Journal of Animal & Plant Sciences, 34(2): 2024, Page: 336-350
ISSN (print): 1018-7081; ISSN (online): 2309-8694
https://doi.org/10.36899/JAPS.2024.2.0721

EXPLORING DATA MINING ALGORITHMS FOR PREDICTING DUCK EGG WEIGHT BASED ON EGG QUALITY CHARACTERISTICS

L. Dahloum[1,*], Q. Benameur[1], and A. Yakubu[2]

[1]Laboratoire Agrobiotechnologie, Ressources génétiques, et Modélisation (AGROBIOGEN), Abdelhamid Ibn Badis University, PO. Box 188, Mostaganem 27000, Algeria
[2]Department of Animal Science, Faculty of Agriculture, Nasarawa State University, Keffi, Shabu-Lafia Campus, P.M.B. 135, Lafia 950101, Nigeria
[*]Corresponding author's email: Lahouari.dahloum@univ-mosta.dz

ABSTRACT

The present investigation aimed to compare the performance of twomachine learning algorithms, Artificial Neural Network (ANN), and Classification and Regression Tree (CART), alongside the Automatic Linear Modelling (ALM), and the traditional Multivariate Linear Regression model (MLR) to predict the egg weight (EWT) of Mallard duck from some egg traits including egg length (EL), egg width (EWd), egg shape index (ESI), eggshell weight (ESW), albumen weight (AW), albumen height (AH), yolk weight (YW), yolk height (YH), yolk diameter (YD), and Haugh unit (HU). The Pearson correlation between observed and predicted values (*r*), coefficient of determination (R^2), adjusted coefficient of determination (R^2_{adj}), Root Mean Squared Error (RMSE), and Relative Approximation Error (RAE) were used to estimate model performance. EWT hada strong correlation with egg dimensions (EL and EWd, r=0.752 and 0.790, respectively), AW (r= 0.815), and YW (r= 0.784). The R^2_{adj} values were 0.981, 0.970, 0.964 and 0.897, for ANN, ALM, MLR, and CARTmodels, respectively. The lowest RMSE was found for ANN (0.753), while the highest RMSE was observed for CART (1.778). Overall, the ensemble models proposed in this study yielded similar results, with the ANN algorithm showing a marginally superior predictive performance compared to ALM, CART, and MLR models. This finding suggests that ANN could be considered the most suitable for the prediction of egg weight in Mallard duck.

Keywords: egg weight, Mallard duck, artificial neural network, automatic linear modelling, classification and regression tree, multivariate linear regression.

Published first online January 29, 2024 Published final March 31, 2024

INTRODUCTION

Eggs have acquired greater importance as an inexpensive and high-quality protein (Almeida *et al.,*2020). Eggs are common ingredients used by the food industry, predominantly for their taste and functional properties. Additionally, eggs contain numerous biologically active compounds that remain largely unexplored, but they hold significant potential for applications in the medical, pharmaceutical, and biotechnological industries (Anton *et al.,*2006; Zhang *et al.,* 2021). Moreover, egg quality characteristics such us egg weight, proportions of shell, yolk and albumen and nutrient composition can considerably affect the growing embryo during incubation and chick performance (İpek and Sözcü, 2013). Hence, continuous evaluation of different egg quality traits has become one of the major points of concern in modern poultry production (Wang *et al.,* 2017). Besides the chicken, ducks are the most significant poultry species (Bello *et al.,* 2022). Duck production is one of the branches of poultry production that supplies protein, eggs, and fatty liver (El-Deghadi *et al.,* 2022). Duck eggs are more nutritient than chicken eggs because they contain less water (Ismoyowati and Sumarmono, 2019).The egg production in the most productive duck breeds reaches about 250 to 300 eggs per year (Abd EL-Hack *et al.,* 2019). However, the economic importance and contribution of ducks to food security vary considerably between continents and countries (Pingel, 2011). In Algeria, the Mallard ducks are abundant, but their breeding is relatively undeveloped and restricted to traditional farms due to the lack of information on the nutritional value of ducks. To the best of our knowledge, no work has been undertaken to date to characterize the egg from duck in Algeria.

In animal research, several studies have made use of traditional statistical methods such as correlations, simple regression, and multivariate linear regression to estimate the relationships between traits of economic importance. Nevertheless, these conventional methods have not been found sufficient enough to model complex relationships. Specifically, the presence of strong relationships among predictors also known as multicollinearity compromises the results of multivariable regression analyses due to the inflation of the standard errors of the parameters, resulting in a reduction in the reliability of the final regression model (Kim, 2019). Moreover, traditional approaches follow strict statistical assumptions and data requirements. Difficulties caused from multicollinearity in regression analysis have been reported by different researchers (Eyduran *et al.,* 2010; Khorshidi-Jalali *et al*., 2019; Yakubu,2010; Dahloum *et al*., 2016).

An alternative to traditional statistics is statistical learning, also known as data mining (DM). DM is the use of computer-based methods to accurately model the nonlinear and complex relationship between the dependent variable and predictors in huge datasets (Pinto da Costa and Cabral, 2022).

Among various methods belonging to DM, the most commonly used algorithms include Multivariate Adaptive Regression Splines (MARS), Artificial Neural Networks (ANNs), and decision trees (DT) such us Chi-square Automatic Interaction Detection (CHAID), Classification and Regression Trees (CART), and Quick Unbiased Efficient Statistical Trees (QUEST). These methods, along with others such as Support Vector Machines (SVM), Random Forest Regression (RFR), and k-Nearest Neighbors (k-NN) have been preferred due to the advantages they possess, including the ability to handle nonlinear and noisy data, the absence of assumptions regarding the underlying distribution of values of the input variables, robustness against multicollinearity (Mendeş and Akkartal, 2009), suitability for high-dimensional data, simplicity, computational speed, high accuracy, and ease of interpretation.

Data mining applications have gained so much momentum in animal science recently (Grzesiak and Zaborski, 2012). Salawu *et al.* (2014) used ANN to predict the body weights of Rabbits. ANN was also successfully applied to predict and model milk yield in cows (Gocheva-Ilieva *et al.,*2022), and sheep (Karadas *et al.,* 2017). Almeida *et al.* (2020) applied ANN to predict zootechnical and management data in commercial laying hens farms. On the other hand, Nasser and Abu-Naser (2019) employed ANN for predicting the animal category. Eyduran *et al.* (2017) compared the predictive ability of MLR, CART, CHAID, and ANN in body weight prediction from some body measurements of the indigenous Beetal goat. Lee *et al.* (2020) estimated the carcass weight of Hanwoo cattle as a function of body measurements of Hanwoo cattle by using MLR, PLS (Partial least squares) regression, and ANN. For the prediction of body weight in sheep breeds, Tirink (2022) evaluated the ability of BRNN (Bayesian Regularized Neural Network), SVM, RFR, and MARS algorithms. Eyduran *et al.* (2013) applied RTM (Regression tree method) to predict the 305-d milk yield of Brown Swiss cattle. Grzesiak *et al.* (2010) used CF (classification functions), LR (logistic regression), ANN, and MARS for the detection of cows with artificial insemination difficulties.

In regard to establishing egg quality standards, Orhan *et al.* (2016) applied MLR, RR (Ridge Regression), and CHAID algorithm to predict egg weight based on albumen weight, yolk weight, and shell weight in commercial layer hybrids. In quail, Çelik *et al.* (2017) compared the predictive performance of CHAID, exhaustive CHAID and CART in the estimation of egg weight from some egg quality traits measurements.In another study, Sengul *et al.* (2020) compared Grossman-Koops, cubic and segmented polynomial models with MARS algorithm for predicting egg production in the Chukar partridge and found that the MARS predictive model can serve as a better alternative to classical non-linear models in predicting cumulative egg production.González Ariza *et al.* (2022) developed a stepwise discriminant canonical analysis to cluster eggs across hen genotypes considering egg quality attributes. Çelik *et al* (2016) investigated the effect of some egg quality traits (egg weight, egg width, egg height, and shape index) on fertility of eggs of Japanese quail with different colored feathers with the aid of CART data-mining algorithm.

As yet, no other studies are available on the egg quality characteristics of ducks using robust computational methods, and our results are the first to be reported. Therefore, the purpose of this study was to estimate and compare the ability of ALM, ANN, CART, and MLR models in the prediction of duck egg weight from some egg characteristics measurements based on several goodness of fit criteria.

MATERIALS AND METHODS

Material: Data were obtained from a total of 173 freshly laid eggs of Mallard ducks (35-50 wk old), directly collected from 20 smallholders in the province of Tiaret (35°55′52″ N, 0°08′24″E), located in northwest Algeria. The region is characterized by a semi-arid climate, and it is well known for its agricultural potential and livestock production.

