Kafkas Universitesi Veteriner Fakultesi Dergisi
ISSN: 1300-6045 e-ISSN: 1309-2251
Journal Home-Page: http://vetdergikafkas.org

Kafkas Univ Vet Fak Derg
28 (3): 411-419, 2022
DOI: 10.9775/kvfd.2022.27164

Research Article

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

Cem TIRINK [1,a (*)]

[1] Igdir University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics, TR-76000 Iğdır - TÜRKİYE
[a] ORCID: 0000-0001-6902-5837

Article ID: KVFD-2022-27164 Received: 24.01.2022 Accepted: 11.05.2022 Published Online: 18.05.2022

Abstract: In this study, it is aimed to compare several data mining and artificial neural network algorithms to predict body weight from biometric measurements for the Thalli sheep breed. For this purpose, the prediction capabilities of Bayesian Regularized Neural Network (BRNN), Support Vector Regression (SVR), Random Forest Regression (RFR) and Multivariate Adaptive Regression Splines (MARS) algorithms were comparatively investigated. To measure the predictive performances of the evaluated algorithms, body measurements such as body length, heart girth, ear length, ear width, head width, head length, withers height, rump length, rump width neck length, neck width of Thalli sheep were used for predicting the body weight. In this context, 270 female Thalli sheep were used to predict body weight. Model comparison criteria such as root-mean square error (RMSE), standard deviation ratio (SDR), performance index (PI), global relative approximation error (RAE), mean absolute percentage error (MAPE), Pearson's correlation coefficient (r), determination of coefficient (R^2) and Akaike's information criteria (AIC) were used to compare all algorithms. In conclusion, the MARS algorithm can be recommended to enable breeders to obtain an elite population of Thalli sheep breed.

Keywords: *Bayesian regularized neural network, Multivariate adaptive regression splines, Random forest regression, Support vector regression, Thalli sheep*

Thalli Koyunlarında Biyometrik Ölçümlerden Vücut Ağırlığı Tahmini İçin Bayesian Regularized Neural Network, Random Forest Regresyon, Support Vector Regresyon ve Çok Değişkenli Regresyon Uzanımları Algoritmalarının Karşılaştırılması

Öz: Bu çalışmada, Thalli koyun ırkı için biyometrik ölçümlerden vücut ağırlığını tahmin etmek için çeşitli veri madenciliği ve yapay sinir ağı algoritmalarının karşılaştırılması amaçlanmıştır. Bu amaçla BRNN, SVR, RFR ve MARS algoritmalarının tahmin performansları karşılaştırmalı olarak incelenmiştir. Değerlendirilen algoritmaların tahmin performanslarını ölçmek amacıyla vücut uzunluğu, göğüs çevresi, kulak uzunluğu, kulak genişliği, baş genişliği, baş uzunluğu, cidago yüksekliği, sağrı uzunluğu, sağrı genişliği boyun uzunluğu ve boyun genişliği gibi vücut ölçüleri canlı ağırlığını tahmin etmek için Thalli ırkı koyunlar kullanılmıştır. Bu kapsamda canlı ağırlık tahmini için 270 adet dişi Thalli koyunu kullanılmıştır. Tüm algoritmaların karşılaştırılmasında RMSE, SDR, PI, RAE, MAPE, r, R^2 ve AIC gibi model karşılaştırma kriterleri kullanılmıştır. Sonuç olarak, yetiştiricilerin elit bir Thalli koyun ırkı popülasyonu elde etmelerini sağlamak için MARS algoritması önerilebilir.

Anahtar sözcükler: *Bayesian regularized neural network, Çok değişkenli regresyon uzanımları, Random forest regresyon, Support vektör regresyon, Thalli koyunu*

Introduction

According to the FAO 2019-year database, there are 14.810.000 head meat sheep in Pakistan [1]. In total, Pakistan has 31 sheep breeds reared for meat, milk and wool products [2]. Among those, the Thalli breed is a thin-tailed sheep breed kept under tropical regions of Punjab province located in Pakistan, and medium size breed that has white body color, brown/black head with black spots on its legs. Sheep is an invaluable small ruminant breed that is used in many civilizations not only to obtain animal products such as meat, milk and fleece, but also to improve the rural economy [3]. Body weight is the major economical trait for all meat animals because income for farmers is directly gained by the weight of the animal. More sustained attention has been drawn to describe the relationship between body weight and linear biometric measurements (body measurements) for increasing meat production.

How to cite this article?

Tırınk C: 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. *Kafkas Univ Vet Fak Derg,* 28 (3): 411-419, 2022.
DOI: 10.9775/kvfd.2022.27164

(*) Corresponding Author

Tel: +90 541 694 9094
E-mail: cem.tirink@gmail.com

Body measurements of sheep may reflect production performance and genetic characteristics as well as growth and development in sheep [4]. Some body measurements taken at early growth periods were reported to be beneficial for early selection to improve superior offspring in body weight in next generation and these measurements provide helpful in practice for sheep breeders who are willing to predict body weight, which is essential in flock management.

In rural conditions where there is no weighing scale, estimation of live weight in sheep by using body measurements provides an important advantage in flock management by making it easier to know the optimum feed amount per sheep in the herd, marketing price, medical doses, optimum slaughtering time [5,6]. Estimation based on body measurements used as one of the positive effects on body weight is considered as an indirect selection criterion in animal breeding [3]. In this framework, the best way to ascertain the effective body measurements are to implement reliable statistical techniques such as multivariate analysis methods and data mining algorithms for performing phenotypically breed description of sheep.

