) had the highest prediction accuracy based on MAE. The Support Vector Regressor's (SVR) performance was the least effective, with the lowest R^2 and highest MSE, RMSE, and MAE values.

With the exception of the neural network, Ruchay *et al*. (2022) indicated that four of the machine learning methods utilized in this study were capable of producing predictive models for live weight, as had been demonstrated for Hereford cattle. XGB algorithm was previously described by Hamadani *et al*. (2022) as an effective algorithm for the prediction of sheep genetic merit, with a 0.915 correlation coefficient, low prediction error, and R^2 value of 0.88.

Table 1: Comparison of predictive model algorithms for Nigerian cattle breeds' body weight

Algorithm	R^2	MSE	RMSE	MAE
Extreme Gradient Boosting (XGB)	0.957	410.324	20.256	8.769
Random Forest (RF)	0.954	438.420	20.938	6.796
K-Nearest Neigbour Regressor (KNN)	0.882	1129.245	33.604	26.011
Neural Network (NN)	0.647	3368.861	58.042	40.648
Support Vector Regressor (SVR)	0.304	6651.867	81.559	67.626

R^2 = coefficient of determination, MSE = mean square error, RMSE = root mean square error, MAE = mean absolute error

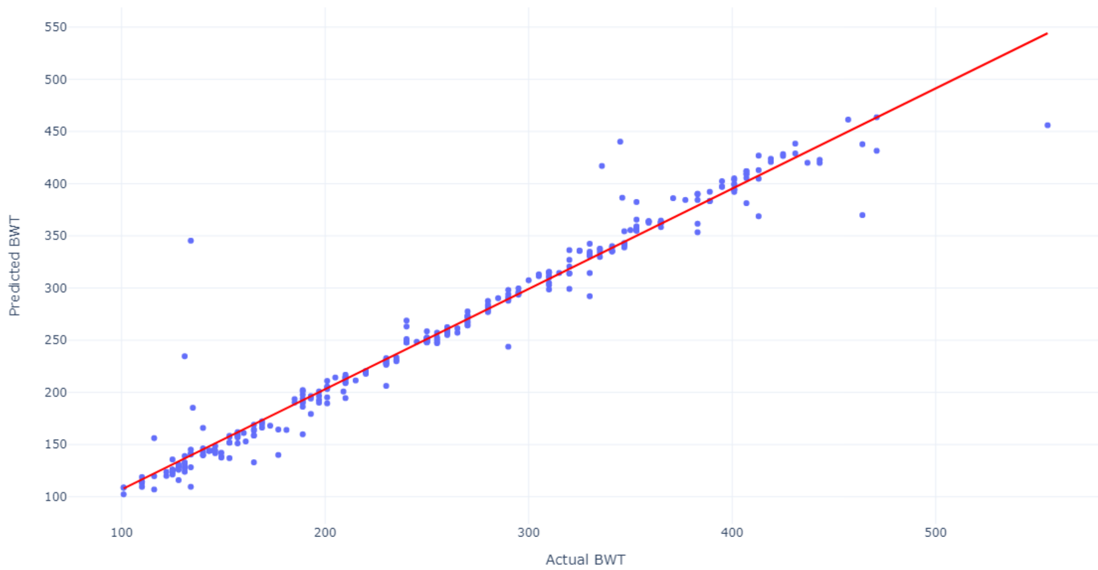

Figure 1: Predictive performance of Extreme Gradient Boosting (XGB) for three breeds of cattle

