ISSN: 2754-477X

Journal of Food Technology & Nutrition Sciences

SCIENTIFIC Research and Community

Research Article

Open Access

Differentiation of Exotic Chicken Breeds based on Egg Quality Traits using Classification Tree Algorithm

Kefelegn Kebede* and Ashenafi Getachew Megersa

School of Animal and Range Sciences, Haramaya University, Ethiopia

ABSTRACT

This study aimed to evaluate the effect of breed on egg quality traits and to discriminate chicken breeds (i.e., Bovan Brouwn (BV), Fayoumi (FM) and Sasso (SS)) based on egg quality traits using classification tree algorithm (CTA). For this purpose, a total of 1696 eggs were utilized and the traits egg weight, albumen height, albumen weight, egg length, egg width, egg weight, shell weight, shell thickness, yolk colour, yolk height and yolk weight, were recorded. Significant differences ($P<0.05$) were found in egg weight, albumen height, albumen weight, and egg width, with Bovans exhibiting the highest LSM values, indicating superior egg quality. Conversely, Sasso demonstrated superior yolk and shell traits with the highest LSM values for yolk weight, yolk height, shell weight, and egg length. Yolk colour differed significantly, with Sasso displaying the darkest yolks. No significant differences were observed in shell thickness. The CTA effectively discriminated between breeds based on egg quality traits, with the model explaining 90% of the variance ($R^2 = 0.90$) in the training dataset and 76% ($R^2 = 0.76$) in the validation dataset. Key predictors identified included egg weight, albumen height, yolk weight, albumen weight, and egg length. The model achieved high classification accuracy, with ROC AUC values of 0.99 for Bovans, 1.00 for Fayoumi, and 0.98 for Sasso, underscoring its robustness. These findings highlight the critical role of specific egg quality traits in breed differentiation and suggest that CTA is a valuable tool for breeders to enhance productivity and quality by selecting for desirable traits. Further research into the genetic underpinnings of these traits could refine classification models and improve breeding outcomes.

***Corresponding author**

Kefelegn Kebede, School of Animal and Range Sciences, Haramaya University, Tel: +251900467171; Ethiopia.

Received: July 13, 2024; **Accepted:** July 20, 2024; **Published:** July 27, 2024

Keywords: Breed Differentiation, Egg Quality Traits, Machine Learning Algorithm, Poultry

Abbreviations:
AH: Albumen Height
AW: Albumen Weight
EL: Egg Length
EWi: Egg Width
EW: Egg Weight
SW: Shell Weight
ST: Shell Thickness
YC: Yolk Colour
YH: Yolk Height
YW: Yolk Weight
CTA: Classification Tree Algorithm
LSM: Least Square Means
BV: Bovan
FM: Fayoumi
SS: Sasso
ROC: Receiver Operating Characteristic
AUC: Area Under the Curve

Introduction

The statistical analysis of categorical response variables in livestock and poultry, such as mortality, calving ease, stillbirth, litter size, fertility, breed discrimination is commonly conducted using logistic regression, canonical discriminant analysis, and data transformation methods. These methodologies, however, depend on assumptions of normality, constant variance, linearity, and non-multicollinearity. Violations of these assumptions can result in errors in the interpretation of results. To address these challenges, it is crucial to complement traditional analyses with more robust statistical procedures that do not rely on these assumptions. Classification and Regression Tree (CART) technique, a method within the domain of data mining, offers a viable alternative. Unlike traditional methods, CART does not require assumptions of normality, constant variance, or linearity. It provides several advantages, including the ability to conduct multivariate analysis, its nonparametric nature, the capability to handle multiway splits, applicability to various types of variables, resilience to data transformations, and effective handling of missing values. Furthermore, CART offers a graphical representation of data, facilitating the interpretation of complex interactions [1-7].

Data mining algorithms, including CART, are widely used in business and medicine, and they also have significant applications in animal science research. For instance, used CART to identify factors affecting the number of lambs reared from a fertilized ewe. employed CART methods to predict Polish Merino lamb mortality between birth and weaning, while explored CART to evaluate the relationship between PrP genotypes and litter size in various breeds. used CART to analyze the effects of cage density, genotype, and season on Japanese quail fertility. Similarly, applied CART to assess how egg quality characteristics influenced fertility in Japanese quail, identifying specific egg dimensions associated with higher fertility rates [1,3,4,8-12].

CART has also been utilized for selecting measures to prepare cows for artificial insemination, evaluating relationships between milk yield and udder traits in goats, and identifying genetic, physiological, and environmental parameters influencing milk quality. Additionally, utilized ANOVA with transformations, converting calving ease data into a standardized normal distribution using frequencies and Snell's transformation to normalize the distribution and ensure homogeneity of variance [13-16].

Despite its potential, the application of CART in poultry science in Ethiopia remains limited. Moreover, there is a paucity of information on discriminating chicken breeds based on egg quality traits using the CART algorithm. Therefore, the objectives of this study are to evaluate the effect of breed on egg quality traits and to discriminate chicken breeds based on egg quality traits using the CART data mining algorithm. This information will aid breeders in making informed decisions about breeding strategies, focusing on selecting animals with desirable traits [17].

Materials and Methods

Description of the Study Area

The study was conducted at the poultry research farm of Haramaya University, situated 505 km to the east of Addis Ababa. The site is positioned at an elevation of 1980 meters above sea level, with coordinates of 90 26'N latitude and 420 3'E longitude. The average annual maximum and minimum temperatures are recorded at 23.4°C and 8.25°C, respectively, while the region experiences an average annual rainfall of 741.6 mm.