Many studies have been reported about the prediction of body weight from body measurements in different animal species such as sheep [7-9], cattle [10,11], rabbit [12], dog [13] and camel [14]. In the literature, there are many practical approaches to estimate body weight for sheep breeds by body measurements within the scope of multiple regression [15], Classification and Regression Tree (CART) and Chi-square automatic interaction detection (CHAID) and MARS algorithm [6] and artificial neural networks [16]. However, the application of Bayesian Regularized Neural Network, Random Forest Regression and Support Vector Regression is still rare for prediction of body weight in sheep breeds. For Thalli sheep, there is a dearth of information about revealing predictive performances of various data mining and artificial neural network algorithms to model the causal relationship between body measurements and body weight in Thalli sheep. Thalli sheep related further studies are necessary for developing economic situation of smallholder farms under tropical regions of Pakistan. Thalli sheep are a breed originating from the Thall region, with a black mouth, Roman nose, black long ears, and white colored other parts of the body.

In literature, there is an absence of information about Thalli sheep raised in Pakistan such as the description of the inbreeding effect on before weaning period of growth features, identification of influential environmental factors on before weaning time characteristics [17] and quantitative genetic evaluation of after weaning period characteristics [18]. Nevertheless, information on the body weight prediction in Thalli sheep by body measurements emphasize insufficiently. The prediction is of great importance for making much better decisions on flock management, breed standards, breeding schemes and conserving gene reserves of the Thalli sheep. In this respect, sophistical statistical techniques can help to produce more reliable estimates within the scope of indirect selection criteria to be applied in sheep and to reveal body measurements that affect body weight. In recent years, use of these techniques such as Artificial Neural Networks (ANNs), Classification and Regression Tree (CART), Exhaustive Chi-square Automatic Interaction Detector (Exhaustive CHAID), Chi-square Automatic Interaction Detector (CHAID) and Multivariate Adaptive Regression Splines (MARS) has gained importance for the prediction of body weight in various sheep breeds [3,9,16]. The current study has been carried out both to fill this gap in the literature and to compare the prediction performances of these algorithms.

Material and Methods

The research was carried out at Livestock Experiment Station in Punjab, Pakistan. The animals were sent to graze on the pasture between 9:00 am and 3:00 pm. The animals were given water twice a day. They were vaccinated for enteroxemia, peste des petits ruminants and pox. Wheat bran was daily given 250 g/animal during the pregnancy period. The rams were kept separately, and the natural mating methods were done in August or September. As a material, 270 female Thalli sheep were used. The age range of Thalli sheep used in the study is between 1-2 years old. The data on body measurements were recorded based on some phenotypic traits in Thalli sheep. A digital weighing device is used to determine body weight and a flexible measuring tape (special tape) to record different body measurements. Body measurements such as body length, heart girth, ear length, ear width, head width, head length, withers height, rump length, rump width, neck length, neck width were used to predict of body weight.

Statistical Analysis

The data set can be assumed normally distributed by the Kolmogorov-Smirnov normality test (P>0.05). Descriptive statistics of all quantitative characteristics for Thalli sheep breed were reported as mean, standard deviation. The Thalli sheep data was partitioned two different data sets, 75% - 25% training and test sets, respectively. Additionally, Bayesian Regularized Neural Network, MARS, Random Forest Regression and Support Vector Regression algorithms were used to estimate body weight from body measurements in Thalli sheep.

Statistical evaluations were made using the R software [19]. To provide information about the structure of the data, descriptive statistics were performed. Descriptive statistics for all variables were estimated by using "psych" package in R environment [20]. In addition, correlation plot drawn by "corrplot" package in R software [21]. The "caret" packages in the R software were used to perform the analyzes of the BRNN and MARS algorithms [22]. The random forest regression was performed by using "randomForest" packages [23]. Also, support vector regression algorithm was performed by using "e1071" package in R software [24]. To display the performances of the constructed BRNN, MARS, Random Forest and Support Vector Regression models, the "ehaGoF" package was employed [25].

Bayesian Regularized Neural Network Algorithm (BRNN)

Artificial Neural Networks (ANNs) are known as mathematical models utilized in many scientific fields with the scope of solving prediction problems [26]. ANNs as one of the powerful artificial intelligence algorithms is structurally similar to the human brain and can be manipulated for sequential, nominal, scale-dependent variables. ANNs topologically consist of three layers such as input layer, hidden layer and output layer, respectively. The input layer consists of explanatory variables that the hidden layer depends on to start the process. The hidden layer consists of the activation functions and computes the weights of the explanatory variables in order to explore the effects of explanatory variables on the response variable [27]. Two types of ANNs algorithms such as radial basis functions neural networks (RBFNN) and bayesian regularized neural networks (BRNN) enable analysts to construct better models in predictive performance in comparison with linear models [26].

The BRNN function fits into a neural network that has input, hidden and output layers as described as given below [28,29].

$$y_i = g(x_i) + e_i = \sum_{k=1}^{s} w_k g_k \left(b_k + \sum_{j=1}^{p} x_{ij}\beta_k^{[k]}\right) + e_i, \ i = 1, \dots, n \quad (1)$$

where, s is number of neurons for hidden layers, b_k is the bias for the kth neuron for k=1, …, s, w_k is the weight of the kth neurons for k=1, …, s, $\beta_k^{[k]}$ is the weight of the jth input of the network, x_{ij} is the input of jth predictor in ith observation, e_i is the error term of the model, g_k is the activation function that equation as given below:

$$g_k(x) = \frac{\exp(2x)-1}{\exp(2x)+1} \quad (2)$$

It uses Nguyen and Widrow algorithm to assign starting weights while performing this function and Gauss-Newton algorithm to provide optimization [30]. The Nguyen and Widrow initialization algorithm generates the initial weight and bias values to evenly distribute the active regions of neurons over the input area [31,32].

