Proceedings of the Pakistan Academy of Sciences:
A. Physical and Computational Sciences: 58(2): 29-38 (2021)

ISSN (Print): 2518-4245; ISSN (Online): 2518-4253
https://doi.org/10.53560/PPASA(58-2)603

Pakistan Academy of Sciences

Research Article

Body Weight Prediction of Thalli Sheep Reared in Southern Punjab Using Different Data Mining Algorithms

Ansar Abbas[1*], Muhammad Aman Ullah[2], and Abdul Waheed[3]

[1]Govt. Degree College for Boys Makhdoom Rashid, Multan, Pakistan
[2]Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan
[3]Department of Livestock and Poultry Production, Bahauddin Zakariya University, Multan, Pakistan

Abstract: This study is conducted to predict the body weight (BW) for Thalli sheep of southern Punjab from different body measurements. In the BW prediction, several body measurements viz., withers height, body length, head length, head width, ear length, ear width, neck length, neck width, heart girth, rump length, rump width, tail length, barrel depth and sacral pelvic width are used as predictors. The data mining algorithms such as Chi-square Automatic Interaction Detector (CHAID), Exhaustive CHAID, Classification and Regression Tree (CART) and Artificial Neural Network (ANN) are used to predict the BW for a total of 85 female Thalli sheep. The data set is partitioned into training (80 %) and test (20 %) sets before the algorithms are used. The minimum number of parent (4) and child nodes (2) are set in order to ensure their predictive ability. The R^2 % and RMSE values for CHAID, Exhaustive CHAID, ANN and CART algorithms are 67.38(1.003), 64.37(1.049), 61.45(1.093) and 59.02(1.125), respectively. The most significant predictor is BL in the BW prediction of Thalli sheep. The heaviest BW average of 9.596 kg is obtained from the subgroup of those having BL > 25.000 inches. On behalf of the several goodness of fit criteria, we conclude that the CHAID algorithm performance is better in order to predict the BW of Thalli sheep and more suitable decision tree diagram visually. Also, the obtained CHAID results may help to determine body measurements positively associated with BW for developing better selection strategies with the scope of indirect selection criteria.

Keywords: ANN, Body Measurements, Body Weight, CART, CHAID, Exhaustive CHAID.

1. INTRODUCTION

Sheep are generally reared as a small ruminant utilized in rural development in almost all societies with various uses [1]. The ultimate aim of sheep production is to obtain the enhanced productivity of the yielding traits such as; milk and meat produced from the animals. The actual BW of a sheep is one of the determining factors in finding the proper medicinal dose, feed amount and price of sheep [2]. Sheep breeders generally presume the live weight of a sheep by touching or visually. The sale and purchase of animals are usually made by bargaining or on the source of their physical look. However, in this way of selling the farmers usually did not get the actual price of their animals and the big part of the profit is manipulated by the middleman [3]. They also estimate the live weight by using the weighing machine. However, it is really difficult to measure BW in village surroundings where measurement scale (i.e., animal weighing machine) and regular records of sheep are not available [4]. Under these circumstances, biometrical researchers (breeders) often used different morphological measurements in the prediction of live BW of sheep and to increase wealthy animal breeding systems [5].

In the literature, different morphological characteristics (i.e., withers height, body length, chest girth, barrel depth or heart girth etc.) have been utilized to predict BW [6-8]. The various prediction equation of live weight has been estimated by applying different types of regression (one variable, two or more than two variables or ridge regression) [9-12]. In these traditional methods, if there is a very strong relationship greater than 0.80-0.90 among

Received: January 2021; Accepted: June 2021
*Corresponding Author: Ansar Abbas <ansarashri@gmail.com>

predictors, the biased parameter estimation is available because of the multicollinearity problem for each morphological trait and it is very difficult to correctly infer the influence of morphological trait on BW, see further details in [13, 14].

The limitations of regression equations have led researchers to look at alternative models. In recent years, different researchers planned to use different data mining algorithms i.e., CHAID [14-21], Exhaustive CHAID [17, 22-24] and CART [14, 20, 25-28], are reported to be superior compared with multiple linear regression, cluster analysis, logistic regression and discriminant analysis. Moreover, these algorithms obtaining homogenous subgroups in a little while are not affected by the problem of multicollinearity, missing data and outliers [29]. Also, the supremacy of multivariate adaptive regression splines (MARS) is reported for the prediction of BW in sheep [21, 30-32], goat [28], camel [33] and cattle [34]. These algorithms are non-parametric methods that are commonly used for nominal, ordinal and scale variables [14, 22].