Chicken Breeds Management

For this experiment, three exotic chicken breeds - Bovan Brown, Fayoumi, and Sasso breeds - were utilized. These breeds were raised under uniform housing and feeding conditions to maintain consistency. Initially, the chickens were kept in brooder houses with incandescent heating lamps for the first eight weeks. Subsequently, they were moved to the grower house during the growth phase and to the layer house, utilizing the deep litter system, during the laying phase.

Throughout the experiment, the chickens had unrestricted access to water and they were also provided with a recommended feed level to meet their nutritional needs. Initially, a standard ration with 20% crude protein (CP) and 2800 Kcal/kg metabolizable energy (ME) was given for the first eight weeks. During the growth phase (9 to 20 weeks), the feed contained 16% CP and 2800 Kcal/kg ME. Finally, during the laying period, the feed was adjusted to 16.50% CP and 1750 Kcal/kg ME to cater to the specific requirements of the hens.

To safeguard the health and welfare of the chickens, they were vaccinated against major viral diseases. Furthermore, a veterinary professional closely monitored the chickens to ensure their optimal health throughout the study duration.

Data Collection and Traits Measured

A total of 1696 eggs (600 from Bovan Brown, 741 from Fayoumi, and 355 from Sasso) were used for this study. Albumen height, albumen weight, egg length, egg width, egg weight, shell weight, yolk colour, yolk height, yolk weight, and shell thickness were then recorded. Egg weight, egg length and width were measured on unbroken eggs using a sensitive weighing scale and digital calliper, respectively. Then, the internal egg quality parameters were measured by breaking out the eggs on flat glass. After their length was measured using a tripod micrometre, albumen and yolk were carefully separated from each other to measure albumen and yolk weight. Finally, the eggshell was gently washed and air-dried for 48 hours to determine the eggshell thickness.

Statistical Data Analysis

For all statistical analyses in this study, JMP Pro version 17.2 was used [18].

Analysis of Variance

Egg quality traits (EW, AH, AW, YH, YW, SW, YC, EL, and EWi) were subjected to one-way analysis of variance using the general linear model procedure of JMP Pro to determine the effect of breed. Treatment means were separated using Duncan's multiple range test at a 95% confidence interval.

The linear model employed was:

$$Y_{ij} = \mu + B_i + \varepsilon_{ij}$$

where:

Y_{ij} = Observed value of the egg quality trait

μ = Overall mean

B_i = Fixed effect of the i[th] breed (i = Bovans Brown, Fayoumi, and Sasso)

ε_{ij} = Random residual error term

Classification Tree Algorithm (CTA)

CTAs are analytic techniques and data-mining tools that facilitate the creation of easily comprehensible graphical models to describe and predict phenomena expressed on both nominal and ordinal scales. They are particularly useful for the preliminary selection of traits that have a statistical effect on the response variable.

The CTA process begins at the root node, which contains only the response variable and no fragmentation. The data is then split into binary nodes, recursively creating child nodes until homogeneous subsets are achieved. Each split aim to maximize homogeneity within the resulting subsets, continuing until the index of homogeneity meets specified criteria. This process involves assigning a predicted outcome class to each node and continuing the splits until further division is impossible. The final subsets, which are not subject to further division, are called terminal nodes (leaves). The number of leaves indicates the tree size, while the depth is determined by the number of edges between the root and the most distant leaves.

CTA uses p-values with a Bonferroni correction as the splitting criterion. Pruning is employed to remove redundant branches and enhance the accuracy of the model. The quality of the CTA models was assessed using several metrics: average squared error, cumulative lift, Kolmogorov–Smirnov statistics, misclassification rate, and the area under the ROC curve. Lower average squared error and misclassification rate, alongside higher cumulative lift, Kolmogorov–Smirnov statistics, and ROC area, indicated better model quality [7,13,19-22].

Results and Discussions

Analysis of Variance: Table 1 below presents the least square means (±SE) for various egg quality traits of the chicken breeds: Bovan Brown, Fayoumi, and Sasso. The traits showed a wide range of variability.

Table 1: Least Square Means (±SE) of Egg Quality Traits of the Exotic Breeds

Breed			
Traits(units)	Bovans	Fayoumi	Sasso
EW (g)	$63.46^{a}\pm0.21$	$44.33^{c}\pm0.18$	$62.07^{b}\pm0.27$
AH (cm)	$10.85^{a}\pm0.06$	$6.65^{c}\pm0.06$	$7.18^{b}\pm0.08$
AW (g)	$42.04^{a}\pm0.15$	$26.07^{c}\pm0.14$	$37.59^{b}\pm0.20$
YH (cm)	$14.36^{b}\pm0.04$	$14.00^{c}\pm0.04$	$15.05^{a}\pm0.06$
YW (g)	$15.71^{b}\pm0.07$	$14.34^{c}\pm0.06$	$18.41^{a}\pm0.09$
SW (g)	$4.99^{b}\pm0.03$	$4.00^{c}\pm0.03$	$5.14^{a}\pm0.05$
EL (cm)	$5.45^{b}\pm0.01$	$5.05^{c}\pm0.01$	$5.74^{a}\pm0.02$
EWi (cm)	$4.23^{a}\pm0.01$	$3.76^{c}\pm0.01$	$4.18^{b}\pm0.01$
ST (cm)	$0.81^{a}\pm0.01$	$0.79^{a}\pm0.01$	$0.78^{a}\pm0.01$
YC (-)	$1.23^{c}\pm0.02$	$1.37^{b}\pm0.02$	$1.54^{a}\pm0.03$

^{a,b,c} when different superscripts are indicated in the same row for a given trait, it means that there is a significant (P < 0.05) effect of breed. AH = albumen height, AW = albumen weight, EL = egg length, EW = egg weight, EWi = egg width, SW = shell weight, YC = yolk colour, YH = yolk height, YW = yolk weight, and ST = shell thickness.