To evaluate the internal and external egg quality traits, eggs were transported at 4℃ to the lab within 24h. The external egg traits recorded included egg weight (EWT, g), egg length (EL, mm), egg width (EWd, mm), egg shape index (ESI), and eggshell weight (ESW, g). Regarding the internal egg quality, the parameters measured were albumen weight (AW, g), albumen height (AH, mm), yolk weight (YW, g), yolk height (YH, mm), yolk diameter (YD, mm), and Haugh unit (HU). EWT was determined to the nearest 0.01g using an electronic scale. EWd, EL, and YD were determined with a digital caliper accurate to 0.1mm, while AH and YH were determined using a tripod micrometer with a precision of 0.01mm. The eggshell weight (ESW, g) was determined according to Sun *et al.* (2019) and Inca *et al.* (2020). ESI and HU were evaluated according to the following equations:

$$ESI = \frac{EWd}{EL} \times 100 \quad (1)$$

$$HU = 100 \log (AH + 7.7 - 1.7EWT^{0.37}) \quad (2)$$

Methods

Statistical analysis: The internal and external duck egg quality traits estimates were analyzed using some descriptive statistics (minimum, maximum, mean, standard error, and coefficient of variation). All phenotypic variables are given in Table 1. Linear associations between the egg traits were estimated by Pearson's correlation coefficient. In this study, EWT was the dependent variable, while the remaining 10 egg characteristics were the input explanatory variables (covariates).

Multiple linear regression analysis: Multiple linear regression analysis is a form of regression analysis commonly used for modeling the relationship between a dependent variable (regressand) and a set of independent variables (regressor) by a linear regression equation (Tabrizi and Sancar, 2017). To assess multicollinearity, the Variance Inflation Factor (VIF) is commonly used (Daoud, 2017). Multicollinearity is present when the VIF is higher than 5 to 10 (Tranmer *et al.,* 2020, Kim, 2019). The variables YD, YH, YW, AW, and ESW have been selected as input variables using the stepwise technique to predict egg weight according to the following formula:

$$EWT = a + b_i X_i + \ldots + b_i X_i \quad (3)$$

where EWT is the body weight, *a* is the regression intercept, b_i is the *ith* partial regression coefficients of the *ith* egg trait, and Xi is the *ith* egg trait.

Automatic Linear Modelling: Automatic Linear Modelling (ALM) is not as commonly used as the other computational methods but has gained popularity in recent years (Genç and Mendeş, 2021). ALM serves as a valuable screening tool, automating the process of selecting the most suitable subset of predictors, which is particularly crucial when dealing with a large number of predictors (Oshima and Dell-Ross, 2016; Genç and Mendeş, 2021). In the study, variables with VIF values > 10 were identified, and systematically removed to mitigate multicollinearity effects. The same predictor variables fitted into the MLR were used to generate the ALM. The selected ALM model was configured as a standard model with the forward stepwise as model selection method, and the Akaike's Information Criterion (AIC) for evaluating marginal contribution.

Machine learning models

Artificial Neural Network: An Artificial Neural Network (ANN) is a computing system based on the way biological nervous systems, such as the human brain (Dastres and Soori, 2021). The ANN is a methodology that takes into account nonlinearities in the relationship between the input and output information (Savegnago *et al.,* 2011). It consists of a set of interconnected neurons linked with weighted connections (Li *et al.,*2018).

In the current study, Multilayer Perception (MLP) with one hidden layer and Back Propagation network was used. The network was trained with 70% of the whole dataset and tested (model validation) with 30% of the dataset. The input layer consists of nodes corresponding to the 10 egg characteristic traits used for predicting egg weight. The hyperbolic tangent function and the linear activation function were employed for the hidden and output layers in ANN according to Yakubu and Nimyak (2020). The output layer has been configured with a single output node dedicated to estimating the egg weight. The weights and biases of this layer have been optimized during the model training process. Every other option in the ANN was set to default.

Classification and Regression Tree (CART): CART stands for Classification and Regression Trees (Breiman *et al*., 1984). It is a powerful predictive algorithm widely used in machine learning. CART models can be categorized based on the dependent variable. Categorical outcome variables require the use of a classification tree, while continuous outcomes utilize regression trees (Wray and Byers, 2020; Razi and Athappilly, 2005). CART constructs a binary decision tree structure where each fork represents a predictor variable, and each node provides a prediction for the target variable (Lee *et al*., 2010; Ali *et al*., 2015; Wray and Byers, 2020). In general, CART analysis begins with a single node, also known as the 'Parent node', while subsequent nodes that undergo further partitioning are termed 'child nodes'. The nodes where partitioning concludes, indicating homogeneity or purity, are commonly known as 'terminal nodes' or 'leaves'. CART looks for splits that minimize the prediction squared error (the least–squared deviation). The prediction in each leaf is based on the weighted mean for node (Maimon and Rokach, 2005). In the study, the dataset was initially divided into two distinct subsets, namely the training set, comprising 70% of the total data, and the test set, which accounted for the remaining 30%. The minimum observation count in parent and child nodes was set to 10:5 in order to improve the model predictive ability.

Comparison of the models quality: The quality of the assigned models was assessed and compared using the following specific statistical parameters according to Grzesiak and Zaborski (2012).

Pearson correlation coefficient between the observed and the predicted values

$$r_{Y_iY_{ip}} = \frac{cov(Y_i,Y_{ip})}{S_{Y_i}S_{Y_{ip}}} \quad (4)$$

Root mean square error

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - Y_{ip}\right)^2} \quad (5)$$

Global relative approximation error (RAE)

$$RAE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i-Y_{ip})^2}{\sum_{i=1}^{n} Y_i^2}} \quad (6)$$

Coefficient of determination

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i-Y_{ip})^2}{\sum_{i=1}^{n}(Y_i-\bar{Y})^2} \quad (7)$$

The Adjusted coefficient of determination

$$R^2_{adj} = 1 - \left[\frac{(1-R^2)(n-1)}{n-k-1}\right] \quad (8)$$

Where: Y_i is the actual egg weight value of ith egg, Y_{ip} is the predicted egg weight value of ith egg, $\bar{Y}$ isthe mean of the actual body weight values. n: the total sample size, and k the number of the independent variables in the model not including the constant.

All the computations were performed using the SPSS statistical software version 25.0.The significance level in all the analyses was set at $p<0.05$.

RESULTS

The Pearson correlation coefficients *(r)* between measured egg weight and external and internal quality traits of the egg are presented in Table 2. Correlations ranged from 0.000 to 0.945. EWT was highly and positively correlated with egg dimensions (EL and EWd, r= 0.752 and 0.790, respectively, $p<0.01$), AW (r= 0.815, $p<0.01$), and YW (r= 0.784, $p<0.01$). Low and negligible correlations were observed between EWT and ESI (r= −0.115, $p>$ 0.05) and between EWT and AH (r= −0.055, $p>$ 0.05). A negative significant correlation was also found between EWT and HU (r= −0.302, $p<0.01$). AW and YW showed a significant association with egg dimensions (EL and EWd) ranging from 0.602 to 0.672 ($p<0.01$). A highly significant, weak, and negative correlation (r= −0.195, $p<0.01$) was found between AH and YH.

Comparison of classification performances of the algorithms: In the present study, first, all 10 explanatory variables were included in the MLR model to predict EWT. The ANOVA results showed that the MLR model fitted was statistically significant (F= 486.74, $p<0.001$). When considering all the 10 predictors, the percentage of the EWT variance explained by the model is equal to 96.8%. The high values of VIF (>10) obtained for some of the independent variables is a sign of multicollinearity in the model. In the current study, the multicollinearity issues were found in EL, EWd, ESI, AH, and HU(Table 2). The estimated parameters of the MLR model are summarized in Table 3.

The explanatory variables with significant influence in determining the EWT were AW, YW, and ESW. As a result, the EWT prediction equation was EWT= 1.47+0.965AW+0.984YW+0.999ESW along with R^2=0.966, indicating that 96.6% of the total variation in the EWT is explained by these three variables. With the positive coefficients, an increment in EWT would be expected as AW, YW, and ESW increased.

The performance quality criteria of MLR, ALM, ANN, and CART models for the prediction of egg weight are summarized in Table 4. In the current study, the model exhibiting the highest values of r, R^2, and R^2_{adj}, along with the lowest values of RAE and RMSE, was selected as the most suitable model. The associations between the observed and the predicted egg weight using ALM, ANN, and CART are shown in Figures 1, 2, and 3. The Pearson correlation coefficient (r) between the observed and the predicted egg weight was highestin ANN (0.990)compared to ALM (0.984), MLR (0.982), andCART. (0.950). Similarly, the R^2 values for these models were 0.982, 0.970, 0.966, and 0.903, respectively. The R^2_{adj} values followed a similar pattern, with respective values of 0.981, 0.970, 0.964, and 0.897. The ANN model exhibited the lowest RMSE and RAE values of 0.753 and 0.012 in contrast to ALM (0.985,0.016), MLR (1.046, 0.017), and CART (1.778, 0.029).