Support Vector Regression (SVR)

An important branch of the support vector machine, which is one of the machine learning algorithms, is the support vector regression (SVR) algorithm [33]. While the statistical method used in classification is called support vector classification (SVC), the statistical method used with modeling and prediction is called SVR [34-36]. Since SVR is a supervised learning method, the performance of SVR varies depending on the training and test dataset [37].

In linear SVR model, the main goal of SVR is to define a function f(x) that can have the maximum deviation (ε) from the training set and should be as straight as possible. Training data points are placed within the limit between −ε to +ε [37]. However, most of these studies cannot be modeled within the scope of linearity. Therefore, in the case of nonlinear SVR, the input data is matched to a higher dimensional Hilbert space $(\mathcal{H})$ so that the regression line can be linear [33].

The nonlinear regression hyperplane to be obtained is as follows.

$$\hat{y} = \langle w, \phi(x) \rangle + b \quad (3)$$

In this equation, w is a weight vector, $\phi(x)$ is non-linear kernel functions, $\langle .,. \rangle$ indicates vector inner product and is a bias term. There are many nonlinear kernel functions and one of these kernel functions is gaussian radial basis function kernel. The kernel function used in this study is the gausian radial basis function.

Random Forest Regression (RFR)

Random Forests is a popular method among multivariate statistical methods because of its easy applicability in classification and regression type problems. The Random Forest algorithm, which adds a layer of randomness to the bagging algorithm, was proposed by Breiman [38]. The RFR algorithm is a learning algorithm by combines sets of regression trees. A regression tree is represented as a set of constraints that are applied hierarchically from root to leaf of the tree [39,40]. The biggest advantage of this algorithm is that it can be easily used in the case of nonlinearity.

The algorithm requires a process that includes three stages [23]. The first procedure is to build the number of trees (n_{tree}) bootstrap samples from original data. The second procedure is to develop an un-pruned classification or regression tree for each sample. The last procedure is to estimate the new data from the tree. For the Thalli sheep data set, model parameters such as n_{tree} and the number of variables tried at each split are selected (m_{try}) 500 and 3, respectively.

Multivariate Adaptive Regression Splines Algorithm (MARS)

One of the tree-based algorithms are used to solve regression-type problems while estimating based on quantitative traits [41-43]. To solve classification problems, The Multivariate Adaptive Regression (MARS) algorithm, which is a non-parametric regression technique that enables more effective identification of nonlinear and interaction effects between response and explanatory variables, was proposed by Friedman [44] and derived from the CART algorithm. In the algorithm, there is no need for any assumptions about both the distribution of variables and the relationships between variables [45,46]. The algorithm is a non-parametric regression technique in which various slopes in the training data set are divided into individual segmented linear segments (splines) [45]. Splines connect seamlessly and form connection points called "knot". Candidate nodes are randomly placed within the range of each estimator, so the model estimation to be made with the MARS algorithm is more flexible and interpretable with the help of piecewise linear regressions [45].

The algorithm consists of two different stages, a forward and backward pass stage [47]. The first stage for the algorithm is the forward pass stage. At this stage, the algorithm starts with an intercept in the first model and to improve the model recursively includes the basic function pairs with the least training error. The forward pass stage characteristically produces an over-fitted pattern that reaches maximum complexity [44]. The model constructed from the forward pass stage fits particularly good. Nevertheless, its generalization ability can be weak for a data set before an undetermined constructed model which means an overfitting problem. The basic functions that provide the least amount to the prediction model are eliminated in the backward pass stage and this situation is used in the solution of this problem [6,47].

At the beginning of the analysis, the multicollinearity between the explanatory variables was checked and it was found that there was no multicollinearity between the variables. To estimate BW using the training data set, the cross-validation method was used to decide the best MARS model among 324 MARS model with degree = 1:10 and nprune = 2:38 in determining the number of terms to be selected in the model. Ten-fold cross-validation was used for MARS model in the training data set.

RMSE, rRMSE, SDR, PI, RAE, MAPE, r, R^2 and Adj-R^2 criteria were used to compare the performance of the model. To compare the model performances were made according to the lowest RMSE, rRMSE, SDR, PI, RAE, MAPE values and the highest r, R^2 and Adj-R^2 value [48].

Results

Mean and standard deviation as the descriptive statistics of each trait in the present study for Thalli sheep are given in *Table 1.*

Fig. 1 showed that Pearson's correlation coefficient to determine the relationship between body measurements and BW. All correlation coefficients were determined to be statistically significant (P<0.01).

To compare the algorithms, some model comparison criteria were used for determining the performances of the algorithms in *Table 2*. Using metrics, the performances of all models at each stage were evaluated with test data from the respective stages and compared to determine the best model.