Least square mean differences in egg weight, albumen height, albumen weight, and egg width among the breeds were significant (P < 0.05), which was in the order of BV > SS > FF. The effect of breed on egg weight, albumen height, albumen weight, and egg width has also been reported by which agrees with the findings of the present study. Bovans exhibited the highest LSM value for egg weight (63.46 g), whereas Sasso (62.07 g) and Fayoumi (44.33 g) demonstrated comparatively lower values. This suggests that Bovans and Sasso's breeds yield eggs with notably higher weights in comparison to Fayoumi. The average egg weight of Bovan and Sasso observed in the present study was found to be higher than the findings of [23-24].

Likewise, Bovans displayed the highest LSM value for albumen height (10.85 cm), followed by Sasso (7.18 cm), while Fayoumi exhibited the lowest value (6.65 cm). These findings indicate that the Bovans breed produces eggs with significantly greater albumen height relative to the other two breeds. The average height of albumen obtained in the present study was found to be higher than the findings reported by [23].

The analysis of albumen weight revealed that Bovans showed the highest LSM value (42.04 g), whereas Sasso (37.59 g) and Fayoumi displayed relatively lower values (26.07 g). Consequently, the Bovans breed tends to produce eggs with markedly higher albumen weight compared to the other two breeds. The average estimates of albumen weight obtained in the present study for Fayoumi agree with the findings of while the average albumen weight was lower than the findings of the present study for Bovan and Sasso [23,25].

Regarding egg width, Bovans presented the highest LSM value (4.23 cm), followed closely by Sasso (4.18 cm), with Fayoumi rendering the lowest value (3.76 cm). This suggests that the Bovans breed demonstrates a propensity for producing eggs with significantly greater widths in comparison to the other two breeds. The average estimate of egg width obtained in the present study was near the findings of, who observed an average egg width of 41.4 mm. While reported the average egg weight of RIR x Fayomi to be 56.5 g which was higher than the findings of the present study [23,25].

Significant differences were observed (P < 0.05) in least square mean values for yolk height, yolk weight, shell weight, and egg length across the breeds, with the order being SS > BV > FF. Significant (P<0.05) effect of breed on yolk height, yolk weight, shell weight, and egg length observed in the findings of the present study has also been reported by [23-27].

In terms of yolk weight (YW), Sasso displayed the highest LSM value (18.41 cm), followed by Bovans (15.71 cm), while Fayoumi exhibited the lowest value (14.34 cm). These findings indicate that the Sasso breed tends to yield eggs with significantly greater yolk weight in comparison to the other two breeds. Reported the average yolk weight of 12.69 g, which is lower than the current findings [23].

Regarding yolk height (YH), Sasso exhibited the highest LSM value (15.05 cm), followed by Bovans (14.36 cm), while Fayoumi demonstrated the lowest value (14.00 cm). These results suggest that the Sasso breed tends to produce eggs with significantly greater yolk height relative to the other two breeds. Similarly, Sasso demonstrated the highest LSM value for shell weight (5.14 g), followed by Bovans (4.99 g), with Fayoumi displaying the lowest value (4.00 g). This indicates that the Sasso breed tends to produce eggs with significantly higher shell weights compared to the other two breeds. Regarding egg length, Sasso showcased the highest LSM value (5.74 cm), followed by Bovans (5.45 cm), while Fayoumi exhibited the lowest value (5.05 cm). These outcomes suggest that the Sasso breed has a propensity for yielding eggs with significantly greater lengths in comparison to the other two breeds. The average egg length reported in the available literature by is lengthier than the findings of the present study [23,25,26].

Shell thickness did not show a significant effect of breed (p > 0.05). The mean shell thickness obtained in the current findings was higher than that found by [28,29].

Lastly, yolk colour exhibited a significant effect of breed (p < 0.05), which was in the order of SS > FF > BV. Sasso had the highest LSM value for yolk colour (1.54), followed by Fayoumi (1.37), and Bovans had the lowest value (1.23). These findings suggest that the Sasso breed tends to produce eggs with significantly darker yolk colour compared to the other two breeds. The average egg length obtained in the present study was found to be higher than the findings of [26].

Classification Tree Algorithm (CTA)

The study employed a CTA to discriminate between chicken breeds based on various egg quality traits. This analysis aimed to partition the egg quality traits and evaluate their respective contributions to the differentiation of chicken breeds. The performance of the CTA model was assessed using both training and validation datasets, yielding promising results.

The training dataset comprised 1272 observations and resulted in a classification tree with 14 splits. The model explained 90% of the variance ($R^2 = 0.90$) in egg quality traits among the breeds in the training set. This high R^2 value indicates that the CTA model effectively captured the relationships between egg quality traits and breed differentiation. These findings are consistent with previous research that has demonstrated the utility of machine learning algorithms in poultry science for trait prediction and classification [13,19].

For the validation dataset, which included 424 observations, the model explained 76% of the variance ($R^2 = 0.76$). Although slightly lower than the training dataset's R^2, this value still indicates a good fit between the egg quality traits and breeds, confirming the robustness and generalizability of the CTA model. The reduction in R^2 is expected and common when transitioning from training to validation datasets due to the potential overfitting of the model to the training data [3].

The CTA diagram (Figure 1) constructed in the study provides a visual representation of the relationships between the egg quality traits and the breeds. This diagram is a valuable tool for understanding how specific traits contribute to breed differentiation. For instance, it visually elucidates which traits are pivotal for splitting the dataset at various nodes, thereby highlighting the critical predictors of breed classification. The ability of the CTA to visually and statistically elucidate these relationships makes it a powerful tool in poultry breeding research [5].