The ANN model included input nodes consisting of 10 explanatory variables, hidden nodes comprising a bias term and seven H terms (H1:1 - H1:7), and the dependent variable EWT (Figure 4). Black lines indicate positive weights, while blue lines indicate negative weights. Line thickness is in proportion to the relative magnitude of each weight. In the ANN model, the most influential parameters for predicting EWT were AW, YW, HU, ESW, and AH (Table 5). These were followed by YD, EL, ESI, and EWd, while YH contributed the least to EWT determination. Both ALM and ANN algorithms ranked AW and YW as the most influential variables, while the order of importance for other variables varied between the models. In the ALM model, AW, YW and ESW, were identified as the significant explanatory variables automatically selected for predicting whole egg weight (Table 6**).** Conversely, the CART algorithm revealed a different set of significant input variables for predicting Mallard duck egg weight.

The relative contribution of each of the explanatory variables to the regression tree is presented in Table 7. With respect to the normalized importance, the contribution of egg width to the tree was at the highest ratio of 100%. It was followed by egg length (84.3%), albumen weight (82.0%), yolk weight (80.3%), yolk diameter (37.4%), eggshell weight (36.9%), yolk height

(30.9%), Haugh unit (15.9%), egg shape index (7.9%) and albumen height (3.2%).

The regression tree using CART algorithm is shown in Figure 5. The tree was built with five variables (EWd, AW, EL, YW and ESI). The tree was mostly influenced by EWd while the least influence was exhibited by ESI. A total of ten terminal (homogeneous) nodes (nodes 7, 9, 10, 11, 12, 13, 15, 16, 17 and 18), on which decisions are made, were formed. Node 0, which is the root node, provided information on the descriptive statistics where the total number of observations was 173 and mean egg weight was 59.312 g with standard deviation of 5.728. Based on the influence of egg width, Node 0 was partitioned into non-homogeneous nodes 1 and 2 with predicted mean egg weight of 54.676 g and 63.394 g, respectively. Node 1, on the basis of egg width,was split further into node 3 (Ewd≤ 41.446 mm) and node 4 (Ewd> 41.446 mm) while node 2, based on albumen weight, was divided into node 5 (AW≤ 33.901 g) and node 6 (AW>33.401 g). The respective predicted mean egg weights in both cases were 61.429 g and 66.451 g. On the basis of albumen weight, node 3 was further partitioned into homogeneous node 7 (AW≤ 25.902; EWT= 47.797; SD= 2.690) and non-homogeneous node 8 (AW> 25.902; EWT= 53.118; SD= 1.822). Node 4,based on egg length, branched into homogeneous node 9 (EL≤ 59.525; EWT= 56.186; SD=1.575) and homogeneous node 10 (EL> 59.525; EWT = 60.483; SD=1.754). Node5 was influenced by yolk weight and was split further into two homogeneous nodes 11 and 12. At node 11, a total of 32 eggs with YW≤ 22.865 were grouped with a predicted mean egg weight of 60.039 g and SD of 1.754. With a total of 32 eggs and YW>22.865, node 12 had predicted mean egg weight of 63.281 g and SD of 2.142. Using egg length as a splitting criterion, node 6 was partitioned into terminal node 13 (EL≤ 60.950; EWT = 63.825; SD=1.548) and non-terminal node 14 (EL>60.950; EWT = 67.935; SD=3.090). Node 8 was further divided into two terminal nodes 15 and 16 using albumen weight as the splitting variable. While node15 with AW≤29.519 had predicted mean egg weight of 52.382 g and SD of 1.292, node 16 with AW>29.519 was characterized by predicted mean egg weight of 55.000 g and SD of 1.655. Node 14 was split into two terminal nodes 17 and 18 using egg shape index as the criterion with a small improvement of 0.446. With ESI≤70.723, the predicted egg weight and standard deviation of node 17 were 66.847 g and 1.983. At node 18 with ESI>70.723, the predicted egg weight of 71.051 g (SD= 3.742) was the highest among all the ten terminal nodes.

Table 1. Basic statitistics for different traits (n=173).

Trait under study	Minimum	Maximum	Mean	Std Dev	CV (%)
External traits					
Egg weight	42.840	76.070	59.310	5.720	9.640
Egg lenght	45.690	65.830	59.200	2.800	4.730
Egg width	31.310	45.690	42.320	1.670	3.950
Egg shape index	65.090	76.720	71.560	2.510	3.510
Internal traits					
Yolk diameter	35.330	49.310	47.220	5.850	12.390
Yolk height	7.800	20.960	16.670	1.500	9.000
Albumlen height	5.160	17.430	7.380	1.250	16.940
Yolk weight	15.200	30.030	21.450	2.830	13.190
Eggshell weight	4.600	11.000	6.150	1.270	20.650
Abumen weight	21.300	45.380	31.680	3.570	11.270
Haugh units	70.290	123.590	85.640	7.070	8.260

Std Dev: standard deviation; CV: coefficient of variation

Table 2. Pearson's correlation between egg traits (n=173).

	EL	EWd	ESI	YD	YH	AH	YW	ESW	AW	HU	VIF
EWT	0,752**	0,790**	−0,115^{ns}	0,319**	0,346**	−0,055^{ns}	0,784**	0,422**	0,815**	−0,302**	
EL	1	0,680**	−0,560**	0,321**	0,348**	−0,080^{ns}	0,672**	0,275**	0,602**	−0,270**	844,181
EWd		1	0,226**	0,246**	0,283**	−0,031^{ns}	0,648**	0,322**	0,663**	−0,237**	643,568
ESI			1	−0,148*	−0,153*	0,065^{ns}	−0,170*	−0,006^{ns}	−0,056^{ns}	0,087^{ns}	488,916
YD				1	0,101^{ns}	−0,021^{ns}	0,435**	0,085^{ns}	0,147*	−0,094^{ns}	1,273
YH					1	−0,195**	0,374**	0,003^{ns}	0,279**	−0,181*	1,679
AH						1	−0,040^{ns}	−0,288**	0,052^{ns}	**0,945****	27,468
YW							1	0,281**	0,390**	−0,237**	3,314
ESW								1	0,105^{ns}	−0,426**	1,576
AW									1	−0,138**	2.620
HU										1	29,800

VIF: variance inflation factor.

Table 3.The estimated MLR parameters.

Model term	Estimate	SE	*t*-value	*p*-value	VIF
Intercept	1.478	0,837	1.765	0.079	
Albumen weight	0.965	0.025	39.321	<0.000	1.180
Yolk weight	0.984	0.032	30.735	<0.000	1.267
Eggshell weight	0.999	0.066	15.123	<0.000	1.086
SE: standard error					

Table 4.Predictive performance of MLR, ALM, and ANN.

Model	r	R^2	R^2_{adj}	RMSE	RAE
MLR	0,982	0,966	0,964	1,046	0,017
ALM	0,984	0,970	0,970	0,985	0,016
ANN	0,990	0,982	0,981	0.753	0,012
CART	0.950	0,903	0,897	1,778	0,029

Table 5.Variable importance in the prediction of EWT using ANN.

Model term	Importance	Normalized importance (%)
Egg lenght	0.060	18.6
Egg width	0.040	12.8
Egg shape index	0.042	13.5
Yolk diameter	0.061	19.5
Yolk height	0.013	4.1
Albumen height	0.085	27.1
Yolk weight	0.205	65.2
Eggshell weight	0.091	29.4
Albumen weight	0.312	100.0
Haugh unit	0.092	29.8

Table 6.Variable importance in the prediction of EWT using ALM.

Model term	Coefficient	SE	*t*-value	*p*-value	Importance
Intercept	0.019	0.809	0.023	0.981	
Albumen weight	1.023	0.024	41.844	<0.000	0.582
Yolk weight	0.958	0.030	31.454	<0.000	0.329
Eggshell weight	1.035	0.063	16.337	<0.000	0.089

SE: standard error

Table 7.Variable importance in the prediction of EWT using CART.