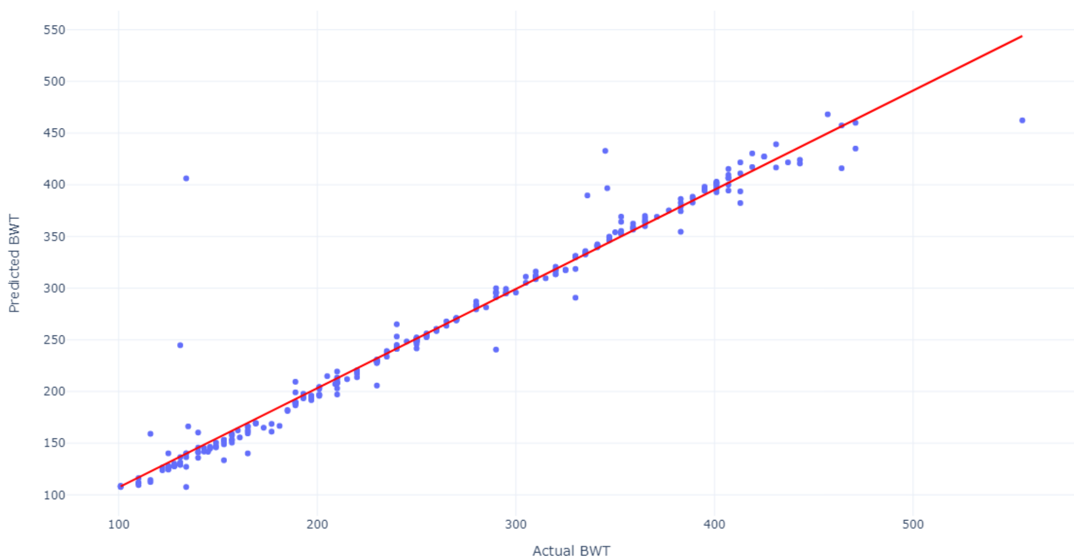

Figure 2: Predictive performance of Random Forest for three breeds of cattle

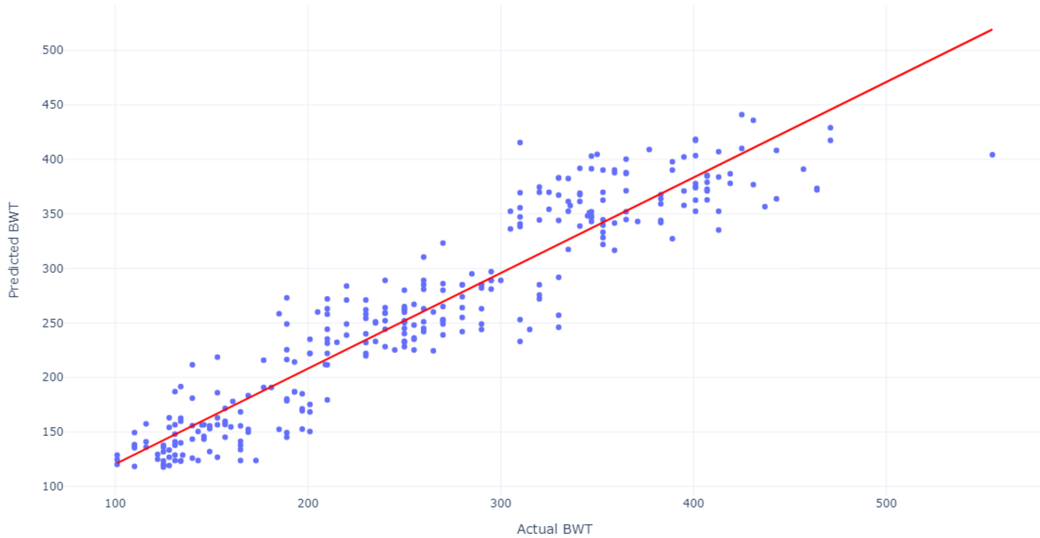

Figure 3: Predictive performance of K-Nearest Neighbour Regressor (KNN) for three breeds of cattle