In Pakistan, numerous studies on BW prediction in sheep and goat are available e.g., Mohammad et al. [4] predict the BW of Balochi sheep based on different body measurements using the regression tree method. In 2015, Ali et al. [14] predict the BW from body length, withers height, chest girth, paunch girth, face length, length between ears, length of ears, width and length of tail of 6-9 months Harnai sheep. In their study, they used the CART, CHAID, Exhaustive CHAID and ANN algorithms. Another study by Eyduran et al. [22] also predict the BW of Pakistani Beetal goat based on six different predictors i.e., head girth, neck length, diagonal body length, belly sprung, shank circumference and rump height. They also used the CART, CHAID and ANN algorithms for BW prediction and the results were compared with the multiple linear regression. However, no researchers predict the BW for Thalli sheep of south Punjab, Pakistan. This research gap motivated us to conduct a study for BW prediction of Thali sheep of south Punjab using different data mining algorithms. The main purpose of our study is to predict BW for Thali sheep of southern Punjab from different body measurements. We also compared the performance of different data mining algorithms viz. CART, CHAID, Exhaustive CHAID and ANN under study in order to predict BW.

2. MATERIALS AND METHODS

The current animal data were provided from Thalli sheep locally found in southern Punjab, Pakistan. The same breed was also available with livestock farmers as well as government livestock stations. In the present study, 85 female sheep at varying ages ranging from 1 to 12 months were included. Random sampling was used for sheep selection. All healthy sheep who didn't take any medication, not any physical disability were included in the study.

Data collection activity was made by the same person in order to avoid the between individual variation. The required information was taken through a self-administered questionnaire, comprised of age and morphologically measurements. All the morphologically measurements i.e., withers height (WH), body length (BL), head length (HL), head width (HW), ear length (EL), ear width (EW), neck length (NL), neck width (NW), heart girth (HG), rump length (RL), rump width (RW), tail length (TL), barrel depth (BD) and sacral-pelvic-width (SPW) were recorded according to standard procedures. Measurements were made in inches and taken in standing position of sheep. Initially, descriptive analysis i.e., means, standard deviation (SD) and percentage co-efficient of variation (C.V %) of each quantitative variable were reported in Table 1.

The BW of sheep was predicted using four different methods i.e., CHAID, Exhaustive CHAID, CART and ANN algorithms. A CHAID algorithm based on chi-square test of association and is used to classify those subsets of predictors that best describe the dependent variable. A CHAID algorithm forms a decision tree structure by recursively dividing a subset into many homogenous subsets (nodes) consisting alike responses of predictand variable as soon as possible starting from the root node [35-36]. The term regression tree is used for the tree that its dependent variable is scale [37]. In CHAID algorithm, the dependent variable can be continuous and categorical but, the independent variables are categorical variables only and can have more than two categories. CHAID can create multiple splits [24, 38]. The basic aim of CHAID algorithm is to reduce variance within nodes in the dependent variable during constructing regression tree diagram. In the advancement of CHAID algorithm, Exhaustive CHAID is based on three-stage-data mining algorithm (i.e., merging, partitioning, and stopping) algorithms that recursively use multi-way splitting to form homogenous subsets on the basis of Bonferroni adjustment until the differences between the observed and the estimated values in response variable are minimal [15, 22, 39]. Exhaustive CHAID has the same splitting and stopping rules like CHAID; however, the merging step is more exhaustive than CHAID, by continuing to merge categories of the predictor variable until only two super categories are left. The Exhaustive CHAID can find the best split for each predictor variable [40].

Breiman et al. [41] developed CART algorithm. It is a recursive splitting method and is used both for regression and classification problems. In CART algorithm, the dependent variable is scale and independent variables can be scale or categorical. CART algorithm creates a binary split [24, 38]. In CART, the best input variable is chosen by using a range of diversity procedures [1]. CART algorithm creates more homogenous sub-groups than CHAID algorithm using pruning [42]. By default, the maximum number of levels (tree depths) is 5 for CART and 3 for CHAID algorithm. A 10-fold cross validation criteria was applied and the minimum number of cases for parent and child node was set at 4:2 in order to correctly model the relationship between response and independent variables and also to get best possible decision tree structure.

Artificial neural network (ANN) biologically resembles the human brain. It consists of three layers i.e., input, hidden and output layers and are used with one hidden layer on the source of MLP which is also called a feed forward neural network to predict BW from body measurements [14, 43]. The data were at random dividing into a training set (80 %) and test set (20 %). All of the above stated methods- CHAID, Exhaustive CHAID, CART data mining algorithms and ANN are available in statistical software “Statistical Package for Social Sciences (SPSS)” version 23.0 which were utilized for predicting BW of sheep on the basis of different morphological characteristics. Moreover, Bonferroni adjustment was performed for both CHAID algorithms in order to calculate adjusted p values of F values [44].

2.1. Model Quality Criteria

In our study, the model selection criteria was based on R^2 (%), Adj. R^2 (%), r, CV (%), SD ratio, RAE, RMSE, MAD, MSE, MAPE, AIC and ME. we selected those best algorithm that have highest r values, R^2 (%), Adj. R^2 (%), but the lowest RMSE, RAE, SD ratio, CV(%), MAPE, RAE, MAD, MSE and ME values, respectively. All statistical notations were obtained from a paper [14, 16, 22, 27-28], and a review written by Grzesiak and Zaborski [45]. The formulas of these criteria are given below:

Coefficient of determination (%)

$$R^2(\%) = \left[1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\right] * 100$$

Adjusted Coefficient of Determination (%)

$$Adj.R^2(\%) = \left[1 - \frac{\frac{1}{n-k-1}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\right] * 100$$

Coefficient of Variation (C.V%)

$$C.V(\%) = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2}}{\bar{Y}} * 100$$

Standard Deviation Ratio (SD_{ratio})

$$SD_{ratio} = \sqrt{\frac{\frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2}{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Global relative approximation error (RAE)

$$RAE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} Y_i^2}}$$

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}$$

Pearson correlation coefficient between actual and predicted values.