Model term	Importance	Normalized importance (%)
Egg width	26.742	100.0
Egg lenght	22.550	84.3
Albumen weight	21.928	82.0
Yolk weight	21.482	80.3
Yolk diameter	9.999	37.4
Eeggshell weight	9.878	36.9
Yolk height	8.253	30.9
Haugh unit	4.255	15.9
Egg shape index	2.103	7.9
Albumen height	0.854	3.2

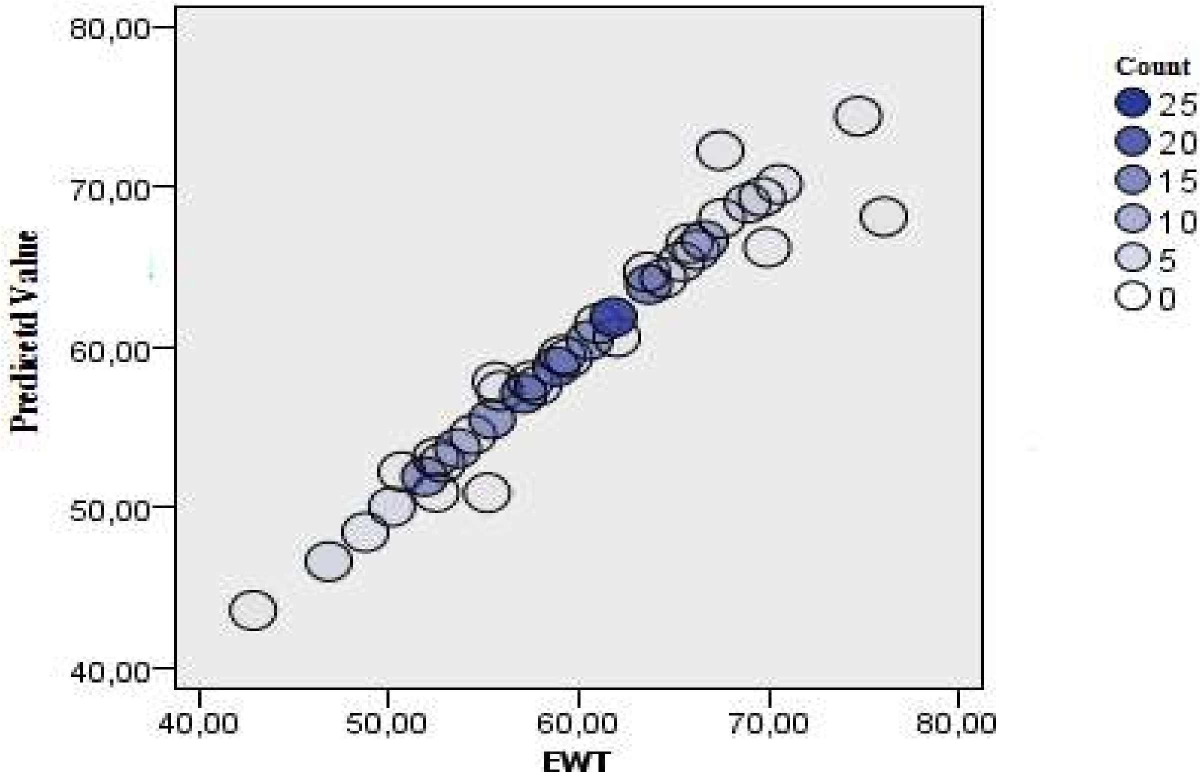

Figure 1.The scatter plot of observed and predicted egg weight using ALM

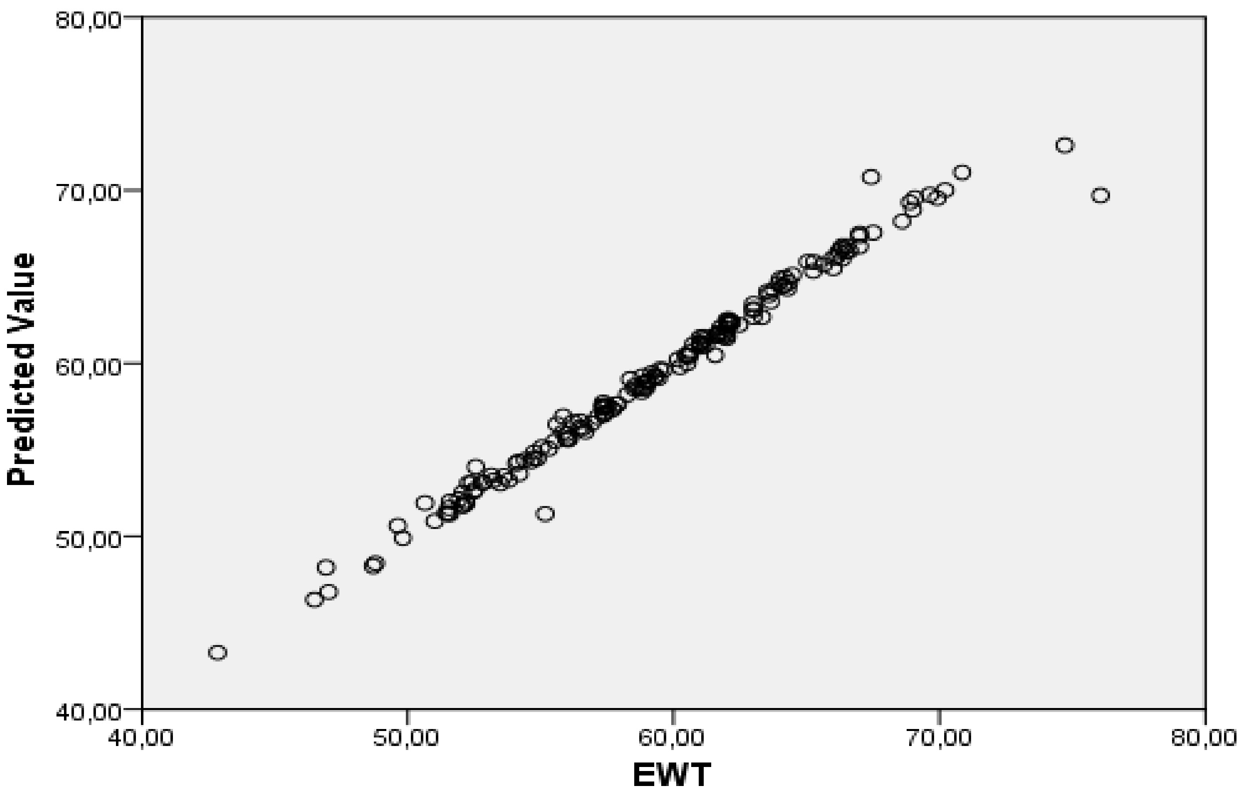

Figure 2.The scattter plot of observed and predicted egg weight using ANN

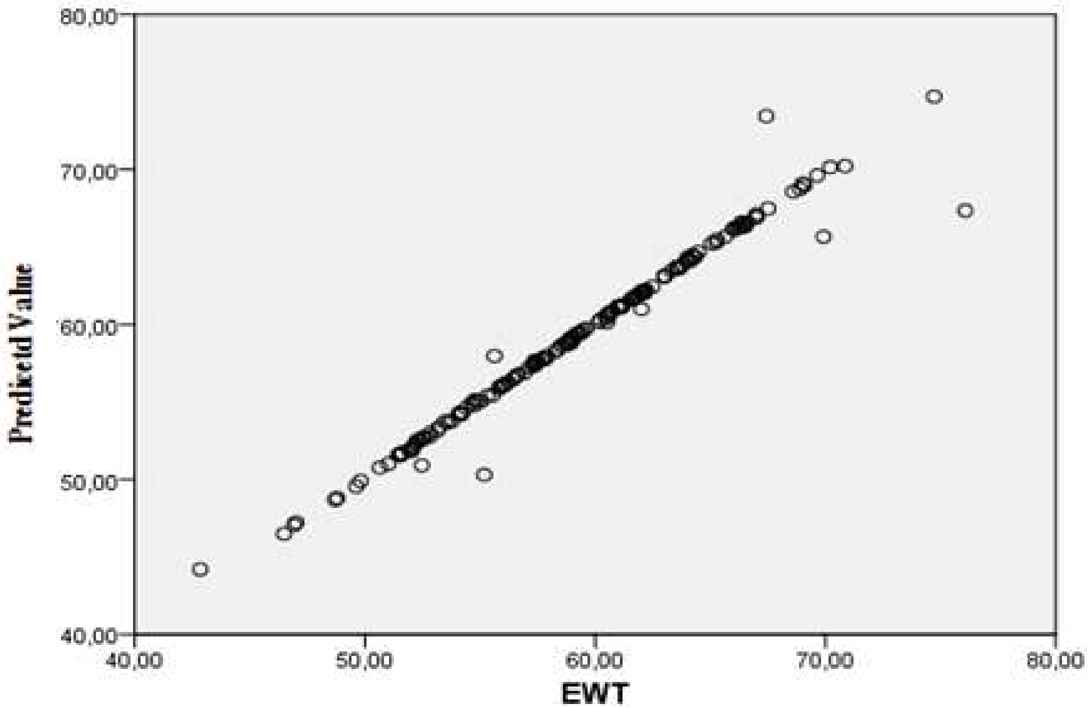

Figure 3.The scatter plot of observed and predicted egg weight using CART

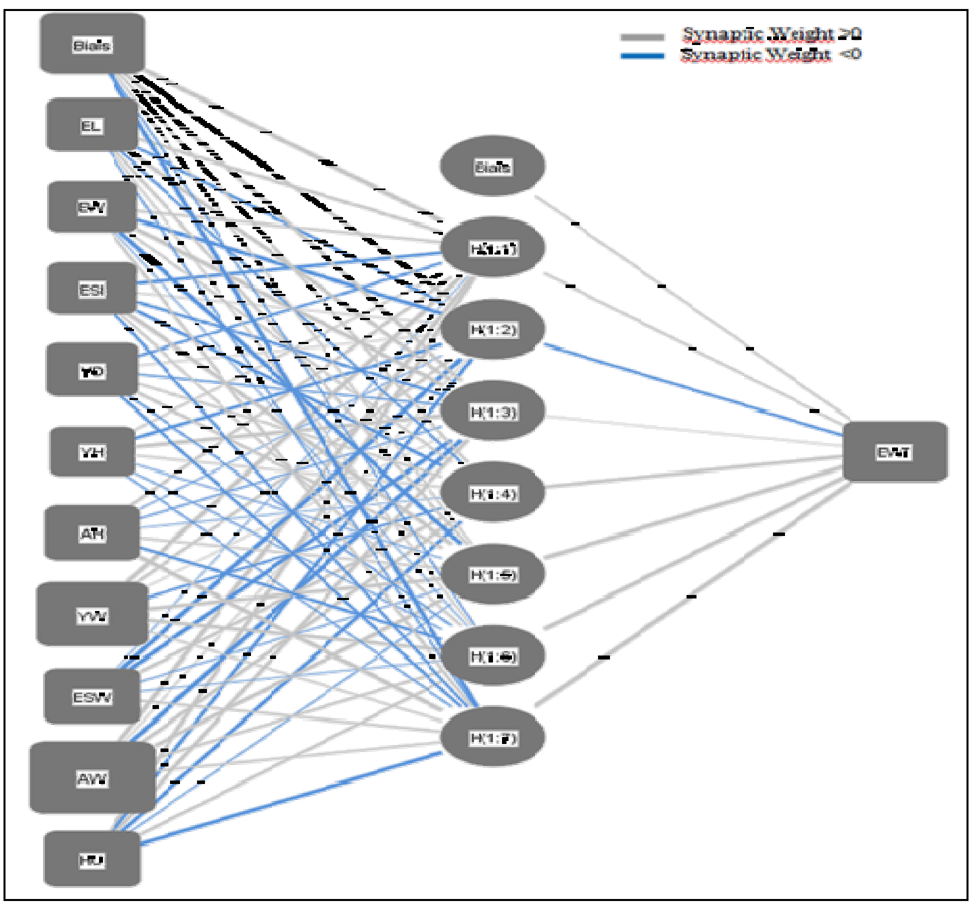

Figure 4.Diagram of the ANN architecture

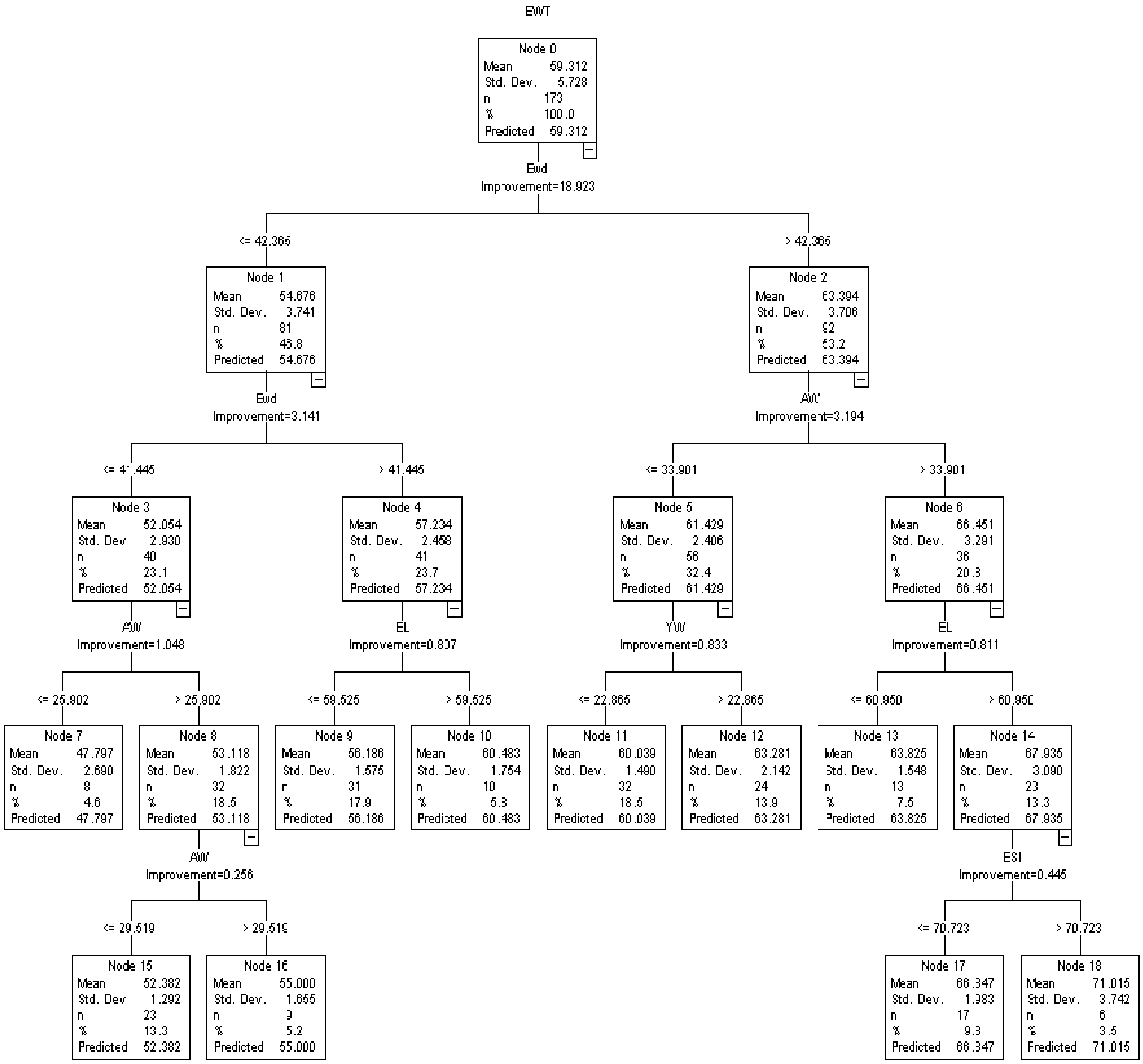

Figure 5. CART tree model for EWT prediction

DISCUSSION

The data used in the current study were estimated in terms of basic statistics. The mean egg weight and size obtained in this study are higher than those reported by Labbaci *et al.* (2014) in Mallard ducks at Lake Tonga (Northeastern Algeria). The mean values for egg length, egg width, eggshell weight, and albumen weight are similar to those reported in Alabio duck (Hartati *et al.,*2021). Contrastingly, the mean egg weight and most of the other external and internal quality trait averages were lower than the results obtained in Nigerian Muscovy duck reared under different management systems, except albumen height, yolk diameter,and Haugh unit (Etuk *et al.,*2012). Similarly, in Pekin duck eggs, Indarsih *et al.* (2021) found a mean weight of 67.5±5.9g, with a mean length of 60.7±3.1mm, mean width of 44.7±0.9mm, and an average shape index of 76.2±1.7 and 70.9±2.8 for rounded and elongated eggs, respectively. These authors demonstrated that the shape index is a suitable parameter for sex identification in Pekin duck. Lin *et al.