Turkish Journal of Agriculture - Food Science and Technology, 10(4): 587-594, 2022
DOI: https://doi.org/10.24925/turjaf.v10i4.587-594.4670

Turkish Journal of Agriculture - Food Science and Technology

Available online, ISSN: 2148-127X | www.agrifoodscience.com | Turkish Science and Technology Publishing (TURSTEP)

Live Weight Prediction in Norduz Sheep Using Machine Learning Algorithms

Cihan Çakmakçı[1,2,a,*]

[1]*Department of Agricultural Biotechnology-Animal Biotechnology Unit, Faculty of Agriculture, Van Yüzüncü Yıl University, Van, Turkey.*
[2]*Laboratory of Biometeorology and Ethology, Department of Animal Science, Faculty of Animal Science and Food Engineering, University of Sao Paulo, Pirassununga, Brazil.*
[*]*Corresponding author*

A R T I C L E I N F O

Research Article

Received : 06/09/2021
Accepted : 13/01/2022

Keywords:
Machine learning
Variable importance
Correlation matrix
Morphological measurements
Live weight

A B S T R A C T

The objective of this study was to compare predictive performances of four machine learning (ML) models: Support Vector Machines with Radial Basis Function Kernel (SVMR), Classification and Regression Trees (CART), Random Forest (RF) and Model Average Neural Networks (MANN) to predict live weight from morphological measurements of Norduz sheep (n=93). Seven morphological measurements; chest girth (CG), chest width (CW), chest depth (CD), height at withers (HW), body length (BL), heigth at rump (HR) and rump width (RW) were used to predict live weigth (LW) of Norduz sheep. All morphological measurements were positively correlated to LW. Live weight had the highest correlation with CG and the lowest correlation with HR. Initially, highly correlated predictors were removed from the data set. The remaining predictors were then subjected to variable selection procedures using the Boruta algorithm. The results of Boruta confirmed the importance of the four predictors HW, BL, CW, and CD. However, HR confirmed to be unimportant was excluded from the dataset. The ML models were trained on selected predictors. The results showed that the prediction performance validated using the test dataset indicated that RF had the lowest values of Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percent Error (MAPE). The permutation-based variable importance scores indicate that CW and CD were the most important variables in predicting LW. The actual LW had the highest significant positive correlations with the values predicted by SVMR and RF, and followed by ANN and CART models respectively. There were no differences between the means of actual and predicted LWs by machine learning models. The fact that the models generalized well on the testing data sets indicates that machine learning algorithms have valid predictive patterns and are effective methods in LW weight of Norduz sheep. Considering runtime of the models, although the CART model had the lowest computational cost, it had the worst performance. The MANN algorithm, on the other hand, required a longer runtime to process the same dataset.

[a] *cakmakcicihan@gmail.com* *https://orcid.org/0000-0001-6512-9268*

Introduction

Live weight information plays an important role in managing performance-related parameters affecting livestock productivity, such as animal growth, uniform housing, space requirements, optimize feeding practices, monitoring herd health, and predicting and controlling marketing weights (Kashiha et al., 2014). Regular monitoring and management of changes in live weight in sheep significantly related to the productive and reproductive efficiency of the ewes, and to the progeny performance such as meat and fleece yield, as well as their effects on the economic profitability of the enterprise (Brown et al., 2015). As farms continue to grow in size, even small changes in production practices can have a large impact on the overall profitability of the business in animal production (Kashiha et al., 2014). Therefore, accurate monitoring of changes in body weight is critical for making effective management decisions for an efficient livestock production especially for meat production. Studies have shown that morphological measurements have significant correlations with body weight in sheep (Ferra et al., 2010; Mavule et al., 2013) and goat (Dakhlan et al., 2020) and are reliable features in the prediction of live weight in different animal species (Menesatti et al., 2014; Gomes et al., 2016; Fernandes et al., 2019; Meghelli et al., 2020). Since morphological measurements associated with body weight and body condition scores in farm animals (Chacón et al., 2011; Mekparyup et al., 2013; Olaniyi et al., 2018) can reflect growth, development and production performance as well as genetic characteristics, it is important to monitor changes in these values (Zhang et al., 2018). However, as the number of predictors increases with an interaction effect between two or more of them, the estimation task becomes more complicated. Machine learning provides ideally suited methodologies for extracting insights from this type of data (Valletta et al., 2017). Machine learning, also referred to as statistical learning, is a branch of artificial intelligence dedicated to the study of algorithms for prediction and inference (Morota et al., 2018). Additionally, machine learning can tackle a wide variety of tasks, including classification, clustering, and regression of an outcome of interest by multiple factors and elucidating the effects that contribute to it (Valletta et al., 2017). According to recent review (Benos et al., 2021), machine learning algorithms are classified into four categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Supervised machine learning, which is also known as predictive model, is used for tasks that involve the prediction of a given output using other features in the data set (Boehmke and Greenwell, 2019). When the objective of supervised learning is to predict the outcome of a numerical continuum, it is referred to as a regression problem (Boehmke and Greenwell, 2019), which has been used to predict live weight in cattle (Aytekin et al., 2018), goat (Dakhlan et al., 2020) and carcass traits from morphological measurements (Pabiou et al., 2011; Shahinfar et al., 2019; Lee et al., 2020). Using different machine learning algorithms body weights can be predicted with high accuracy by morphological measurements in different species of farm animals (Burke et al., 2004; Brown et al., 2015). Live weights were also predicted accurately using morphological measurements extracted from digital image processing, such as shoulder height, chest depth, and body length in sheep (Menesatti et al., 2014), body area pigs (Kashiha et al., 2014), several morphological measurements in cattle (Gomes et al., 2016; Miller et al., 2019), and lateral body surface images in buffaloes (Negretti et al., 2007).