The metric methods used were evaluated for both the train and the test set, and the methods were compared to find the best method. According to *Table 2,* the performance in the test for each model was determined to be weaker than

Table 1. *Descriptive statistics for each measurement*

Variables	Mean	Standard Deviation
Withers height	63.92	7.66
Body length	63.71	9.08
Head length	24.63	4.31
Head width	9.53	1.64
Ear length	26.5	2.82
Ear width	11.69	1.31
Neck length	24.26	3.54
Neck width	15.62	2.25
Heart girth	68.15	9.58
Rump length	12.93	2.54
Rump width	18.16	4.36
Body weight	23.77	6.89

	withers_height	body_length	head_length	head_width	ear_length	ear_width	neck_length	neck_width	heart_girth	rump_length	rump_width	body_weight
withers_height	1.00											
body_length	0.87	1.00										
head_length	0.75	0.78	1.00									
head_width	0.65	0.61	0.52	1.00								
ear_length	0.62	0.57	0.48	0.45	1.00							
ear_width	0.62	0.54	0.41	0.55	0.75	1.00						
neck_length	0.72	0.68	0.64	0.61	0.48	0.48	1.00					
neck_width	0.72	0.77	0.65	0.55	0.43	0.45	0.65	1.00				
heart_girth	0.88	0.83	0.74	0.69	0.62	0.60	0.74	0.70	1.00			
rump_length	0.71	0.71	0.62	0.59	0.51	0.42	0.54	0.53	0.74	1.00		
rump_width	0.48	0.57	0.69	0.06	0.26	0.14	0.30	0.49	0.44	0.43	1.00	
body_weight	0.75	0.78	0.80	0.57	0.50	0.46	0.62	0.60	0.76	0.65	0.54	1.00

Fig 1. Correlation matrix

Table 2. Model comparison criteria for BRNN, SVR, RFR and MARS algorithms

Model Comparison Criteria	BRNN		SVR		RFR		MARS	
	Train	Test	Train	Test	Train	Test	Train	Test
Root mean square error (RMSE)	11.093	10.41	7.272	11.21	11.797	11.684	9.375	10.076
Standard deviation ratio (SDR)	0.477	0.476	0.386	0.496	0.492	0.507	0.439	0.460
Pearson's correlation coefficients (PC)	0.879	0.880	0.922	0.868	0.871	0.863	0.899	0.888
Performance index (PI)	7.467	7.197	5.909	7.514	7.733	7.694	6.793	7.05
Relative approximation error (RAE)	0.018	0.017	0.012	0.018	0.019	0.019	0.015	0.016
Mean absolute percentage error (MAPE)	12.063	13.177	7.764	12.537	11.615	12.884	10.39	12.086
Coefficient of determination (R^2)	0.772	0.758	0.851	0.74	0.758	0.729	0.808	0.766
Akaike's information Criterion (AIC)	490.881	154.620	404.737	159.511	503.435	162.245	474.571	170.467

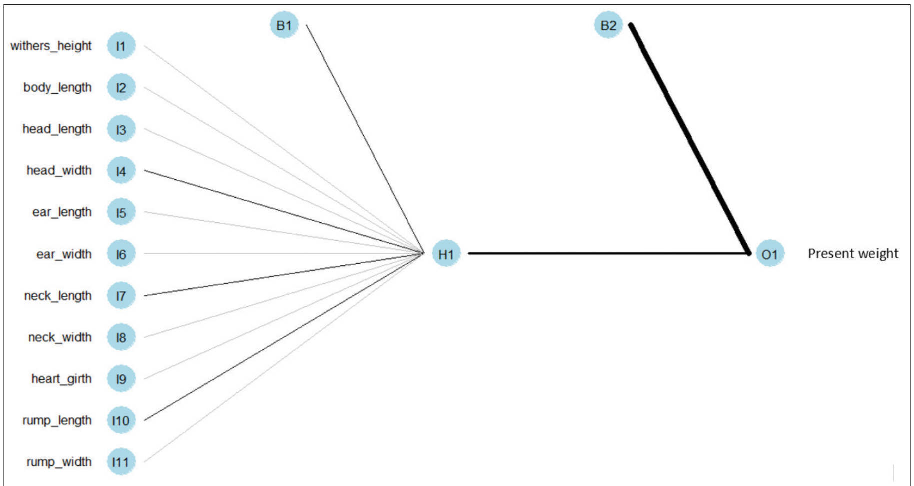

Fig 2. BRNN architecture to predict BW

the training dataset. When all metrics were evaluated, it turned out that the algorithm with the best fit for both the train and the test set was the MARS algorithm.

The Bayesian regularized neural network (BRNN) models were trained using the training data sets to predict the body weight. BRNN network algorithm produced a topologically 11-1-1 structure (the number of neurons in the input, hidden and output layers, respectively *(Fig. 2)* for the body weight prediction. In addition, the optimum BRNN model was achieved in the 12th epoch of the training phase. The network with one neuron gave the lowest RMSE value *(Fig. 3)*. The goodness of fit criteria revealed that BRNN algorithm produced the best fit among the candidate algorithms *(Table 2)*. Sensitivity analysis was carried out to estimate the relative importance values of the explanatory variables on body weight *(Fig. 4)*.

At the beginning of the SVR algorithm was trained training data. After the training process the SVR were examined to estimate the body weight for Thalli sheep breed. Gaussian radial basis kernel function for SVR estimation of body weight. The reliability of the model depends on the selection of parameters such as cost (C) and epsilon. These parameters were tested for various values and analysis was applied for C and epsilon values, which would give the most reliable model. Sensitivity analysis was carried out to estimate relative importance values of the explanatory variables on body weight for SVR *(Fig. 5)*.

For RFR, the model performance metrics were given in *Table 2*. In addition, sensitivity analysis was carried out to estimate relative importance values of the explanatory variables on body weight for RFR *(Fig. 6)*.