* (2016) reported higher mean values of egg weights of 65.0±3.9 and 67.0±4.2g in Shan Ma laying ducks, at 210 and 300 days of age, respectively. There were also reports of higher mean values in comparison to the present findings for egg weight in Domyati duck (Egyptian local breed) and Khaki-Campbell duck with 61.4±6.5and 64.3±3.4g, respectively (El-Deghadi *et al.,* 2022). The reasons for diverse opinions among researchers regarding some egg characteristics are multifactorial suchas genetic factors, layer's age, body weight, health, nutrition, and egg storage conditions (Roberts, 2004; Dahloum *et al.,* 2018; Alkan and Türker, 2021; Çelik *et al.,*2021). In this sense, Reyna and Burggren (2017) reported a decrease in the fertility of duck eggs stored for more than six days from laying to incubation. In addition, Ipek and Sozcu (2017) found thatheavier eggs from Pekin ducks had better hatchability than the light and medium eggs.

The results of the correlation analyses showed highly significant associations between the independent variables. The strong correlation between EWT and both AW and YW suggests that these parameters can change at a significant level depending on the change that can occur in the egg weight. This finding is consistent with the results of several previous studies on other poultry species such us quails (Ouaffai *et al.,*2018), Guinea Fowl (Onunkwo and Okoro, 2015), partridge (Alkan *et al.,*2014), and commercial layers (Orhan *et al.,* 2016). The statistically non-significant and negative phenotypic correlation value found in the present study between the egg weight and the egg shape index is consistent with the findings of several previous studies (Olawumi and Ogunlade, 2008; Alkan and Türker, 2021; Jang, 2022).