The high R^2 values obtained in both the training and validation datasets underscore the importance of egg quality traits in breed differentiation. These results suggest that breeders can rely on such models for making informed decisions about selection and breeding strategies. By focusing on the key egg quality traits identified by the CTA, breeders can enhance the productivity and quality of their flocks. This aligns with the findings of other studies that have applied machine learning techniques to animal breeding and genetics, demonstrating their potential to improve breeding outcomes [7].

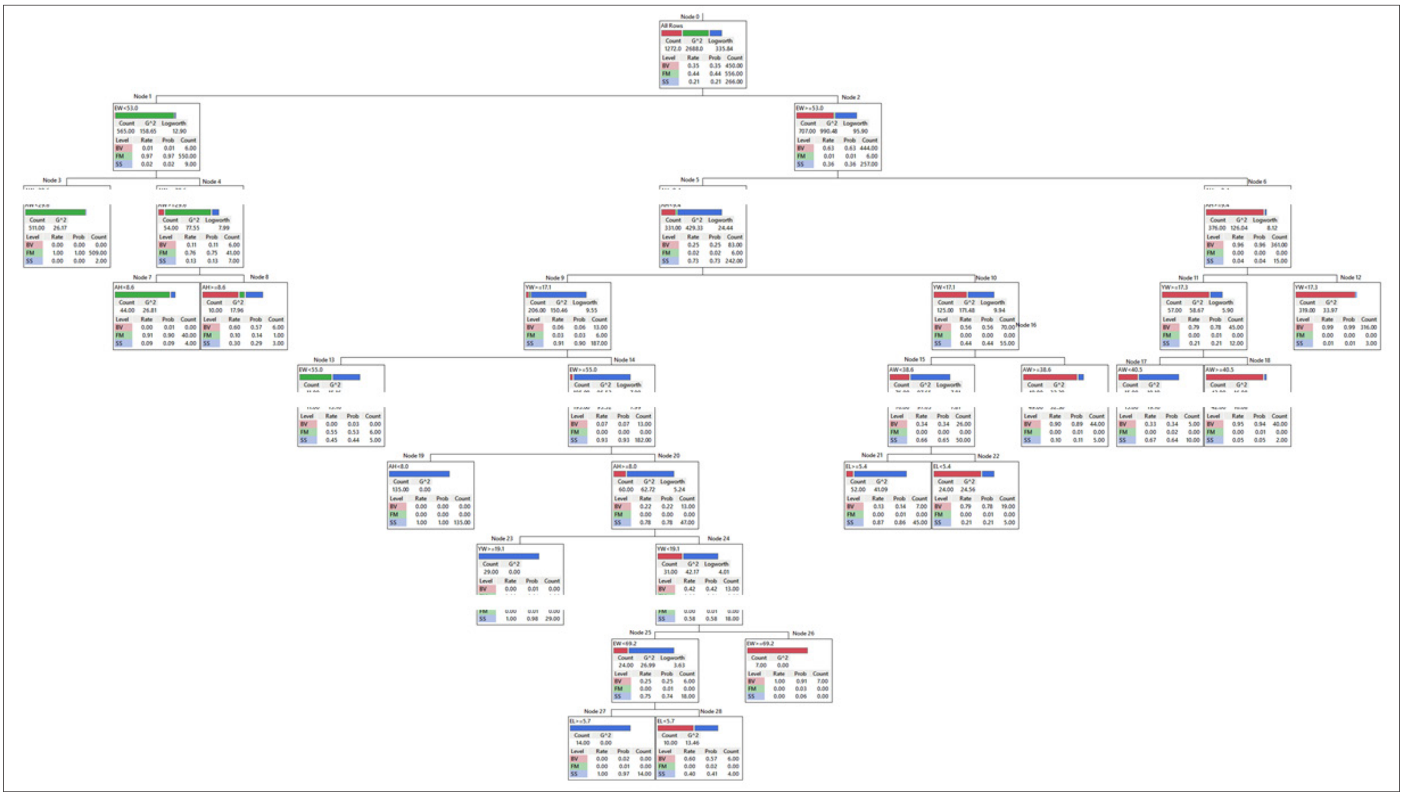

Figure 1: Classification Tree Diagram Constructed by CTA

The root node (Node 0) encompasses all 1,272 observations, with a G-squared value of 2,688 and a logworth value of 862.8, indicating a significant relationship between egg quality traits and chicken breeds. Node 1 includes 565 observations with egg weight less than 53g, displaying a G-squared value of 158.7 and a logworth value of 13.9. This node is further divided into Node 2 and Node 3.

Node 2 contains observations with egg weight less than 53g and albumen weight below 29.6g. There are 511 observations in this node, with a G-squared value of 26.2. Node 3 represents observations with egg weight less than 53g and albumen weight of 29.6g or more. This node includes 54 observations, with a G-squared value of 77.6 and a logworth value of 6.8.

Node 4, which includes 707 observations with egg weight of 53g or more, has a G-squared value of 990.5 and a logworth value of 158.2. This node splits into Node 5 and Node 6.

Node 5 contains observations with egg weight of 53g or more and albumen height less than 9.4 cm. There are 331 observations in this node, with a G-squared value of 429.3 and a logworth value of 34.8. Node 6 represents observations with egg weight of 53g or more and albumen height of 9.4 cm or more, comprising 376 observations, with a G-squared value of 126.0 and a logworth value of 8.2.

Node 6 is further subdivided into Node 7 and Node 8. Node 7 includes observations with egg weight of 53g or more, albumen height of 9.4 cm or more, and yolk weight of 17.3g or more. This node contains 57 observations, with a G-squared value of 58.7 and a logworth value of 5.2. Node 8 represents observations with egg weight of 53g or more, albumen height of 9.4 cm or more, and yolk weight less than 17.3g. This node includes 319 observations, with a G-squared value of 34.0.