The aim of this study is to predict live weights of Norduz sheep from morphological measurements using different machine learning algorithms. Validating the potential of machine learning methods for predicting live weight from morphological measurements could enable the development of prototype software for automated live weight prediction in the future via the integration of digital image processing techniques. Given the influence of farm management practices on production elements, it is very difficult to assess each individual animal adequately on a continuous and regular basis using conventional assessment methods that cause stress as livestock holdings expand. Supporting farmers via the application of precision livestock technology such as digital image processing and the integration of advanced machine learning algorithms may help increase livestock production efficiency, while also giving the farmer with more free time.

Materials and Methods

This study's dataset included morphological measurements taken from 93 Norduz ewes (aged 3-4 years), a fat-tailed breed native to the province of Van in Eastern Anatolia, Turkey. Seven morphological measurements were used to predict live weight (LW) of ewes. These morphological measurements were chest girth (CG), chest width (CW), chest depth (CD), height at withers (HW), body length (BL), height at rump (HR) and rump width (RW) (Table 1). Morphological measurements were carried out with the assistance of two handlers. The handlers ensured that the measures were taken accurately by keeping the ewe in the correct position. Live weights were measured manually on a static weighing platform after the morphological measurements were taken, to be used as a gold standard for predictive models. To minimize gutfill error, animals were deprived of feed for 12 h the night before weighing.

Statistical Analysis and Model Building

Modelling framework includes preprocessing the data, partitioning the data into training and testing sets, identifying optimal tuning parameters, building models, and estimating predictive performance (Kuhn and Johnson, 2019). The steps of the machine learning process are given in Figure 1. A major goal of the machine learning process is to find an algorithm that needed not only fits well to the past data, but more importantly, most accurately predicts future outcome, which is called the generalizability of the algorithm, based on a set of features (Boehmke and Greenwell, 2019). In order to evaluate the generalizability performance of the optimal model, the dataset was split into training and test datasets. Training dataset was used to select features, train algorithms, tune hyperparameters, and compare models as well as all of the other procedures needed to choose an optimal model (Boehmke and Greenwell, 2019), while the unseen testing data set was used to qualify performance of the models (Kuhn and Johnson, 2013).

Preprocessing

Initially, all variables were used as the candidate variables. The dataset was split into two parts: 75% training and 25% testing. The training data set (n=69) was used to create the model, while the unseen test data set (n=24) was used to qualify performance of the models.

Table 1. Summary of variable statistics

Value	LW (kg)	HW (cm)	BL (cm)	CD (cm)	CW (cm)	HR (cm)	RW (cm)	CG (cm)
Mean	54.7	72.9	66.7	33.2	21.1	71.1	22.7	93.0
Std.Dev.	8.3	2.6	3.3	2.1	2.2	2.9	1.9	5.7
Minimum	39.1	66.0	60.0	22.0	17.0	64.0	19.0	80.0
Median	53.9	73.0	66.5	33.0	21.0	71.0	22.5	93.0
Maximum	76.6	80.0	73.0	38.0	26.0	76.0	27.5	106.0

LW: live weight, HW: height at withers, BL: body length, CD: chest depth, CW: chest width, HR: height at rump, RW: rump width and CG: chest girth.

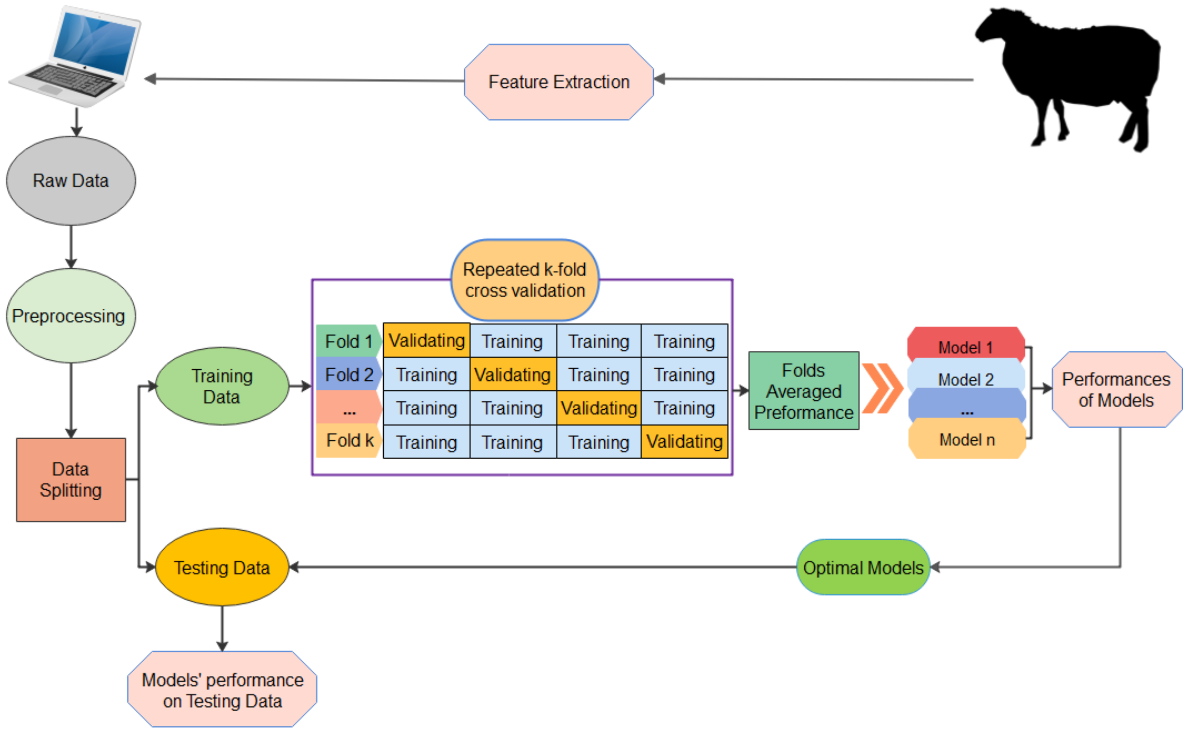


Figure 1. A workflow diagram that indicates each step of the machine learning process