The negative significant relationship between EWT and HU stands in contrast to the insignificant correlation reported by other researchers (Debnath and Ghosh, 2015; Vekić *et al.,* 2022). Haugh unit is an important index to evaluate egg protein qualityand reflect egg freshness (Gao *et al.,* 2022). The negative associations of HU with AW, YW, and YD obtained in the current study agree with the previous report on indigenous chickens (Bekele *et al.,* 2022). In Alabio Duck, Hartati *et al.* (2021) found that EL was positively correlated with ESW, AW, and YW (r=0.28, 0.53, and 0.52, respectively, $p<0.05$).

Predictive performance of ANN,ALM, CART, and MLR: Data mining techniques can be a good option to describe complex associations between variables (Canga *et al.,*2021). In the study, the ALM, ANN, and CART algorithms yielded different sets of significant predictors due to the distinct methodologies and criteria they employ for variable selection and model construction. These variations reflect the inherent differences in their strategies for modelling and predicting egg weight.

In the poultry field, Bolzan *et al.* (2008) explored the use of ANN to predict the hatchability of artificially incubated eggs derived from a 39-week-old Cobb 500 broiler breeder flock. ALM and ANN were fitted with the intent to predict hatchability and mortality in muscovy ducks (Yakubu *et al.,*2019), and to forecast heat stress index in Sasso hens (Yakubu *et al.,*2018). Çelik *et al.* (2021) evaluated the performance of CART and MARS in the prediction of egg weight of the quails. Canga *et al.* (2021) aimed to predict egg weight from egg quality traits in Lohman LSL Classic white hybrid laying hens, with the help of the MARS data mining algorithm.

The application of data mining techniques was also successfully investigated to estimate egg weight in many poultry species. This is the first modelling study to determine Mallard duck egg weight using a combination of external and internal egg characteristics, employing ANN and CART algorithms, along with the ALM technique. However, in this study, it was not possible to make an adequate comparison with other studies owing to the use of different poultry species, traits, variables, sample sizes, and different computational methods. Portillo-Salgado *et al.* (2021) demonstrated that both the decision tree technique based on the CHAID algorithm and MLR can be used reliably for predictive estimates of egg weight from external traits of Guinea fowl as they showed similar accuracy (R^2= 74.0 and 75.0%, respectively). Çelik*et al.* (2017) investigated the ability of the CART, CHAID, and Exhaustive CHAID algorithms in the prediction of quail egg weight. In their study, the Pearson correlation coefficients (*r*) between actual and predicted egg weight values for CHAID, Exhaustive CHAID, and CART algorithms were 90.6%, 92.7%, and 92.0%, respectively. In the same order, the R^2 values were 82.06%, 85.86%, and 84.66%, while the R^2_{adj} values were 82.06%, 85.85%, and 84.66%. The RAE estimates were 0.087 for all algorithms and the estimates of RMSE were 0.453, 0.402, and 0.419, respectively. The results indicate that the Exhaustive CHAID algorithm is very effective for determining internal and external quality features in quail eggs. In another study on quail, Çelik *et al.* (2021) demonstrated that MARS showed much better predictive performance than CART for the prediction of egg weight with R^2 values of 85.0% and 72.8%, respectively. The reported values of R^2 are smaller than the R^2 values obtained in the current study. Çiftsüren and Akkol (2018) used RR (Ridge regression), LASSO (Least Absolute Shrinkage and Selection Operator), and EN (Elastic net) regression models to determine egg yolk and egg albumen weights from some external egg characteristics in Japanese quail. In their study, it was revealed that LASSO was the best model due to its high predictive accuracy. For egg yolk weight, the goodness of fit of the regression estimating equations was 58.34%, 59.17%, and 59.11% for RR, LASSO, and EN methods, respectively. For egg albumen weight the goodness of fit of the regression equations was 75.60%, 75.94%, and 75.81% for the respective RR, LASSO, and EN methods. In the study conducted by Alapatt *et al.* (2022) to determine the egg weight in White Leghorn Chicken from some internal and external egg traits using different methods, the EN regression was identified as the best predictive model (R^2_{adj}=86.5%) followed by RR (R^2_{adj}=81.13%), RFR (R^2_{adj}= 65.02%), LASSO (R^2_{adj}= 29.67%), and CART algorithm (R^2_{adj}= 29.45%). In the prediction of the egg weight from albumen weight, yolk weight, and shell weight in commercial layer hybrids, Orhan *et al.* (2016) demonstrates the superiority of the CHAID algorithm with higher accuracy (R^2= 99.98%) compared to MLR (R^2= 93.4%) and RR (R^2= 93.15%). Canga *et al.* (2021) applied the MARS data mining algorithmto predict egg weight from egg quality traits in Lohman LSL Classic white hybrid laying hens and achieved sufficient fit with the mean predictive performance measures estimated as 61.0%, 0.779, and 0.430 for R^2, *r*, and SD ratio, respectively. In indigenous free-range chickens, Liswaniso *et al.* (2021) preferred CHAID and CART algorithms to predict the egg weight from egg length, egg width, shell weight, shell thickness, albumen weight, yolk height, yolk width, and yolk weight. For CHAID algorithm, the goodness of fit was R^2= 82.3%, R^2_{adj}=82.3%, RMSE=2.23, RAE=0.04, and SD ratio=0.04. In the case of the CART algorithm, the results were estimated to be 59.3%, 59.3%, 2.32, 0.07, and 0.24, respectively.

Conclusion: Based on the results of the present study, it was concluded that the ANN algorithm was slightly more efficient for egg weight determination in Mallard ducks based on some internal and external egg traits as illustrated by its lower error measurements compared to ALM, MLR, and CART algorithm. These findings may assist poultry researchers and producers to choose the best predictors to increase egg quality in ducks by the selection of high-performance genotypes.

Conflict of Interest: The authors declare that they have no conflict of interest.

Author Contributions: LD and QB contributed to the conception and design of the study. LD organized the database. LD and AY conducted the statistical analyses. LD wrote the original draft. AY revised the manuscript. All authors read and approved the submitted version.

Acknowledgments: The authors are grateful for the financial support providedby the Directorate-General for Scientific Research and Technological Development (DGRSDT), Algeria. We also wish to thank all the smallholder farmers for providing the egg samples for use in this study.

REFERENCES

Abd El-Hack, M.E., C.B. Hurtado, D.M. Toro,M. Alagawany, E.M.Abdelfattah and S.S. Elnesr (2019). Fertility and hatchability in duck eggs.Worlds Poult. Sci. J. 75:4, 599–608. https://doi:10.1017/S0043933919000060

Ali, M., E. Eyduran., M.M. Tariq., C. Tirink., F. Abbas., M.A. Bajwa *et al* (2015). Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post weaning period from some biometrical characteristics in Harnai Sheep. Pakistan J. Zool, 47:1579-1585. https://doi.org/10.17582/journal.pjz/2018.50.1.189.195

Alkan S and İ. Türker (2021). Effects of Egg Shape Index on Egg Quality in Partridges. Ordu Univ. J. Sci. Tech. 11,140−151. https://doi:10.54370/ordubtd.996530

Alkan, S, A.Galiç, T.Karsli and K. Karabağ (2014). Effects of egg weight on egg quality traits in partridge (Alectoris Chukar). J. Appl. Anim. Res.0, 1–7. https://doi:10.1080/09712119.2014.980419

Almeida, L.G.B., E.B Oliveira, T.Q. Furian, K. A. Borges, D. T, Rocha, C. T. P. Salle and H. L. S. Moraes(2020). Artificial Neural Networks on Eggs Production Data Management. Acta Sci. Vet. 48: 1732. https://doi: 10.22456/1679-9216.101462

Anton, M., F. Nau and Y. Nys (2006). Bioactive egg components and their potential uses. Worlds Poult. Sci. J. 62, 429-438. https://doi:10.1017/S004393390600105X

Bekele, B., A. Melesse, W. Esatu and T. Dessie (2022). Production performance and egg quality evaluation of indigenous chickens across different agro-ecologies of Southern Ethiopia.Vet Integr Sci. 20,133−145. https://doi: 10.12982/VIS.2022.012

Bello, S. F., A. C. Adeola and Q. Nie (2022). The study of candidate genes in the improvement of egg production in ducks. Poult. Sci. 101, 7. https://doi:10.1016/j.psj.2022.101850

Bolzan, A.C., R.A.F. Machado and J.C.Z. Piaia (2008). Egg Hatchability Prediction by Multiple Linear Regression and Artificial Neural Networks. Braz J Poultry Sci. 10, 97–102.

Breiman L., J. Friedman., R. Olshen and C. Stone (1984). Classification and Regression Trees. Chapman and Hall, Wadsworth (New York).