Node 5 also splits into Node 9, which contains observations with egg weight of 53g or more, albumen height less than 9.4 cm, and yolk weight of 17.1g or more. This node has 206 observations, with a G-squared value of 150.5 and a logworth value of 8.9. Node 9 is further divided into Node 10 and Node 11. Node 10, which includes observations with yolk colour less than 12, has 137 observations and a G-squared value of 37.1. Node 11 represents observations with yolk colour of 12 or more, comprising 69 observations, with a G-squared value of 90.2 and a logworth value of 6.6.

Node 1 is further divided based on yolk weight into Node 16 and Node 17. Node 16 includes observations with yolk weight less than 15.5g, containing 282 observations with a G-squared value of 22.1. Node 17 represents observations with yolk weight of 15.5g or more, with 283 observations, a G-squared value of 32.3, and a logworth value of 2.4.

Node 16 splits into Node 18 and Node 19. Node 18 includes observations with yolk colour less than 10, containing 220 observations with a G-squared value of 15.0. Node 19 represents observations with yolk colour of 10 or more, comprising 62 observations with a G-squared value of 1.9.

Node 17 divides further into Node 20 and Node 21. Node 20 includes observations with albumen weight less than 30.1g, containing 187 observations with a G-squared value of 27.1 and a logworth value of 2.0. Node 21 represents observations with albumen weight of 30.1g or more, with 96 observations and a G-squared value of 4.0.

Node 20 is subdivided by yolk colour into Node 22 and Node 23. Node 22 includes observations with yolk colour less than 11, containing 127 observations with a G-squared value of 19.4. Node 23 represents observations with yolk colour of 11 or more, comprising 60 observations with a G-squared value of 4.9. Node 2 also splits based on yolk weight into Node 24 and Node 25.

For practical use, graphical trees should be smaller with clearly arranged decision steps to simplify interpretations for farmers and consultants [30].

Variable Importance

The findings from the current study underscore the critical importance of specific egg quality traits in differentiating between chicken breeds. The data presented in Table 2 reveal that egg weight (EW), albumen height (AH), yolk weight (YW), albumen weight (AW), and egg length (EL) are key predictors in the classification model, while other traits such as yolk colour (YC), yolk height (YH), egg width (EWi), shell weight (SW), and shell thickness (ST) do not significantly contribute to breed differentiation.

Egg weight emerged as the most influential predictor, with three splits and a portion of contribution of 0.66. This high level of importance suggests that variations in egg weight are a primary factor in distinguishing among the three breeds (BV, FM, and SS). This finding aligns with previous studies that have highlighted egg weight as a significant trait in poultry breeding and classification [13,19].

Albumen height and yolk weight also play substantial roles, each contributing through three splits with portions of 0.21 and 0.07, respectively. These results indicate that higher albumen height and specific ranges of yolk weight are associated with certain breeds, reinforcing the findings of prior research which emphasized the importance of albumen and yolk quality traits in breed differentiation. The contribution of albumen weight, although smaller (0.05), further underscores the relevance of albumen-related traits in the classification model [3].

Egg length, despite having a smaller impact (0.02 portion of contribution), still contributes to breed differentiation with its two splits. This suggests that while egg length is less critical than other traits, it still provides valuable information for the classification tree model. This finding is consistent with studies that have identified egg dimensions as useful predictors in poultry classification models [7].

On the other hand, traits such as yolk colour, yolk height, egg width, shell weight, and shell thickness did not contribute to the classification model, as indicated by their zero splits and zero portion of contribution. This lack of significance could be due to the homogeneity of these traits across the breeds studied or their lesser impact on breed differentiation. This aligns with findings from previous studies where certain egg quality traits were found to have minimal impact on classification outcomes [5].

Table 2: Egg Quality Traits' Important in Differentiation the Chicken Breeds

Trait	Number of Splits	G2	Portion
EW	3	1593.8	0.66
AH	3	500.7	0.21
YW	3	161.3	0.07
AW	3	120.0	0.05
EL	2	45.5	0.02
YC	0	0	0
YH	0	0	0
EWi	0	0	0
SW	0	0	0
ST	0	0	0

AH = albumen height, AW = albumen weight, EL = egg length, EW = egg weight, EWi = egg width, SW = shell weight, YC = yolk colour, YH = yolk height, YW = yolk weight, and ST = shell thickness.

Terminal Leaf Report

The application of the Classification Tree Algorithm (CTA) to discriminate between chicken breeds based on egg quality traits has yielded promising results, highlighting the effectiveness of this method in identifying breed-specific characteristics. The terminal leaf report provided probabilities and count values for each breed prediction, demonstrating the model's accuracy and revealing significant patterns in egg quality traits.

The results indicate that egg quality traits are effective discriminators of chicken breeds. For instance, in the leaf labelled "EW<53.0&AW<29.6," the model predicts a 100% probability of the Fayoumi breed, with a count of 509 observations. This suggests that lower egg weight (EW) and albumen weight (AW) are strong indicators of the Fayoumi breed. Similarly, in the leaf labelled "EW≥53&AH≥9.4&YW<17.3," the model predicts a 99% probability of the Bovan breed, with a count of 316 observations.

This indicates that the Bovan breed is more likely associated with higher egg weight and albumen height (AH).

The analysis also reveals several interesting patterns. Nodes with low egg weight and low albumen weight are more likely to be associated with the Fayoumi breed, while nodes with high egg weight and high albumen height are more likely to be associated with the Bovan breed. Additionally, the inclusion of yolk weight (YW) as a predictor appears to be significant in further distinguishing the breeds, with higher yolk weight values being associated with the Sasso breed. These findings are consistent with previous research that has highlighted the importance of specific egg quality traits in breed differentiation [13,19].