Numeric variables were standardized by centering and scaling to have zero mean and unit variance, which provides a common comparable unit of measure across all variables. A correlation matrix was generated using the "ggstatsplot" R package (Patil, 2021) to keep all pairwise correlations below the threshold of 0.75 and eliminate highly correlated variables in order to avoid collinearity. The best subset of predictors for LW were selected from the training data using the Boruta algorithm implemented in "Boruta" R package (Kursa and Rudnicki, 2010). The machine learning models were then trained on the optimum subset. Hyperparameters, specific to the ML algorithm, were tuned by cross validation (Valletta et al., 2017). Repeated 10 times 10-fold cross validation resampling method was used to increase the precision of the estimated generalization error (Boehmke and Greenwell, 2019). Random Forest (RF), Support Vector Machines with Radial Basis Function Kernel (SVMR), Classification and Regression Trees (CART) and Model Average Neural Networks (MANN) were performed using "caret" R package (Kuhn, 2020) to predict live weight of Norduz sheep based on a set of selected morphological measurements. The "ggstatsplot" R package was used for the comparisons of live weight means predicted by machine learning models. All analysis were performed using R programming language, R version 4.1.2 (RCoreTeam, 2021).

Evaluation of Model Performance

The metric used to assess the effectiveness of a model to predict the outcome is very important and can influence the conclusions. The most effective approach to assessing model performance is to assess the predictive accuracy by loss functions (Boehmke and Greenwell, 2019; Kuhn and Johnson, 2019). Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percent Error (MAPE) were generated with the "Metrics" R package (Hamner and Frasco, 2018) and used as loss functions to assess the performance of predictive regression models. The "vip" R package was used to find the most influential variables in the models using permutation based variable importance (Greenwell and Boehmke, 2020). The "ggplot2" R package was used to visualize the results (Wickham, 2016).

Results and Discussion

Preprocessing and Model Building

Correlation matrix of whole data set indicated that LW had the highest correlation with CG (Figure 2A). In line with these results, Önk et al. (2018) reported that CG had the highest correlation with live weight in Tuj lambs. Higher correlation between CG and LW were also reported in cattle (Tebug et al., 2016; Weber et al., 2020a) and in goat (Abd-Allah et al., 2019). In the present study, however, the predictors CG and RW were the highly correlated variables and were removed from data set (Figure 2B). The result of Boruta analysis are given in Table 2 and displayed in Figure 3 indicated that the predictor HW, BL, CW and CD were variables confirmed to be important were selected as the final predictors. However the predictor HR was not important and removed from the dataset. Considering training data set, LW had the highest correlation with CD ($r=+0.77$, $P<0.05$) and CW ($r=+0.72$, $P<0.05$) fallowed by BL ($r=+0.57$, $P<0.05$) and the lowest with HR ($r=+0.33$, $P<0.05$).

Model performance, and correlations between actual and predicted live weights

Results of the present study suggested that morphological measurements have potential to predict LWs of sheep. Repeated 10 times 10-fold cross validation resampling results for test datasets across the models are given in Table 3. The results showed that the prediction accuracy validated using the test dataset indicated that RF model outperformed all other models with the lowest values of MAE, RMSE and MAPE. Although the CART models took a substantially less time to train, it was the worst performing model with the highest value of MAE, RMSE, and MAPE among the models.

In another study, Shahinfar et al. (2019) reported that using the correlation coefficient and MAE as predictive performance measures, random forest outperformed linear regression, deep learning, gradient boosting tree, k-nearest neighbor and model tree in predicting sheep carcass traits. In their study, Sant'Ana et al. (2021) found that the RF model, which had the lowest MAE values, outperformed the SVM model in predicting body weights in sheep using morphological data extracted from digital images. However, contrary to the results of the present study, Celik et al. (2017) reported that the RMSE and MAPE values of the CART model were lower than those of the ANN model in predicting body weight of Mengali rams using body length, withers height, heart girth, testicular length, scrotal length, and scrotal circumference as predictors. In another study, Huma and Iqbal (2019) compared the performance of generalized linear models, regression trees, support vector machines, and random forests models in predicting the body weight of Balochi sheep using body length, heart girth, withers height, scrotal diameter, scrotal circumference, scrotal length, and testicular length as predictor variables. Similar to the results of the present study, those researchers found that random forests had the lowest MAE, RMSE, and MAPE values on the test dataset, followed by support vector machine and regression trees. Iqbal et al. (2021) on the other hand found that the least square support vector machine model had the lowest values of MAE and RMSE on the test dataset when compared to the ANN model.

Table 2. Variable importance scores.