Canga, D., E. Yavuz and E. Efe (2021). Egg Weight Estimation with Mars Data Mining Algorithm Using R. KSU J. Agric Nat, 24, 242−251. https://doi:10.18016/ksutarimdoga.vi.716880

Çelik, Ş., E. Eyduran,A.Y Şengül and T Şengül (2021). Relationship among egg quality traits in Japanese quails and prediction of egg weight and color using data mining algorithms. Trop Anim Health Prod. 53, 382. https://doi:10.1007/s11250-021-02811-2

Çelik, S., B. Sogut., T. Sengul., E. Eyduran and A. Y. Sengul (2016). Usability of CART algorithm for determining egg quality characteristics influencing fertility in the eggs of Japanese quail. R. Bras. Zootec 45 (11): 645-649.

https://doi.org/10.1590/S1806-92902017001100005
Çelik, Ş., T. Şengul, H Inci,B.Soğut and A Şengul (2017). Estimation of egg weight fromsome external and internal quality characteristics in quail by using various data mining algorithms. Indian. J. Anim. Sci.87: 1524-30.https://doi.10.56093/ijans.v87i12.79871
Çiftsüren, M. N and S.Akkol (2018). Prediction of internal egg quality characteristics and variable selection using regularization methods: ridge, LASSO, and elastic net. Arch. Anim. Breed. 61, 279–284. https://doi:10.5194/aab-61-279-2018
Dahloum, L., A. Yakubu and M. Halbouche (2018). Effects of housing system and plumage colour on egg quality characteristics of indigenous naked-neck chickens. Livest. Res. Rural. Dev.30, 12. Available at:http://www.lrrd.org/lrrd30/12/abdul30206.html
Dahloum, L., N. Moula., M. Halbouche and S Mignon-Grasteau (2016). Phenotypic characterization of the indigenous chickens*(Gallus gallus)* in the northwest of Algeria, Arch. Anim. Breed, 59, 79–90. https://doi.org/10.5194/aab-59-79-2016.
Dastres, R and M. Soori (2021). Artificial Neural Network Systems. Int. J. Imag. Robot, 21:2.
Debnath, B. C and T.K. Ghosh (2015). Phenotypic correlations between some external and internal egg quality traits in Gramapriya layers. Explor. Anim. Medical Res, 5(1): 78-85.
El-Deghadi, A. S., W. A. H. Ali and M. G. Gharib (2022). Study for Some Body Weight and Egg Traits in Domyati and Khaki-Campbell Ducks. Open J. Agricultural Research, *2* (1), 29–36. https://doi: 10.31586/ojar.2022.260
Etuk, I.F., G.S Ojewola, S.F. Abasiekong, K.U. Amaefuleand E. B.Etuk (2012). Egg quality of Muscovy ducks reared under different management systems in the humid tropics. Rev. Cient. UDO Agric. 12, 225−228
Eyduran E., M. Topal and A.Y. Sonmez (2010). Use of factor scores in multiple regression analysis for estimation of body weight by several body measurements in brown trouts (*Salmo trutta fario*). Int. J. Agric. Biol. 12,611-615.
Eyduran, E., D. Zaborski, A.Waheed, S.Çelik, K. Karadas and W. Grzesiak (2017). Comparison of the predictive capabilities of several data mining algorithms and multiple linear regression in the prediction of body weight by means of body measurements in the indigenous Beetal goat of Pakistan. Pakistan J. Zool.49, 257–265. https://doi:10.17582/journal.pjz/2017.49.1.257.265 http://
Eyduran, E., I. Yilmaz, M. M. Tariq and A. Kaygisiz (2013). Estimation of 305-d milk yield using regression tree method in Brown Swiss cattle. J. Anim. Plant. Sci. 23,731−735.
Gao, J., W.Xu, T. Zeng, Y.Tian, C.Wu, S.Liu *et al* (2022). Genome-wide association study of egg-laying traits and egg quality in LingKun chickens. Front. Vet. Sci. 9:877739. https://doi: 10.3389/fvets.2022.877739
Genç, S and M. Mendes (2021). Evaluating performance and determining optimum sample size for regression tree and automatic linear modeling. Arq. bras. med. vet. zootec.73,1391−1402. https://doi:10.1590/1678-4162-12413
Gocheva-Ilieva, S., A. Yordanova and H. Kulina (2022). Predicting the 305-day milk yield of HolsteinFriesian cows depending on the conformation traits and farm using simplified selective ensembles. Mathematics, 10, 1254. https://doi:10.3390/math10081254
González Ariza, A., A. Arando Arbulu, F.J. Navas González,.*et al*. (2022). Data mining-based discriminant analysis as a tool for the study of egg quality in native hen breeds. Sci. Rep,12, 15873. https://doi.org/10.1038/s41598-022-20111-z
Grzesiak, W and D. Zaborski (2012). Examples of the use of data mining methods in animal breeding. In Data Mining Applications in Engineering and Medicine; Karahoca, A, Ed. InTech, Rijeka, Croatia, pp. 303–324.
Grzesiak, W., D. Zaborski, P.Sablik, A. Zukiewicz, A. Dybus and I. Szatkowska (2010).Detection of cows with insemination problems using selected classification models. Comput Electron Agric, 74, 265–273. https://doi:10.1016/j.compag.2010.09.001
Hartati L, N. Muawanah, M. Riyadhi, D. Biyatmoko, A. Sulaiman and G.Simanungkalit (2021). Models of the relationships between egg dimension and egg quality of Alabio duck. (ITAPS 2021). Advances in Biological Sciences Research, 20. Proceedings of the International Conference on Improving Tropical Animal Production for Food Security. https://doi:10.2991/absr.k.220309.027
Inca, S., D.A. Martinez and C. Vilchez (2020). Phenotypic correlation between external and internal egg quality characteristics in 85-week-old laying hens. Int. J. Poult. Sci. 19, 346–355. https://doi:10.3923/ijps.2020.346.355
Indarsih, B., M. H. Tamzil, D.Kisworo and Y. Aprilianti (2021). Egg shape index for sex determination of post-hatch chicks in Pekin ducks: a solution for smallholder duck farming in Lombok Indonesia. Livest. Res. Rural. Dev.33:4. Available at: http://www.lrrd.org/lrrd33/4/3348budii.html

İpek, Aand A. Sözcü (2013). Broiler Chick Quality And Scoring Methods. J. Agricultural Faculty, 27(2), 131-137.

İpek, Aand A. Sözcü (2017). Comparison of hatching egg characteristics, embryo development, yolk absorption, hatch window, and hatchability of pekin duck eggs of different weights. Poult. Sci.96, 3593−3599. https://doi:10.3382/ps/pex181

Ismoyowati and J Sumarmono (2019). Duck Production for Food Security The 1st Animal Science and Food Technology Conference (AnSTC), IOP Conf. Ser.: Earth Environ. Sci. 372 012070

Jang, E (2022). Correlation between Internal and External Egg Quality Indicators in the Early Phase of Hy-Line Brown Laying Hens. Korean J. Poult. Sci. 49,53−60. https://doi:10.5536/KJPS.2022.49.2.53

Karadas, K., M. Tariq, M.M. Tariq and E. Eyduran (2017). Measuring predictive performance of data mining and artificial neural network algorithms for predicting lactation milk yield in indigenous Akkaraman sheep. 49, 1–8. Pakistan J. Zool. https://doi:10.17582/journal.pjz/2017.49.1.1.8

Khorshidi-Jalali., M. M.R. Mohammadabadi., A. Esmailizadeh., A. Barazandeh and O.I. Babenko (2019). Comparison of Artificial Neural Network and Regression Models for Prediction of Body Weight in Raini Cashmere Goat .Iran. J. Appl. Anim. Sci, 3 (9), 453-461.