The model's ability to predict breed-specific probabilities with high accuracy underscores the potential of CTA in poultry science. For breeders, these insights are valuable for making informed decisions about breeding strategies. By identifying and selecting for desirable egg quality traits, breeders can improve the overall productivity and quality of their flocks. The patterns identified in this study, such as the association of low egg weight with the Fayoumi breed and high albumen height with the Bovan breed, can guide selective breeding practices to enhance specific traits.

Furthermore, the inclusion of yolk weight as a key predictor for the Sasso breed suggests that yolk-related traits may play a crucial role in breed differentiation. This aligns with studies that have examined the genetic basis of egg quality traits and their impact on breed characteristics. Future research could explore additional egg quality traits and their genetic underpinnings to further refine the classification models and enhance their predictive power [3,5].

In conclusion, the high accuracy of the CTA in discriminating between chicken breeds based on egg quality traits highlights its utility in poultry breeding. The patterns identified in this study provide valuable insights for breeders, enabling more targeted selection of desirable traits. Further research into the genetic factors influencing these traits will help to refine classification models and improve breeding outcomes.

Table 3: Terminal Leaf Report with Probability and Corresponding Count Values

	Response Probability			Response Count		
	BV	FM	SS	BV	FM	SS
EW<53&AW<29.6	0.00	1.00	0.00	0	509	2
EW<53&AW≥29.6&AH<8.6	0.01	0.90	0.09	0	40	4
EW<53&AW≥29.6&AH≥8.6	0.57	0.14	0.29	6	1	3
EW≥53&AH<9.4&YW≥17.1&EW<55	0.03	0.53	0.44	0	6	5
EW≥53&AH<9.4&YW≥17.1&EW≥55&AH<8	0.00	0.00	1.00	0	0	135
EW≥53&AH<9.4&YW≥17.1&EW≥55&AH≥8&YW≥19.1	0.01	0.01	0.98	0	0	29
EW≥53&AH<9.4&YW≥17.1&EW≥55&AH≥8&YW<19.1&EW<69.2&EL≥5.7	0.02	0.01	0.97	0	0	14
EW≥53&AH<9.4&YW≥17.1&EW≥55&AH≥8&YW<19.1&EW<69.2&EL<5.7	0.57	0.02	0.41	6	0	4
EW≥53&AH<9.4&YW≥17.1&EW≥55&AH≥8&YW<19.1&EW≥69.2	0.91	0.03	0.06	7	0	0
EW≥53&AH<9.4&YW<17.1&AW<38.6&EL≥5.4	0.14	0.01	0.86	7	0	45
EW≥53&AH<9.4&YW<17.1&AW<38.6&EL<5.4	0.78	0.01	0.21	19	0	5
EW≥53&AH<9.4&YW<17.1&AW≥38.6	0.89	0.01	0.11	44	0	5
EW≥53&AH≥9.4&YW≥17.3&AW<40.5	0.34	0.02	0.64	5	0	10
EW≥53&AH≥9.4&YW≥17.3&AW≥40.5	0.94	0.01	0.05	40	0	2
EW≥53&AH≥9.4&YW<17.3	0.99	0.00	0.01	316	0	3

AH = albumen height, AW = albumen weight, EL = egg length, EW = egg weight, and YW = yolk weight.
BV = Bovan brown; FM = Fayoumi; and SS = Sasso.

Confusion Matrix for the Training Data

Table 4: Confusion Matrix for the Training Data Showing Count and Percentages

Breed	Bovan	Fayoumi	Sasso
Bovan	438 (97)	0 (0)	12 (3)
Fayoumi	1 (0)	555 (100)	0 (0)
Sasso	22 (8)	11 (4)	233 (88)

Values in brackets are in percent.

The model's performance of the CTA was evaluated using a confusion matrix, which provides insights into the accuracy and misclassification rates for each breed in the training dataset. For the Bovan Brown breed, the model achieved a classification accuracy of 97%, correctly identifying 438 out of 451 instances. This high accuracy indicates the model's strong ability to distinguish Bovan Brown eggs with minimal errors. Notably, no Bovan Brown eggs were misclassified as Fayoumi, but there were 12 instances (3%) where Bovan eggs were incorrectly classified as Sasso. This suggests a slight overlap in the characteristics of Bovan and Sasso eggs, which could be an area for further refinement in the model.

The Fayoumi breed demonstrated perfect classification accuracy, with all 555 instances correctly identified, resulting in a 100% accuracy rate. The absence of any misclassifications for Fayoumi eggs underscores the model's exceptional performance in distinguishing this breed. This flawless classification aligns with previous research where high accuracy rates were achieved using classification algorithms in livestock studies [13,19].

For the Sasso breed, the model correctly classified 233 out of 266 instances, yielding an accuracy rate of 88%. While this is slightly lower than the rates for Bovan Brown and Fayoumi, it still represents a strong performance. The misclassification rates for Sasso—8% as Bovan and 4% as Fayoumi—indicate areas where the model could be improved. Similar challenges in classification accuracy have been noted in previous studies, where complex biological traits lead to overlapping characteristics among breeds [7].

Overall, the high accuracy rates achieved for each breed validate the effectiveness of the CTA in this context. The results are consistent with findings from other studies that have successfully applied classification tree methods to biological data, demonstrating their utility in animal breeding and genetics [3,5].

These findings are particularly significant for breeders, as the ability to accurately classify eggs based on breed-specific traits can aid in making informed breeding decisions. By identifying and selecting for desirable traits, breeders can improve the quality and productivity of their flocks. Future research could focus on refining the model to further reduce misclassification rates, particularly for breeds with higher overlap in characteristics.