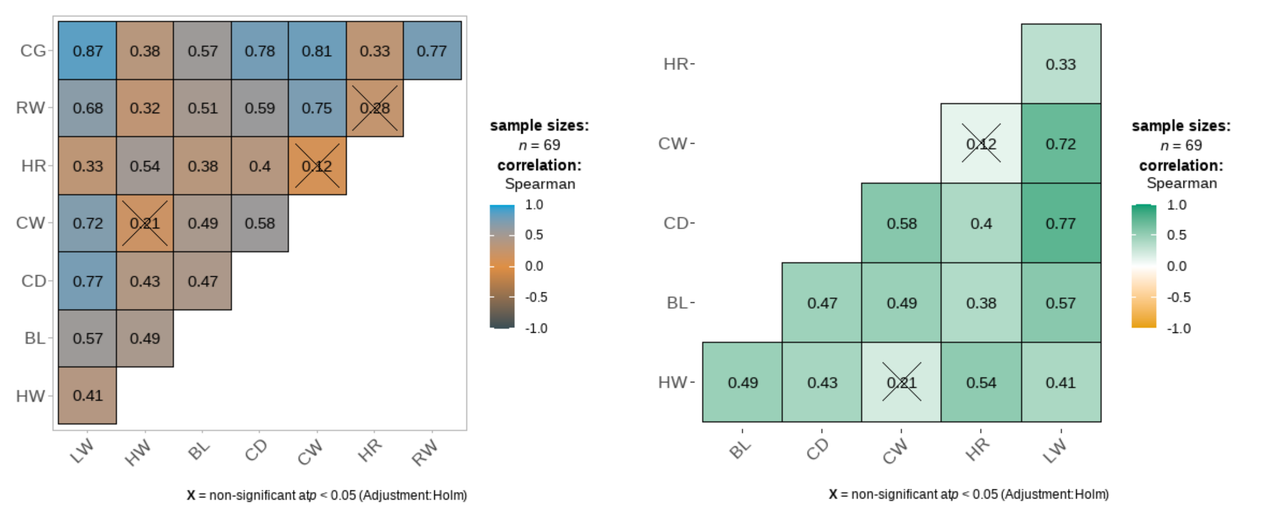
Variables	meanImp	medianImp	minImp	maxImp	normHits	decision
HW	3.872366	3.931149	1.039355	7.071214	0.8229167	Confirmed
BL	9.343600	9.428641	6.736520	11.586554	0.9947917	Confirmed
CD	21.228616	21.273306	18.972473	23.900741	1.0000000	Confirmed
CW	18.854771	18.841887	16.890068	21.162709	1.0000000	Confirmed
HR	1.453639	1.382400	-1.362311	4.791386	0.3697917	Rejected

HW: height at withers, BL: body length, CD: chest depth, CW: chest width and HR: height at rump.

Table 3. Performances of regression models on test dataset.

Model	MAE	RMSE	MAPE	Runtime	Tuning parameters
MANN	4.06	4.97	0.07	37.30857 secs	size = 1, decay = 0.1
CART	4.97	5.69	0.09	1.396198 secs	cp = 0.005
SVMR	3.94	5.07	0.07	2.857357 secs	sigma = 0.3415073, C = 2
RF	3.66	4.47	0.07	4.588363 secs	mtry = 2

MAE: mean absolute error, RMSE: root mean squared error, and MAPE: mean absolute percent error.



(a) (b)

Figure 2. a) Correlation matrix illustrates correlation coefficients between the features in training dataset and b) Correlation matrix illustrates correlation coefficients between features, after removing predictors with an absolute pair-wise correlation of 0.75 or higher in training dataset. CG: chest girth, CD: chest depth, BL: body length, CW: chest width, HR: height at rump, RW: rump width, HW: height at withers.

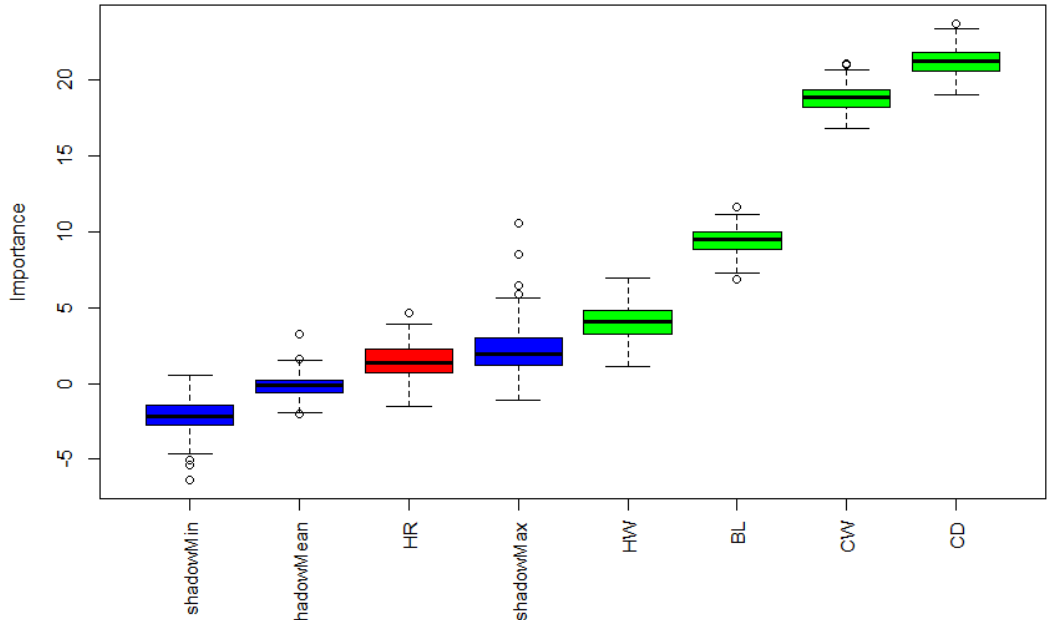


Figure 3. Plot of Boruta result. Blue boxplots correspond to minımum, mean and maximum Z scores of a shadow variable. Boxplots in red and green represent the Z scores of respectively rejected and confirmed variables.
HW: height at withers, BL: body length, CW: chest width and CD: chest depth are the four variables confirmed to be important. HR: height at rump is the only variable that is confirmed to be unimportant.