The optimum MARS model with 9 terms and degree: 1 is selected as the optimum model with the lowest cross-validated RMSE value among 324 candidate MARS models. The optimum MARS model obtained is as bellow in *Table 3*.

Discussion

Methods based on body measurements are widely used in determining the relationship between BW and the structure of the animal species. However, the validity of the statistical method used to estimate BW from these body measurements is also important. In this context, many studies have been carried out for different animal species. In multivariate statistics within the scope of data mining and artificial neural networks, the use of model

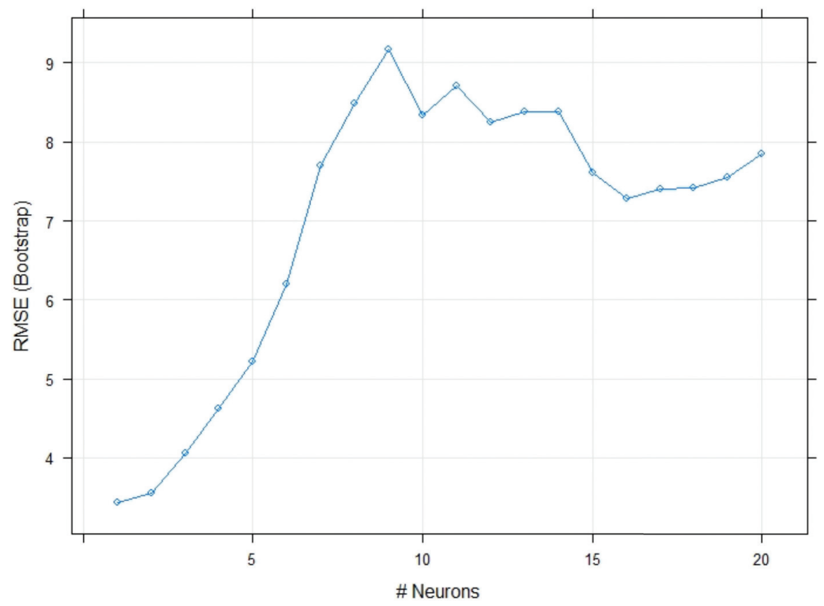

Fig 3. RMSE diagram for the bootsrap BRNN algorithm

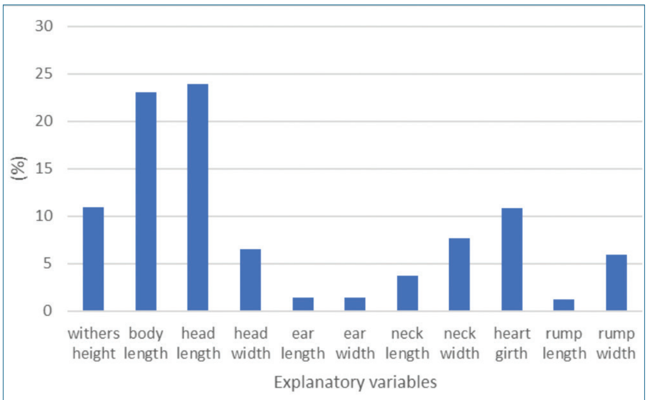

Fig 4. Sensitivity analysis for BRNN model

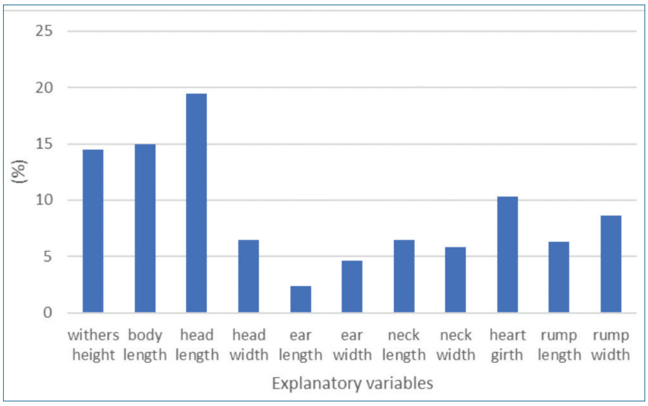

Fig 6. Sensitivity analysis for RFR model

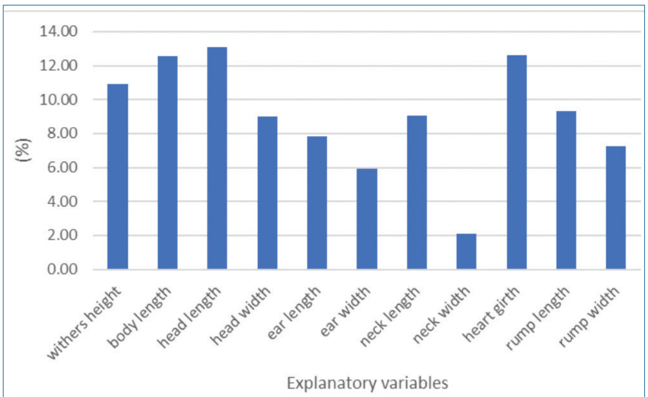

Fig 5. Sensitivity analysis for SVR model

***Table 3.** Obtained optimum MARS model*

Explanatory Variables	Coefficients
Intercept	10.8827049
(body_length-48.26)	0.2558336
(19.05-head_length)	1.5792126
(head_length-19.05)	1.6836423
(head_length-22.86)	-1.2350427
(7.62-head_width)	-2.0680844
(neck_length-27.94)	0.6855337
(12.7-neck_width)	2.0226105
(heart_girth-53.34)	0.1611244

comparison criteria has been proposed in the selection of the best model [12]. In this context, model performances are compared according to the lowest RMSE, rRMSE, SDR, PI, RAE, MAPE values and the highest r, R^2 and Adj-R^2 values [48].