Receiver Operating Characteristic (ROC) on Training Data

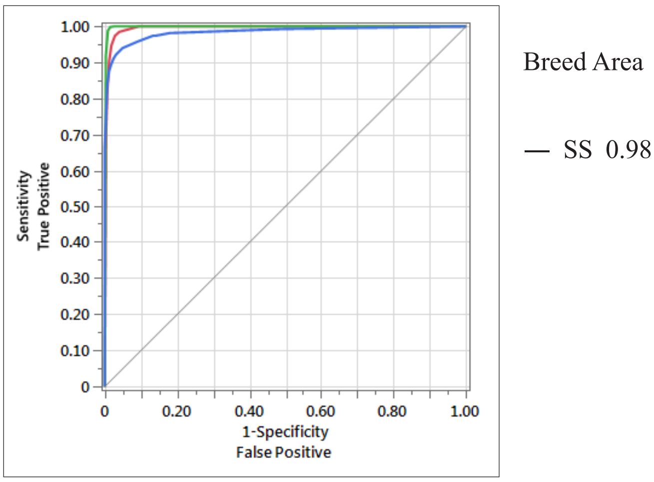

Figure 2: ROC Diagram Constructed by CTA

The model's performance of the CTA was also assessed through Receiver Operating Characteristic (ROC) curve analysis on the training data, providing insight into its discriminative power.

The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) for different threshold settings. The Area Under the Curve (AUC) is a key metric derived from the ROC curve, indicating the overall performance of the classification model. An AUC value of 1.0 represents perfect classification, whereas an AUC of 0.5 suggests no better than random guessing.

The results of the ROC analysis for each breed are as follows:
For the Bovan brown breed (BV), the ROC curve analysis yielded an AUC value of 0.99. This high AUC indicates that the classification tree algorithm can almost perfectly distinguish Bovan brown eggs from those of other breeds. The proximity of the AUC value to 1.0 demonstrates the model's excellent sensitivity and specificity in identifying Bovan brown eggs, with minimal false positive and false negative rates. The findings in the current study aligns with previous studies that have shown the effectiveness of CART in categorizing complex biological data with high accuracy [13,19].

The Fayoumi breed (FM) achieved an AUC value of 1.00. This perfect score suggests that the model can flawlessly differentiate Fayoumi eggs from those of other breeds in the training dataset. An AUC of 1.00 indicates that the model has achieved perfect sensitivity and specificity, meaning there are no misclassifications for the Fayoumi breed. This result is consistent with the findings of, who also reported high classification accuracy using CART for livestock traits [3].

For the Sasso breed (SS), the ROC curve analysis resulted in an AUC value of 0.98. Although slightly lower than that of Bovan and Fayoumi, this AUC still reflects a high level of accuracy in classifying Sasso eggs. The model demonstrates strong discriminative power, with only a small margin for improvement in reducing misclassifications. The strong discriminative power of the model for the Sasso breed, with only a small margin for improvement, underscores the robustness of CART in handling different datasets with high precision [5,7].

Conclusions

The current study used one-way ANOVA to evaluate the effect of breed on egg quality traits and classification tree algorithm to discriminate chicken breeds based on egg quality traits.

The study found significant differences ($P<0.05$) in egg weight, albumen height, albumen weight, and egg width among Bovans (BV), Sasso (SS), and Fayoumi (FF) breeds, in the order of BV > SS > FF. Bovans exhibited the highest least square mean (LSM) values for egg weight (63.46 g), albumen height (10.85 cm), albumen weight (42.04 g), and egg width (4.23 cm), indicating superior egg quality compared to Sasso and Fayoumi. Conversely, significant differences ($P<0.05$) in yolk height, yolk weight, shell weight, and egg length were observed, with the order SS > BV > FF. Sasso had the highest LSM values for yolk weight (18.41 g), yolk height (15.05 cm), shell weight (5.14 g), and egg length (5.74 cm), suggesting superior yolk and shell traits. Yolk colour also differed significantly ($P<0.05$), with Sasso displaying the darkest yolk. No significant differences were found in shell thickness. These findings align with previous studies and highlight breed-specific egg characteristics.

The CTA revealed distinct subgroups of observations based on the predictor variables AW, YW, and SW, highlighting their contribution to the differentiation of chicken breeds. The results of the CTA indicate that the predictor variables, namely albumen weight (AW), yolk weight (YW), and shell weight (SW), play significant roles in differentiating the chicken breeds.

The results of CTA provide a comprehensive understanding of how various egg quality traits interact to influence the differentiation of the chicken breeds. The hierarchical structure of the model allows for the identification of key factors and their relative importance in differentiating the chicken breeds, which can be valuable for researchers, farmers, and industry professionals working with egg production and quality.

The ROC curve analysis confirms the high effectiveness of the classification tree algorithm in differentiating between the three chicken breeds based on egg quality traits. The AUC values of 0.99 for Bovan, 1.00 for Fayoumi, and 0.98 for Sasso underscore the model's robustness and reliability. These results highlight the potential of egg quality traits as reliable predictors for breed discrimination, which can have significant implications for enhancing breeding and management practices in poultry production.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Financial Support
This research was conducted by the financial support of Haramaya University.