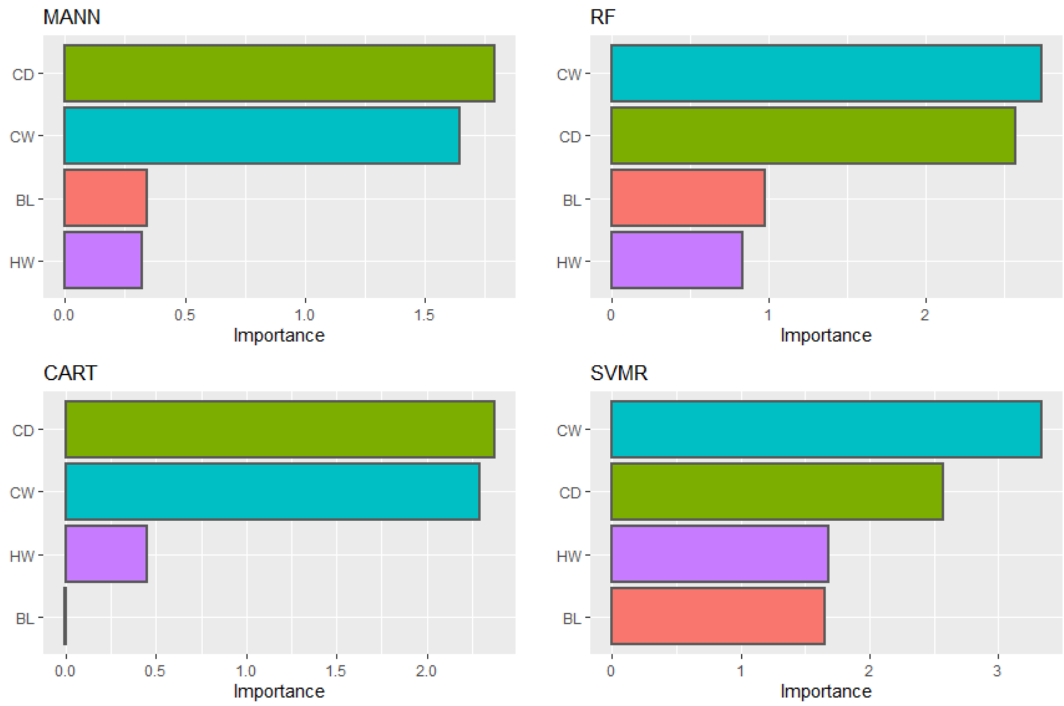


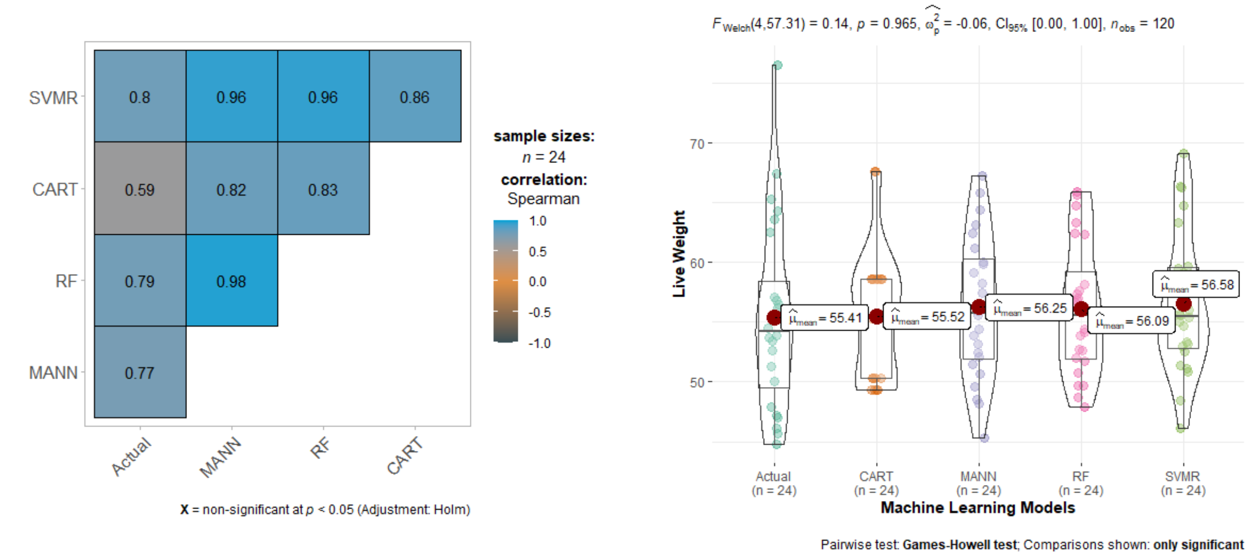
Figure 4. Variable importance scores for the predictors;
HW: height at withers, BL: body length, CD: chest depth and CW: chest width.