With the use of these algorithms, by determining the selection scheme in Thalli sheep, the variables that are effective in BW estimation can be determined, and this will lead to sustainable livestock breeding. However, there are deficiencies in the literature on studies using these algorithms. In this context, it has been determined that only the MARS algorithm and some of the ANN algorithms are used in the literature.

Ali et al.[49] compared the ANN, CART, Exhaustive CHAID and CHAID algorithms in their study on the Harnai sheep breed. When the study was evaluated within the scope of R^2, Exhaustive CHAID was estimated 0.8421, CHAID as 0.8377, CART as 0.82644 and ANN as 0.81999. The proposed method was the Exhaustive CHAID method to define the standards of Harnai sheep breed. Our obtained R^2 values were found lower than the results in comparison of ANN algorithm which is comparable for both studies. Breed and evaluated traits may be the factors of these differences.

Eyduran et al.[50] used CART, CHAID and ANN (RBF, MLP1 and MLP2) algorithms for BW prediction for Beetal goat and this study, which was evaluated within the scope of the correlations found between predicted and actual values, the highest correlation was determined for RBF of the ANN algorithm. In this context, the current study is thought to have different results due to differences between species.

Celik et al.[3] aimed to compare CART, CHAID, Exhaustive CHAID, MARS, MLP, and RBF on Mengali rams. Within the scope of model comparison criteria R^2, SDratio and RMSE the best prediction model was determined as the CART algorithm. However, the MARS algorithm appears to have an R^2 value of 0.88. In the current study, it has been seen that the R^2 value for the train set is similar.

Compared to the results of previous studies, the breeds used in the studies, the age of the animals, the differences in flock management systems and the statistical methods used can be attributed to this wide variation. In this context, compared to results from other studies, it was determined that provide similar results according to the chosen model of evaluation criteria. However, different methods have been proposed in terms of the methods used. Suggesting different statistical methods for BW estimation using body measurements reveals that there is a need for more studies on this subject.

In the literature, variability in BW estimation in sheep may be due to differences in the number of samples used, breed, sex, flock management systems and statistical methods. For this purpose, the correct use of the factors that cause these differences is very important for a sustainable livestock breeding and selection. The correct use of statistical methods, which is one of the factors that cause variability, will create the make a right decision mechanism with more reliable estimates.

The results show that the BW estimation to be made with the MARS algorithm is more reliable within the scope of model comparison criteria. For the MARS algorithm, body length, head length, head width, neck length and neck width measurements were determined to be the most effective BW estimation.

There is no available information to estimate body weight from body measurements within the scope of BRNN, SVR, RFR and MARS algorithm for Thalli sheep. In the present study, body weight was estimated from body measurements by using these algorithms for Thalli sheep breed. Although all algorithms have their own advantages and disadvantages, the model performances obtained from the MARS algorithm were better determined.

In conclusion, the MARS algorithm can be recommended to enable breeders to obtain an elite population of Thalli sheep breed. In addition, it provides to increasing BW as a selection criterion for determining appropriate body measurements and flock management standards. The results of this study, based on model selection criteria for the selection of the most suitable model, showed that data mining and artificial neural network algorithms can be successfully applied to BW estimation based on measured body measurements. Even if there are some differences in the value of the comparison criteria, more reliable models can be obtained by conducting similar studies.

Availability of Data and Materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

The author would like to thank Dr. Abdul Waheed for help in obtaining data.

Financial Support

No financial support was received for conducting this study.

Conflict of Interest

The author declared that there is no conflict of interest.

References

1. FAO: FAOSTAT, 2019. http://www.fao.org/faostat/en/#data/QL; *Accessed*: 11.03.2021.

2. Tariq MM, Bajwa MA, Jawasreh K, Awan MA, Abbas F, Waheed A, Rafeeq M, Wadood A, Khan KU, Rashid N, Atique MA, Bukhari FA: Characterization of four indigenous sheep breeds of Balochistan, Pakistan by random amplified polymorphic DNAs. *Afr J Biotechnol*, 11, 2581-2586, 2012. DOI: 10.5897/AJB11.3196

3. Celik S, Eyduran E, Karadas K, Tariq MM: Comparison of predictive performance of data mining algorithms in predicting body weight in Mengali rams of Pakistan. *Braz J Vet Res Anim Sci*, 46 (11): 863-872, 2017. DOI: 10.1590/S1806-92902017001100005

4. Zhang AL, Wu BP, Wuyun CT, Jiang DX, Xuan EC, Ma FY: Algorithm of sheep body dimension measurement and its applications based on image analysis. *Comput Electron Agric,* 153, 33-45, 2018. DOI: 10.1016/j.compag.2018.07.033

5. Sabbioni A, Beretti V, Superchi P, Ablondi M: Body weight estimation from body measures in Cornigliese sheep breed. *Ital J Anim Sci*, 19 (1): 25-30, 2020. DOI: 10.1080/1828051X.2019.1689189

6. Faraz A, Tirink C, Eyduran E, Waheed A, Tauqir NA, Nabeel MS, Tariq MM: Prediction of live body weight based on body measurements in Thalli sheep under tropical conditions of Pakistan using CART and MARS. *Trop Anim Health Prod*, 53 (2): 1-12, 2021. DOI: 10.1007/s11250-021-02748-6