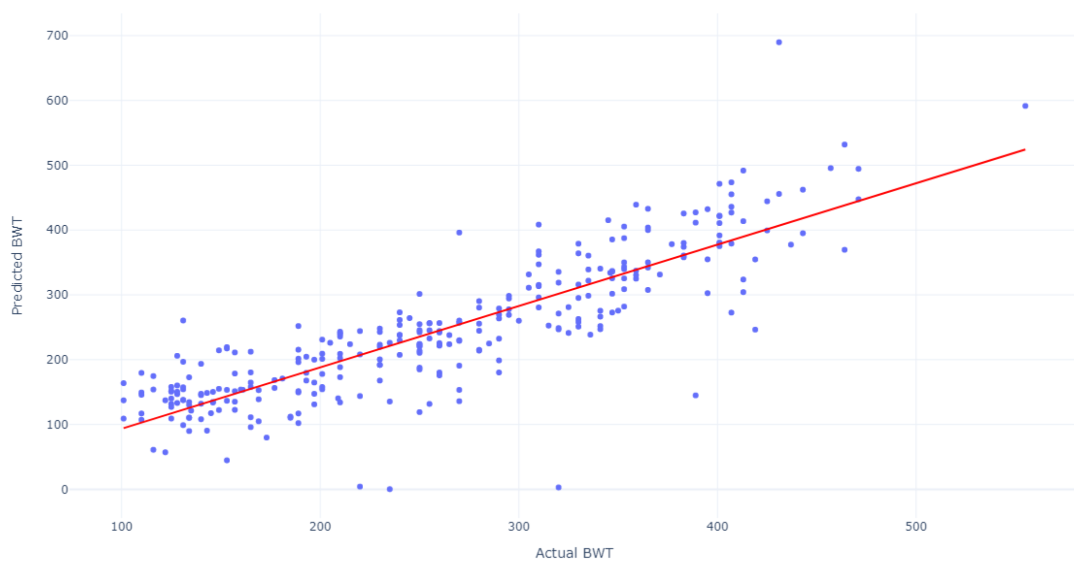

Figure 4: Predictive performance of Neural Network for three breeds of cattle

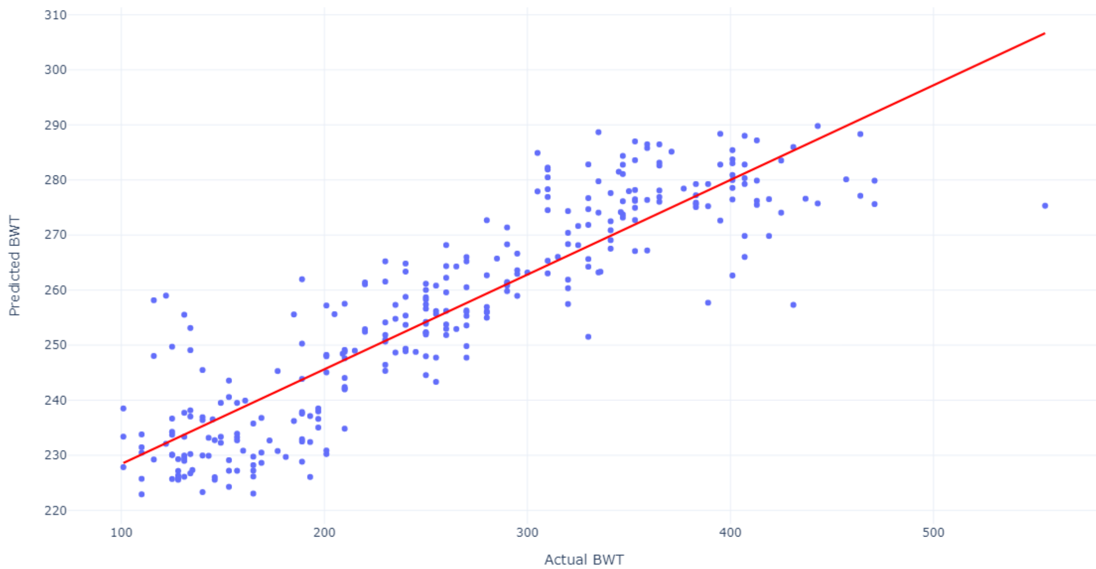

Figure 5: Predictive performance of Support Vector Regressor (SVR) for three breeds of cattle

CONCLUSION AND RECOMMENDATION

Extreme Gradient Boosting (XGB), a machine learning algorithm, is the most effective at predicting the body weight of the three Nigerian cattle breeds studied. This research should lead to the creation of computer applications that will make determining cattle body weight easier. More data from Nigerian cattle breeds should be added to the model.

REFERENCES

Ruchay, A., Kober, V., Dorofeev, K. Kolpakov, V., Dzhulamanov, K., Kalschikov, V.& Guo, H. (2022). Comparative analysis of machine learning algorithms for predicting live weight of Hereford cows. *Computers and Electronics in Agriculture*, *195*, 106837. https://doi.org/10.1016/j.compag.2022.106837.

Ben Ayed, R. and Hanana, M. (2021). Artificial Intelligence to Improve the Food and Agriculture Sector", *Journal of Food Quality*, *2021*, 5584754. https://doi.org/10.1155/2021/5584754

Food and Agricultural Organization of the United Nation. (2012). Phenotypic Characterization of Animal Genetic Resources. Food and Agricultural Organization of the United Nations. Animal Health and Production Guideline No.11. Rome. Retrieved on 15th October, 2022;

Hamadani, A., Ganai, N.A.& Mudasir, S. (2022). Comparison of artificial intelligence algorithms and their ranking for the prediction of genetic merit in sheep. *Scientific Report*, 12, 18726. https://doi.org/10.1038/s41598-022-23499-w