According to the permutation-based variable importance scores used in the present study, the most important predictors of live weight in all machine learning models were CD and CW, whereas the least important predictors were BL and HW (Figure 4). Contrary to the result of the present study, Huma and Iqbal (2019) reported variable importance scores for the random forest model only, with BL being the most important predictors, whereas HW and CG were the least important predictors. In their study, Iqbal et al. (2021) reported variable importance scores only for the ANN model, with HW being one of the most important predictors, whereas BL and CG were among the least important predictors. Although Iqbal et al. (2021) noted in their study that some explanatory variables were highly correlated, indicating multicollinearity, they developed models without feature selection using all measured variables. Therefore, these results are not reasonably comparable to the findings of the present study.

Even if a model is insensitive to a higher number of predictors, beside it is scientifically reasonable to include the minimum possible number of predictors that yield acceptable results, reducing the predictors can sometimes reduce the cost of acquiring data or improve the efficiency of the software used (Kuhn and Johnson, 2019). To address the question of the smallest possible number of variables required to achieve optimal predictive performance and overcome multicollinearity in the present study, correlation matrices and a random forest-based Boruta algorithm were employed to select the most relevant and important variables. Since the approach enables for variable selection to decide which variables are relevant and important for live weight prediction, this approach will save time and effort by eliminating the measurement of irrelevant and unimportant variables.

Correlation matrix shows that the actual live weight values had the highest significant positive correlations with the values predicted by SVMR (r=+0.80, P<0.05), RF (r=+0.79, P<0.05) and ANN (r=+0.77, P<0.05), followed with the lowest value by CART (r=+0.59, P<0.05) model (Figure 5a). However, the results of ANOVA show that there were no statistically significant differences between the means of actual and predicted live weights by machine learning models (Figure 5b). The models generalized well on the testing datasets, suggesting that they had valid predictive patterns, as they were also supported by significant high correlations between actual and predicted values. While repeated 10 times 10-fold cross validation highlighted good predictive performance, the findings varied across machine learning models. Different machine learning methods have been used to predict live weight in sheep (Celik et al., 2017; Huma and Iqbal, 2019; Iqbal et al., 2021; Sant'Ana et al., 2021) or fat tail weight in sheep (Norouzian and Vakili Alavijeh, 2016), cattle (Aytekin et al., 2018; Miller et al., 2019; Weber et al., 2020a) and rabbits (Salawu et al., 2014) reported different results. In the present study, the data was split into two parts as 75% training and 25% testing. Repeated 10 times 10-fold cross validation resampling method was used to increase the precision of the estimated generalization error. In this respect, it differs from previous studies which used different data splitting and model validation methods for prediction of body weight in sheep (Celik et al., 2017; Huma and Iqbal, 2019; Iqbal et al., 2021; Sant'Ana et al., 2021) and cattle (Cominotte et al., 2020; Lee et al., 2020; Weber et al., 2020a; Weber et al., 2020b) or of carcass traits in sheep (Shahinfar et al., 2019) and cattle (Pogorzelska-Przybyłek et al., 2014). Therefore, it was not feasible to draw an accurate comparison between the findings of the present study and the findings of earlier studies that reported results from the use of various machine learning approaches to predict live weight in sheep.

On a more practical level, all model development attempts are constrained by available data. For many situations, the data may have a small sample size, be of poor quality, or be unrepresentative of future samples (Kuhn and Johnson, 2013a). Due to the logistical challenges associated with collecting data on livestock, it is rather common to work with small datasets. The downsides of dealing with limited datasets for machine learning models, on the other hand, are related to overfitting and bias. Therefore, the limitation of the present study was that it was studied with a small data set. However, in order to evaluate the generalizability performance of the models, the dataset was split into training and test datasets. Training dataset was used to select features, train algorithms, tune hyperparameters, while the unseen testing data set was used to qualify performance of the models. Hyperparameters, specific to the ML algorithm, were tuned by repeated 10 times 10-fold cross validation resampling method to increase the precision of the estimated generalization error.



(a) (b)

Figure 5. a) Correlation matrix illustrates correlation coefficients between actual and predicted live weight by different machine learning models, b) Comparison of differences between live weights predicted with different machine learning models.

MANN: model averaged neural network, CART: classification and regression trees, RF: random forest and SVMR: support vector machines with radial basis function kernel.

Conclusion

This study demonstrates that machine learning approaches are effective methods in predicting LWs in Norduz sheep. The prediction performance validated using the test dataset indicate that random forest outperformed MANN, SVMR and CART models with the lowest values of mean absolute error, root mean squared error and mean absolute percent error. Although the CART model had the lowest computation cost, it performed the worst. The results also suggest that morphological measurements are highly significantly correlated with live weight and are reliable variables that can be used in machine learning methods to accurately predict LW in Norduz sheep. The study also shows that although all variables significantly correlate with LW, the Boruta algorithm can be used effectively to select which variables should be used in final models to compare machine learning algorithms. Since the approach enables for variable selection to decide which variables are relevant and important for live weight prediction, this approach will save time and effort by eliminating the measurement of irrelevant and unimportant variables. In light of these findings, future research might focus on identifying features from larger sample sizes (from animals at various stages of development) that are more easily obtained and relevant to the field, by integrating image analysis techniques and different machine learning algorithms, considering alternative feature selection methods.

Authorship Contribution Statement

The author alone is responsible for the content and writing of the paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The author declare that there is no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author would like to acknowledge his gratitude to the staff of research farm at Van Yüzüncü Yıl University for their assistance in animal handling. The author is grateful to three anonymous reviewers for the thorough and constructive comments on prior versions of this manuscript.