7. Cam MA, Olfaz M, Soydan E: Body measurements reflect body weights and carcass yields in Karayaka sheep. *Asian J Anim Vet Adv*, 5 (2): 120-127, 2010. DOI: 10.3923/ajava.2010.120.127

8. Ghotbaldini H, Mohammadabadi M, Nezamabadi-Pour H, Babenko OI, Bushtruk MV, Tkachenko SV: Predicting breeding value of body weight at 6-month age using Artificial Neural Networks in Kermani sheep breed. *Acta Sci Anim Sci*, 41, 2019. DOI: 10.4025/actascianimsci.v41i1.45282

9. Olfaz M, Tirink C, Onder H: Use of CART and CHAID algorithms in Karayaka sheep breeding. *Kafkas Univ Vet Fak Derg*, 25 (1): 105-110, 2019. DOI: 10.9775/kvfd.2018.20388

10. Aytekin I, Eyduran E, Karadas K, Aksahan R, Keskin I: 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. *Pak J Zool*, 50 (1): 189-195, 2018. DOI: 10.17582/journal.pjz/2018.50.1.189.195

11. Weber VAM, de Lima Weber F, da Silva Oliveira A, Astolfi G, Menezes GV, de Andrade Porto JV, Rezende FPC, de Moraes PH, Matsubara ET, Mateus RG, de Araújo TLAC, da Silva LOC, de Queiroz EQA, de Abreu UGP, da Costa Gomes R, Pistori H: Cattle weight estimation using active contour models and regression trees Bagging. *Comput Electron Agric*, 179, 105804, 2020. DOI: 10.1016/j.compag.2020.105804

12. Salawu EO, Abdulraheem M, Shoyombo A, Adepeju A, Davies S, Akinsola O, Nwagu B: Using artificial neural network to predict body weights of rabbits. *Open J Anim Sci*, 4, 182-186, 2014. DOI: 10.4236/ojas.2014.44023

13. Celik S, Yilmaz O: Prediction of body weight of Turkish tazi dogs using data mining Techniques: Classification and Regression Tree (CART) and multivariate adaptive regression splines (MARS), *Pak J Zool*, 50 (2): 575-583, 2018. DOI: 10.17582/journal.pjz/2018.50.2.575.583

14. Fatih A, Celik S, Eyduran E, Tirink C, Tariq MM, Sheikh IS, Faraz A, Waheed A: Use of MARS algorithm for predicting mature weight of different camel *(Camelus dromedarius)* breeds reared in Pakistan and morphological characterization via cluster analysis. *Trop Anim Health Prod*, 53 (1): 1-14, 2021. DOI: 10.1007/s11250-021-02633-2

15. Topal M, Macit M: Prediction of body weight from body measurements in morkaraman sheep. *J Appl Anim Res*, 25 (2): 97-100, 2004. DOI: 10.1080/09712119.2004.9706484

16. Huma ZE, Iqbal F: Predicting the body weight of Balochi sheep using a machine learning approach. *Turk J Vet Anim Sci*, 43 (4): 500-506, 2019. DOI: 10.3906/vet-1812-23

17. Hussain A, Akhtar P, Ali S, Younas M, Yaqoob M, Babar ME, Kaved K, Shakoor A: Factors influencing body weights at different ages in Thalli sheep. *J Anim Plant Sci*, 23 (1): 1-6, 2013.

18. Hussain A, Akhtar P, Ali S, Javed K, Younas M, Shakoor A, Waheed U: Genetic analysis of post-weaning growth traits of Thalli sheep under tropical conditions. *Trop Anim Health Prod*, 46, 1527-1531, 2014. DOI: 10.1007/s11250-014-0670-x

19. R Core Team: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2020. https://www.R-project.org/; *Accessed:* 11.03.2021.

20. Revelle W: Psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, 2020. https://CRAN.R-project.org/package=psych Version = 2.0.12; *Accessed:* 11.03.2021.

21. Wei T, Simko V: R package "corrplot": Visualization of a Correlation Matrix (Version 0.84), 2017. https://github.com/taiyun/corrplot; *Accessed:* 11.03.2021.

22. Kuhn M: Caret: Classification and Regression Training. R package version 6.0-86, 2020. https://CRAN.R-project.org/package=caret; *Accessed:* 11.03.2021.

23. Liaw A, Wiener M: Classification and regression by random forest. *R News*, 2 (3): 18-22, 2002.

24. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F: e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien R package version 1.7-4, 2020. https://CRAN.R-project.org/package=e1071; *Accessed:* 11.03.2021.

25. Eyduran E: ehaGoF: Calculates Goodness of Fit Statistics. R package version 0.1.1., 2020. https://CRAN.R-project.org/package=ehaGoF; *Accessed:* 11.03.2021.