Kim, J.H (2019). Multicollinearity and misleading statistical results. Korean. J. Anesthesiol. 72(6):558-569. https://doi: 10.4097/kja.19087

Labbaci, R., S. Bourbia, Z. Bouslama (2014). Reproductive ecology of mallard duck (*Anas platyrhynchos)* at Tonga lake (North-East of Algeria).Ann. Biol. Res. 5, 72-78. Available at: www.scholarsresearchlibrary.com

Lee, D., S. Lee, B. Cho, C. Wakholi, Y. Seo, S. Cho *et al*. (2020). Estimation of carcass weight of Hanwoo (Korean native cattle) as a function of body measurements using statistical models and a neural network. Asian-Australas J. Anim. Sci, 33, 1633−1641. https://doi: doi:10.5713/ajas.19.0748

Lee, B K., J. Lessler and E.A. Stuart (2010). Improving propensity score weighting using machine learning. Stat. Med. 29 (3): 337–346. https://doi.org/10.1002/sim.3782

Li, G., X. Zhou, J. Liu, Y. Chen, H. Zhang,Y. Chen *et al*. (2018). Comparison of three data mining models for prediction of advanced schistosomiasis prognosis in the Hubei province. PLoS Negl Trop Dis. 12:2. https://doi:10.1371/journal.pntd.0006262

Lin, R.L., H.P.Chen, R.Rouvier, C.Marie-Etancelin (2016). Genetic parameters of body weight, egg production, and shell quality traits in the Shan Ma laying duck *(Anas platyrhynchos)*. Poult. Sci. 95, 514−2519. https://doi: 10.3382/ps/pew222. PMID: 27520070

Liswaniso, S., N.Qin, T.L Tyasi, and I.M. Chimbaka (2021). Use of data mining algorithms CHAID and CART in predicting egg weight from egg quality traits of indigenous free-range chickens in zambia. Adv. Anim. Vet. Sci. 9, 215−220. https://doi:10.17582/journal.aavs/2021/9.2.215.220

Maimon, O and L. Rokach (2005). Introduction to Supervised Methods. In: Maimon, O., Rokach, L. (eds) Data Mining and Knowledge Discovery Handbook. Boston, MA. https://doi.org/10.1007/0-387-25465

Mendeş, M and E. Akkartal (2009). Regression tree analysis for predicting slaughter weight in broilers. Ital. J. Anim. Sci,8:4,615-624,https://doi:10.4081/ijas.2009.615

Nasser, I. M and S. S. Abu-Naser (2019). Artificial Neural Network for Predicting Animals Category. Int. j. acad. appl. res.3, 18−24.

Olawumi, S.O and J.T. Ogunlade (2008). Phenotypic Correlations Between Some External and Internal Egg Quality Traits in the Exotic Isa Brown Layer Breeders.Asian J. Poult. Sci.,2, 30−35. https://doi:10.3923/ajpsaj.2008.30.35

Onunkwo, D.N and I. C. Okoro (2015). Phenotypic Correlations between External and Internal Egg Quality Traits in Three Varieties of Helmeted Guinea Fowl from 28 to 46 Weeks of Age. Int. J. Livest. Res. 5:3. https://doi: 10.5455/ijlr.20150318092255

Orhan, H., E.Eyduran, A.Tatliyer and H. Saygici (2016). Prediction of egg weight from egg quality characteristics via ridge regression and regression tree methods. R. Bras. Zootec. 45, 380−385. https://doi:10.1590/S1806-92902016000700004

Oshima, T. C and T.Dell-Ross (2016). All Possible Regressions Using IBM SPSS: A Practitioner's Guide to Automatic Linear Modeling, Georgia Educational Research Association Conference. Georgia Southern University. Available at: https://digitalcommons.georgiasouthern.edu/gera/2016/2016/1

Ouaffai, A., L. Dahloum, A. Fassih, M. Milagh and M. Halbouche(2018). Performances de croissance, de ponte et qualité de l'œuf chez la caille Japonaise *(Coturnix coturnix japonica)*. Arch. de Zootec, 67,168−176. https://doi: 0.21071/az.v67i258.3651

Pingel, H (2011). Waterfowl Production for Food Security, Lohmann information, 46:2, 32–42.

Pinto da Costa, J and M. Cabral (2022). Statistical Methods with Applications in Data Mining, A Review of the Most Recent Works. Mathematics,10, 993. https://doi:10.3390/math10060993

Portillo-Salgado, R., F.A. Cigarroa-Vázquez, B. Ruiz-Sesma, P. Mendoza-Nazar, A. Hernández-Marín, W. Esponda-Hernández *et al* (2021). Prediction of Egg Weight from External Egg Traits of Guinea Fowl Using Multiple Linear Regression and Regression Tree Methods. Braz J Poultry Sci. 23, 001–006. https://doi:10.1590/1806-9061-2020-1350

Razi, M.A. and K. Athappilly (2005). A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. Expert. Syst.Appl, 29 (1), 65-74. https://doi.org/10.1016/j.eswa.2005.01.006.

Réhault-Godbert,S., N, Guyotand Y. Nys (2019). The Golden Egg: Nutritional Value, Bioactivities, and Emerging Benefits for Human Health. Nutrients; 11, 684. https://doi:10.3390/nu11030684

Reyna, K.S and W.W. Burggren (2017). Altered embryonic development in northern bobwhite quail (*Colinus virginianus*) induced by pre-incubation oscillatory thermal stresses mimicking global warming predictions. PloS One 12 (9): e0184670. https://doi.org/10.1371/journal.pone.0184670

Roberts, J.R (2004). Factors Affecting Egg Internal Quality and Egg Shell Quality in Laying Hens. J. Poult. Sci*,* 41, 161−177. https://doi:10.2141/jpsa.41.161

Salawu, E. O., M. Abdulraheem, A. Shoyombo, A. Adepeju, S. Davies, O. Akinsola and B. Nwagu (2014). Using Artificial Neural Network to Predict Body Weights of Rabbits. Open J. Anim. Sci, 4, 182–186. https://doi: 10.4236/ojas.2014.44023

Savegnago, R.P., B. N. Nunes, S. L. Caetano, A.S. Ferraudo, G.S. Schmidt, M.C. Ledur *et al*. (2011). Comparison of logistic and neural network models to fit to the egg production curve of White Leghorn hens. Poult. Sci. 90, 705−711. https://doi:10.3382/ps.2010-00723

Sengul, T., S. Celik., E. Eyduran and F. Iqbal (2020). Predicting egg production in Chukar partridges using nonlinear models and multivariate adaptive regression splines (MARS) algorithm. Europ. Poult. Sci, 84. https://doi: 10.1399/eps.2020.302

Shahinfar S., H. Mehrabani-Yeganeh., C. Lucas., A. Kalhor., M. Kazemian and K.A. Weigel. (2012). Prediction of breeding values for dairy cattle using artificial neural networks and neuro-fuzzy systems. Comput. Math. Methods Med.4, 1-9. https://doi: 10.1155/2012/127130

Sun, C., J. Liu, N. Yang and G. Xu (2019). Egg quality and egg albumen property of domestic chicken, duck, goose, turkey, quail, and pigeon. Poult. Sci. 0,1–6. https://doi:10.3382/ps/pez259

Tabrizi, S.S and N Sancar (2017). Prediction of Body Mass Index: A comparative study of multiple linear regression, ANN and ANFIS models. 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW, 24-25 Budapest, Hungary. Procedia computer science, 120, 394−401.

Tranmer, M., J. Murphy., M. Elliot and M. Pampaka (2020) Multiple Linear Regression (2nd Edition); Cathie Marsh Institute Working Paper 2020-01. https://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/2020-1-multiple-linear-regression.pdf

Tırınk, C (2022). Comparison of bayesian regularized neural network, random forest regression, support vector regression and multivariate adaptive regression splines algorithms to predict body weight from biometrical measurements in Thalli sheep. Kafk as Univ Vet Fak Derg*,* 28, 411-419. https://doi:10.9775/kvfd.2022.27164

Vekić, M., D. Savić and S. Jotanović (2022). Phenotypic correlations between egg quality traits amid the laying phase of broiler breeder hens Marinko. Contemporary Agriculture. 71,13−19, https://doi:10.2478/contagri-2022-0003

Wang, J., H. Yue, S.Wu, H. Zhang and G.Qi (2017). Nutritional modulation of health, egg quality and environmental pollution of the layers. Anim Nutr. 3, 91-96. https://doi:10.1016/j.aninu.2017.03.001

Wray, C.M and A.L. Byers (2020). Methodological Progress Note: Classification and Regression Tree Analysis. J Hosp Med. 15(9):549-551. https://doi: 10.12788/jhm.3366. PMID: 32195654.

Yakubu, A (2010). Fixing multicollinearity instability in the prediction of body weight from morphometric traits of white Fulani cows, J. Cent. Eur. Agric, 11(4). https://doi.org/10.5513/jcea.v11i4.855

Yakubu, A and P. Nimyak (2020). Use of artificial neural network to model reproductive performance and mortality of non- descript rabbits. Acta Sci.

Anim. Sci. 42. https://doi.org/10.4025/actascianimsci.v42i1.47715

Yakubu, A., L. Dahloum, A. J.Shoyombo and U.M, Yahaya (2019). Modelling hatchability and mortality in muscovy ducks using automatic linear modelling and artificial neural network. J. Indonesian Trop. Anim. Agric. 44, 65−76. https://doi:10.14710/jitaa.44.1.65-76

Yakubu, A., O. I. A. Oluremi and E. I. Ekpo (2018). Predicting heat stress index in Sasso hens using automatic linear modeling and artificial neural network, Int J Biometeorol. https://doi:10.1007/s00484-018-1521-7

Zhang, X., B. Chelliappan.,S. Rajeswari and M. Antonysamy (2021). Recent Advances in Applications of Bioactive Egg Compounds in Nonfood Sectors. Front Bioeng Biotechnol, 9, 738993. https://doi: 10.3389/fbioe.2021.738993.