References

Abd-Allah S, Abd-El Rahman HH, Shoukry MM, Mohamed MI, Salman FM, Abedo AA. 2019. Some body measurements as a management tool for Shami goats raised in subtropical areas in Egypt. Bulletin of the National Research Centre, 43(1). doi: 10.1186/s42269-019-0042-9

Aytekin İ, Eyduran E, Karadas K, Akşahan R, Keskin İ. 2018. Prediction of Fattening Final Live Weight from some Body Measurements and Fattening Period in Young Bulls of Crossbred and Exotic Breeds using MARS Data Mining Algorithm. Pakistan Journal of Zoology, 50(1). doi: 10.17582/journal.pjz/2018.50.1.189.195

Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D. 2021. Machine Learning in Agriculture: A Comprehensive Updated Review. Sensors (Basel), 21(11). doi: 10.3390/s21113758

Boehmke B, Greenwell B. 2019. Hands-On Machine Learning with R. CRC Press.

Brown DJ, Savage DB, Hinch GN, Hatcher S. 2015. Monitoring liveweight in sheep is a valuable management strategy: a review of available technologies. Animal Production Science, 55(4). doi: 10.1071/an13274

Burke J, Nuthall P, McKinnon A. 2004. An analysis of the feasibility of using image processing to estimate the live weight of sheep. FHMG research report., Applied Management and Computing Division, Lincoln University. Canterbury, New Zealand.

Celik S, Eyduran E, Karadas K, Tariq MM. 2017. Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. Revista Brasileira de Zootecnia, 46(11):863-872. doi: 10.1590/s1806-92902017001100005

Chacón E, Macedo F, Velázquez F, Paiva SR, Pineda E, McManus C. 2011. Morphological measurements and body indices for Cuban Creole goats and their crossbreds. Revista Brasileira de Zootecnia, 40(8):1671-1679. doi: 10.1590/s1516-35982011000800007

Cominotte A, Fernandes AFA, Dorea JRR, Rosa GJM, Ladeira MM, van Cleef EHCB, Pereira GL, Baldassini WA, Machado Neto OR. 2020. Automated computer vision system to predict body weight and average daily gain in beef cattle during growing and finishing phases. Livestock Science, 232:103904. doi: 10.1016/j.livsci.2019.103904

Dakhlan A, Saputra A, Hamdani MDI, Sulastri S. 2020. Regression Models and Correlation Analysis for Predicting Body Weight of Female Ettawa Grade Goat using its Body Measurements. Advances in Animal and Veterinary Sciences, 8(11). doi: 10.17582/journal.aavs/2020/8.11.1142.1146

Fernandes AFA, Dorea JRR, Fitzgerald R, Herring W, Rosa GJM. 2019. A novel automated system to acquire biometric and morphological measurements and predict body weight of pigs via 3D computer vision. Journal of Animal Science, 97(1):496-508. doi: 10.1093/jas/sky418

Ferra JdC, Cieslak S, Sartori Filho R, McManus C, Martins CF, Sereno JRB. 2010. Weight and age at puberty and their correlations with morphometric measurements in crossbred breed Suffolk ewe lambs. Revista Brasileira de Zootecnia, 39(1):134-141. doi: 10.1590/s1516-35982010000100018

Gomes RA, Monteiro GR, Assis GJ, Busato KC, Ladeira MM, Chizzotti ML. 2016. Technical note: Estimating body weight and body composition of beef cattle trough digital image analysis. Journal of Animal Science, 94(12):5414-5422. doi: 10.2527/jas2016-0797

Greenwell BM, Boehmke BC. 2020. Variable Importance Plots—An Introduction to the vip Package. The R Journal, 12(1):343. doi: 10.32614/rj-2020-013

Hamner B, Frasco M. 2018. Metrics: Evaluation Metrics for Machine Learning. R package version 0.1.4. https://CRAN.R-project.org/package=Metrics.

Huma ZE, Iqbal F. 2019. Predicting the body weight of Balochi sheep using a machine learning approach. Turkish Journal of Veterinary and Animal Sciences, 43(4):500-506. doi: 10.3906/vet-1812-23

Iqbal F, Raziq A, Huma ZE, Ali M. 2021. An application of least square support vector machine model with parameters optimization for predicting body weight of Harnai sheep breed. Turkish Journal of Veterinary and Animal Sciences, 45(4):716-725. doi: 10.3906/vet-2009-105

Kashiha M, Bahr C, Ott S, Moons CPH, Niewold TA, Ödberg FO, Berckmans D. 2014. Automatic weight estimation of individual pigs using image analysis. Computers and Electronics in Agriculture, 107:38-44. doi: 10.1016/j.compag.2014.06.003

Kuhn M. 2020. caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret.

Kuhn M, Johnson K. 2013. Applied Predictive Modeling. Springer, New York.

Kuhn M, Johnson K. 2019. Feature Engineering and Selection: A Practical Approach for Predictive Models. CRC Press.

Kursa MB, Rudnicki WR. 2010. Feature Selection with the Boruta Package. Journal of Statistical Software, 36(11). doi: 10.18637/jss.v036.i11

Lee DH, Lee SH, Cho BK, Wakholi C, Seo YW, Cho SH, Kang TH, Lee WH. 2020. Estimation of carcass weight of Hanwoo (Korean native cattle) as a function of body measurements using statistical models and a neural network. Asian-Australasian Journal of Animal Sciences, 33(10):1633-1641. doi: 10.5713/ajas.19.0748

Mavule BS, Muchenje V, Bezuidenhout CC, Kunene NW. 2013. Morphological structure of Zulu sheep based on principal component analysis of body measurements. Small Ruminant Research, 111(1-3):23-30. doi: 10.1016/j.smallrumres. 2012.09.008

Meghelli I, Kaouadji Z, Yilmaz O, Cemal I, Karaca O, Gaouar SBS. 2020. Morphometric characterization and estimating body weight of two Algerian camel breeds using morphometric measurements. Tropical Animal Health and Production, 52(5):2505-2512. doi: 10.1007/s11250-020-02204-x

Mekparyup J, Saithanu K, Arunkeeree N. 2013. Estimation of body weight of holstein-friesian cattle with multiple regression analysis. International Journal of Applied Mathematics and Statistics, 44(14):1-7.