26. Pérez-Rodríguez P, Gianola D, Weigel KA, Rosa GJM, Crossa J: Technical Note: An R package for fitting Bayesian regularized neural networks with applications in animal breeding. *J Anim Sci*, 91 (8): 3522-3531, 2013. DOI: 10.2527/jas.2012-6162

27. Kayri M: Predictive abilities of bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: A comparative empirical study on social data. *Math Comput Appl*, 21(2): 20, 2016. DOI: 10.3390/mca21020020

28. MacKay DJC: A practical Bayesian framework for backpropagation networks. *Neural Comput*, 4, 448-472, 1992. DOI: 10.1162/neco.1992.4.3.448

29. Foresee FD, Hagan MT: Gauss-Newton approximation to Bayesian learning. *Proceedings of International Conference on Neural Networks*, 3, 1930-1935, 1997. DOI: 10.1109/ICNN.1997.614194

30. Nguyen D, Widrow B: Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *1990 IJCNN International Joint Conference on Neural Network*, 3, 21-26, 1990. DOI: 10.1109/IJCNN.1990.137819

31. Dolezel P, Skrabane P, Gago L: Weight initialization possibilities for feedforward neural network with linear saturated activation functions. *IFAC-PapersOnLine*, 49 (25): 49-54, 2016. DOI: 10.1016/j.ifacol.2016.12.009

32. Linan MN, Gerardo B, Medina R: Modified weight initialization in the self-organizing map using Nguyen-Widrow initialization algorithm. *J Phys Conf Ser*, 1235:012055, 2019. DOI: 10.1088/1742-6596/1235/1/012055

33. Nguyen QT, Fouchereau R, Frénod E, Gerard C, Sincholle V: Comparison of forecast models of production of dairy cows combining animal and diet parameters. *Comput Electron Agric*, 170, 105258, 2020. DOI: 10.1016/j.compag.2020.105258

34. Smola AJ, Schölkopf B: A tutorial on support vector regression. *Stat Comput*, 14, 199-222, 2004. DOI: 10.1023/B:STCO.0000035301.49549.88

35. Kavaklioglu K: Modeling and prediction of Turkey's electricity consumption using Support Vector Regression. *Appl Energy*, 88, 368-375, 2011. DOI: 10.1016/j.apenergy.2010.07.021

36. Alonso J, Castañón ÁR, Bahamonde A: Support Vector Regression to predict carcass weight in beef cattle in advance of the slaughter. *Comput Electron Agric*, 91, 116-120, 2013. DOI: 10.1016/j.compag.2012.08.009

37. Patel AK, Chatterjee S, Gorai AK: Development of a machine vision system using the support vector machine regression (SVR) algorithm for the online prediction of iron ore grades. 2017 Fifteenth IAPR International Conference on Machine Vision Applications 12, 197-210, 2017. DOI: 10.23919/MVA.2017.7986823

38. Breiman L: Random forests. *Mach Learn*, 45 (1): 5-32, 2001. DOI: 10.1023/A:1010933404324

39. Rodriguez-Galiano V, Mendes MP, Garcia-Soldado MJ, Chica-Olmo M, Riberio L: Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: A case study in an agricultural setting (Southern Spain). *Sci Total Environ*, 476-477, 189-206, 2014. DOI: 10.1016/j.scitotenv.2014.01.001

40. Wang L, Zhou X, Zhu X, Dong Z, Guo W: Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop J*, 4, 212-219, 2016. DOI: 10.1016/j.cj.2016.01.008

41. Akin M, Eyduran E, Niedz RP, Reed BM: Developing hazelnut tissue culture free of ion confounding. *Plant Cell Tissue Organ Cult*, 130 (3): 483-494, 2017. DOI: 10.1007/s11240-017-1238-z

42. Akin M, Eyduran E, Reed BM: Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. *Plant Cell Tissue Organ Cult*, 128 (2): 303-316, 2017. DOI: 10.1007/s11240-016-1110-6

43. Kovalchuk IY, Mukhitdinova Z, Turdiyev T, Madiyeva G, Akin M, Eyduran E, Reed BM: Modeling some mineral nutrient requirements for micropropagated wild apricot shoot cultures. *Plant Cell Tissue Organ Cult*, 129, 325-335, 2017. DOI: 10.1007/s11240-017-1180-0

44. Friedman JH: Multivariate adaptive regression splines. *Ann Stat*, 19 (1): 1-67, 1991. DOI: 10.1214/aos/1176347963

45. Aksoy A, Erturk YE, Eyduran E, Tariq MM: Utility of MARS algorithm for describing non-genetic factors affecting pasture revenue of Morkaraman breed and Romanov x Morkaraman F1 crossbred sheep under semi intensive conditions. *Pak J Zool*, 51(1): 235-240, 2019. DOI: 10.17582/journal.pjz/2019.51.1.235.240

46. Akin M, Eyduran SP, Eyduran E, Reed BM: Analysis of macro nutrient related growth responses using multivariate adaptive regression splines. *Plant Cell Tissue Organ Cult*, 140, 661-670, 2020. DOI: 10.1007/s11240-019-01763-8

47. Arthur CK, Temeng VA, Ziggah YY: Multivariate Adaptive Regression Splines (MARS) approach to blast-induced ground vibration prediction. *Int J Min Reclam Environ*, 34 (3): 198-222, 2020. DOI: 10.1080/17480930.2019.1577940

48. Tatliyer A: The Effects of raising type on performances of some data mining algorithms in lambs. *KSU J Agric Nat*, 23(3), 772-780, 2020. DOI: 10.18016/ksutarimdoga.vi.651232

49. Ali M, Eyduran E, Tariq MM, Tirink C, Abbas F, Bajwa MA, Baloch MH, Nizamani AH, Waheed A, Awan MA, Shah SH, Ahmad Z, Jan S: 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. *Pak J Zool*, 47, 1579-1585, 2015.

50. Eyduran E, Zaborski D, Waheed A, Celik S, Karadas K, Grzesiak W: 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. *Pak J Zool*, 49 (1): 257-265, 2017.