References

1. Piwczyński D (2009) Improvement of performance traits in Polish Merino. Postdoctoral dissertation 135, Appl. Sci. Rep. of University of Technology and Life Sciences in Bydgoszcz, Poland [in Polish, summary in English 1-122.
2. Schreurs NM, Kenyon PR, Mulvaney FJ, Morel PCH, West DM, et al. (2010) Effect of birthweight and birth rank on the survival of single and twin lambs born to ewe lambs. Anim Prod Sci 50: 460-464.
3. Piwczyński D, Nogalski Z, Sitkowska B (2012) Statistical modelling of calving ease and stillbirths in dairy cattle using the classification tree technique. Livestock Science 149: 131-138.
4. Uckardes F, Narinc D, Kucukonder H (2014) Application of classification tree method to determine factors affecting fertility in Japanese quail eggs. Journal of Animal Science Advances 4:1017-1023.
5. Topal M, Aksakal V, Bayram B, Yanar M (2010) An analysis of the factors affecting birth weight and actual milk yield in Swiss Brown cattle using regression tree. Journal of Animal and Veterinary Advances 9: 285-291.
6. Koyuncugil AS and N Ozgulbas (2010) Surveillance Technologies and Early Warning Systems: Data Mining Applications for Risk Detection. IGI Global USA 160-164.
7. Bayram B, Yanar M, Tüzemen N, Şahin A (2015) Using classification and regression tree method for predicting birth weight in Brown Swiss cattle. Indian Journal of Animal Research 49: 541-545.
8. Feldman D, Gross S (2003) Mortgage default: classification trees analysis. The Pinhass Sapir Center for Development Tel-Aviv University. Discussion Paper 3: 1-46.
9. Abu-Hanna A, de Keizer N (2003) Integrating classification trees with local logistic regression in Intensive Care prognosis. Artificial Intelligence in Medicine29: 5-23.
10. Austin PC (2007) A comparison of regression trees, logistic regression, generalized additive models, and multivariate adaptive regression splines for predicting AMI mortality. Stat Med 26: 2937-2957.
11. Grochowska E, Piwczyn´ski D, Portolano B, Mroczkowski S (2014) Analysis of the influence of the PrP genotype on the litter size in Polish sheep using classification trees and logistic regres- sion. Livest Sci 159: 11-17.
12. Celik S, Sogut B, Sengul T, Eyduran E, Sengul AY (2016) Usability of CART algorithm for determining egg quality characteristics influencing fertility in the eggs of Japanese quail. R Bras Zootec 45: 645-649.
13. Grzesiak W, Zaborski D (2012) Examples of the use of data mining methods in animal breeding. Annals of Animal Science 12: 375-388.
14. Eyduran E, I Yilmaz, MM Tariq, A Kaygisiz (2013) Estimation of 305-d milk yield using regression tree method in Brown Swiss cattle. J Anim Plant Sci 23: 731-735.
15. Sawa R, Doi T, Asai T, Watanabe K, Taniguchi T, et al. (2015) Differences in trunk control between early and late pregnancy during gait. Gait & Posture 42: 455-459.
16. Fiedlerova M, Rehak D, Vacek M, Volek J, Fiedler J Simecek, et al. (2008) Analysis of non-genetic factors affecting calving difficulty in the Czech Holstein population. Czech J Anim Sci 53: 284-291.
17. Abbas A, Ullah MA and Waheed A (2021) Body weight prediction using different data mining algorithms in Thalli sheep: A comparative study. Veterinary World 14: 2332-2338.
18. SAS Institute Inc. JMP Pro version 17.2 [Computer Software].
19. Eyduran E, Topal M, Sonmez AY (2016) Use of factor scores in multiple regression analysis for estimation of body weight by several body measurements in brown swiss cattle. Indian Journal of Animal Research 50: 765-769.
20. Orucoglu O (2011) Determination of environmental factors affecting 305-day milk yield of Holstein cows by regression tree method. Suleyman Demirel University Graduate School of Natural and Applied Sciences (Master Thesis).
21. Yadav SK, Bharadwaj BK and Pal S (2011) Data Mining Applications: A comparative study for predecting students' performance. Int J Inn Tech Creat Eng 1: 13-19.
22. Camdeviren HM, Mehmet MM, Ozkan FT, Toros T, Sasmaz T, et al. (2005) Determination of depression risk factors in children and adolescents by regression tree methodology. Acta Med Okayama 59: 19-26.
23. Sinha B (2014) Genetic impact on some of the reproductive and egg quality traits of Vanraja and Gramapriya birds and their crosses. M.V.Sc. Thesis submitted to Bihar Agriculture University, Sabour (Bhagalpur), Bihar for the degree of Master of Veterinary Science (Animal Genetics and Breeding).
24. Alam S (2015) Influence on genotypes on reproductive and egg quality traits of Gramapriya and its crosses with Desi chicken of Bihar. A Thesis submitted to Bihar Agricultural University, sabour. https://www.basu.org.in/wp-content/uploads/2020/07/INFLUENCE-OF-GENOTYPES-ON-REPRODUCTIVE-AND-EGG-QUALITY-TRAITS-OF-GRAMAPRIYA-AND-ITS-CROSSES-WITH-DESI-CHICKEN-OF-BIHARDR.MD_.SAMEER-ALAM14-12-2015.pdf.
25. Niranjan M, Sharma R, Rajkumar U, Chatterjee R, Reddy B, et al. (2008) Egg quality traits in chicken varieties developed for backyard poultry farming in India. Livest Res Rural Dev 20: 12-20.
26. Islam M and Dutta R (2010) Egg quality traits of indigenous, exotic and crossbred chickens (Gallus dometicus L.) in rajshahi, Bangladesh. J Life Earth Sci 5: 63-67.

27. Sharma RP (2014) Development of location specific chicken varieties for rural and tribal sector. Final report of Emeritus Scientist Scheme (ICAR) submitted to Bihar Agriculture University, Sabour (Bhagalpur), Bihar.
28. Jha DK and Prasad S (2013) Production performance of improved varieties and indigenous breed of chicken in Jharkhand. Indian J Poultry Sci 48: 109-112.
29. Alewi M, Melesse A, Teklegiorgis Y (2012) Crossbreeding Effect on Egg Quality Traits of Local Chickens and their F1 Crosses With Rhode Island Red and Fayoumi Chicken Breeds Under Farmpers' Management Conditions. J Anim Sci Adv 2: 697-705.
30. Yakubu A (2012) Application of regression tree methodology in predicting the body weight of Uda sheep. Anim Sci Biotech 45: 484-490.