Menesatti P, Costa C, Antonucci F, Steri R, Pallottino F, Catillo G. 2014. A low-cost stereovision system to estimate size and weight of live sheep. Computers and Electronics in Agriculture, 103:33-38. doi: 10.1016/j.compag.2014.01.018

Miller GA, Hyslop JJ, Barclay D, Edwards A, Thomson W, Duthie C-A. 2019. Using 3D Imaging and Machine Learning to Predict Liveweight and Carcass Characteristics of Live Finishing Beef Cattle. Frontiers in Sustainable Food Systems, 3. doi: 10.3389/fsufs.2019.00030

Morota G, Ventura RV, Silva FF, Koyama M, Fernando SC. 2018. Big Data Analytics and Precision Animal Agriculture Symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. Journal of Animal Science, 96(4):1540-1550. doi: 10.1093/jas/sky014

Negretti P, Bianconi G, Bartocci S, Terramoccia S. 2007. Lateral Trunk Surface as a new parameter to estimate live body weight byVisual Image Analysis. Italian Journal of Animal Science, 6(sup2):1223-1225. doi: 10.4081/ijas.2007.s2.1223

Norouzian MA, Vakili Alavijeh M. 2016. Comparison of artificial neural network and multiple regression analysis for prediction of fat tail weight of sheep. Iranian Journal of Applied Animal Science, 6(4):895-900.

Olaniyi TA, Popoola MA, Olaniyi OA, Faniyi TO, Iniobong U. 2018. Relationship between morphological traits, body indices and body condition score as welfare indicators of Nigerian sheep. Journal of Animal Science and Veterinary Medicine, 3(1):1-5. doi: 10.31248/jasvm2017.078

Pabiou T, Fikse WF, Cromie AR, Keane MG, Näsholm A, Berry DP. 2011. Use of digital images to predict carcass cut yields in cattle. Livestock Science, 137(1-3):130-140. doi: 10.1016/j.livsci.2010.10.012

Patil I. 2021. Visualizations with statistical details: The 'ggstatsplot' approach. Journal of Open Source Software, 6(61):3167. doi: 10.21105/joss.03167

Pogorzelska-Przybyłek P, Nogalski Z, Wielgosz-Groth Z, Winarski R, Sobczuk-Szul M, Łapińska P, Purwin C. 2014. Prediction of the Carcass Value of Young Holstein-Friesian Bulls Based on Live Body Measurements. Annals of Animal Science, 14(2):429-439. doi: 10.2478/aoas-2014-0004

RCoreTeam. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Salawu EO, Abdulraheem M, Shoyombo A, Adepeju A, Davies S, Akinsola O, Nwagu B. 2014. Using Artificial Neural Network to Predict Body Weights of Rabbits. Open Journal of Animal Sciences, 04(04):182-186. doi: 10.4236/ojas.2014.44023

Sant'Ana DA, Pache MCB, Martins J, Soares WP, de Melo SLN, Garcia V, de Moares Weber VA, da Silva Heimbach N, Mateus RG, Pistori H. 2021. Weighing live sheep using computer vision techniques and regression machine learning. Machine Learning with Applications, 5:100076. doi: 10.1016/j.mlwa.2021.100076

Shahinfar S, Kelman K, Kahn L. 2019. Prediction of sheep carcass traits from early-life records using machine learning. Computers and Electronics in Agriculture, 156:159-177. doi: 10.1016/j.compag.2018.11.021

Tebug SF, Missohou A, Sourokou Sabi S, Juga J, Poole EJ, Tapio M, Marshall K. 2016. Using body measurements to estimate live weight of dairy cattle in low-input systems in Senegal. Journal of Applied Animal Research, 46(1):87-93. doi: 10.1080/09712119.2016.1262265

Valletta JJ, Torney C, Kings M, Thornton A, Madden J. 2017. Applications of machine learning in animal behaviour studies. Animal Behaviour, 124:203-220. doi: 10.1016/j.anbehav.2016.12.005

Weber VAdM, Weber FdL, Gomes RdC, Oliveira AdS, Menezes GV, Abreu UGPd, Belete NAdS, Pistori H. 2020a. Prediction of Girolando cattle weight by means of body measurements extracted from images. Revista Brasileira de Zootecnia, 49. doi: 10.37496/rbz4920190110

Weber VAM, Weber FdL, Oliveira AdS, Astolfi G, Menezes GV, de Andrade Porto JV, Rezende FPC, Moraes PHd, Matsubara ET, Mateus RG, de Araújo TLAC, da Silva LOC, de Queiroz EQA, de Abreu UGP, da Costa Gomes R, Pistori H. 2020b. Cattle weight estimation using active contour models and regression trees Bagging. Computers and Electronics in Agriculture, 179:105804. doi: 10.1016/j.compag.2020.105804

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Zhang ALN, Wu BP, Jiang CXH, Xuan DCZ, Ma EYH, Zhang FYA. 2018. Development and validation of a visual image analysis for monitoring the body size of sheep. Journal of Applied Animal Research, 46(1):1004-1015. doi: 10.1080/09712119.2018.1450257.