$r_{Y_i \hat{Y}_i}$

Akaike information criterion (AIC) calculated as

$$AIC = n \ ln\left[\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2\right] + 2k, if \frac{n}{k} > 40$$

or

AIC =

$$n \ ln\left[\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2\right] + 2k + \frac{2k(k+1)}{n-k-1}, otherwis$$

Mean error (ME)

$$M.E = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)$$

Mean absolute deviation (MAD)

$$MAD = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$$

Mean absolute percentage error (MAPE)

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| * 100$$

Mean square error (MSE)

$$\text{MSE} = \frac{\sum|Y_i - \hat{Y}_i|^2}{n}$$

Where y_i and $\hat{y}_i$ are observed and predicted BW values of the ith sheep. $\mathcal{E}_i$ is the residual value of ith sheep, $\bar{Y}$ and $\bar{\mathcal{E}}$ are the mean of actual BW and residual values. n is the size of the sample and k is the number of input variables used in the model.

3. RESULTS AND DISCUSSION

The summary results of different body measurements are given in Table 1.

Descriptive statistics (Mean ± SD) of BW (kg), WH, BL, HL, HW, EL, EW, NL, NW, HG, RL, RW, TL, BD, and SPW of all sheep in inches were: 6.37 (±1.76), 22.43 (±2.97), 21.45 (±3.25), 7.72 (±1.09), 3.28 (±0.76), 9.83 (±1.17), 4.33 (±0.56), 8.49 (±1.31), 5.38 (±0.81), 23.46 (±3.74), 4.27 (±0.83), 5.69 (±1.56), 4.09 (±0.96), 13.95 (±2.63) and 24.89 (±4.59), respectively.

Table 1. Descriptive statistics for body weight (kg.) and different measurements (inches) of Thali sheep.

Trait	Mean	SD	CV (%)
Body weight	6.378	1.768	27.720
Wither height	22.431	2.978	13.276
Body length	21.454	3.253	15.162
Head length	7.720	1.097	14.209
Head width	3.288	0.763	23.205
Ear length	9.831	1.175	11.951
Ear width	4.334	0.566	13.059
Neck length	8.496	1.310	15.419
Neck width	5.382	0.817	15.180
Heart girth	23.461	3.744	15.958
Rump length	4.275	0.835	19.532
Rump width	5.697	1.564	27.453
Tail length	4.090	0.969	23.691
Barrel depth	13.956	2.630	18.844
Sacral pelvic width	24.899	4.594	18.450

The summary results related to the performance of data mining algorithms to predict BW are presented in Table 2.

The correlation coefficient (r) between observed and predicted body weight, estimated for CHAID, Exhaustive CHAID, CART and ANN methods were 0.821, 0.802, 0.768 and 0.784, respectively. For the best algorithm selection, smaller value of SD ratio, CV (%), RAE, RMSE, MAD, MSE, MAPE, ME and AIC should be observed, but should greater in R^2 and Adj. R^2. In 2012, Grzesiak and Zaborski [45] also suggested that if the SD ratio value is less than 0.40 or between 0 and 0.10, the model will be a good-fit or a very good fit. In this study, we also computed the values of SD ratio for each data mining algorithm, which were reported to be 0.571, 0.596, 0.640 and 0.616, respectively. With the same order, percentage coefficient of variation (CV %) were found to be 15.85, 16.56, 17.76 and 17.26, respectively; percentage coefficients of determination (R^2 %) were 67.38, 64.37, 59.02 and 61.45 respectively; adjusted coefficient of determination (Adj-R^2), (%) were 60.29, 56.62, 50.11 and 53.52 respectively; relative approximation error (RAE) estimates were 0.151, 0.158, 0.170 and 0.162 respectively;

root mean square error (RMSE) estimates were 1.003, 1.049, 1.125 and 1.093 respectively; mean absolute deviation (MAD) estimates were 0.708, 0.758, 0.869 and 0.794 respectively; mean square error (MSE) estimates were 1.007, 1.100, 1.265 and 1.192 respectively; mean absolute percentage error (MAPE) estimates were 11.32, 12.43, 17.86 and 15.01 respectively; mean error (ME) estimates were 0.0008, 0.0016, -0.0014 and 0.094, respectively; and Akaike information criterion (AIC) estimates were 37.59, 45.09, 56.98 and 49.47 respectively. From these results, we found that SD ratio, CV (%), RAE, RMSE, MAD, MSE, MAPE and AIC values of the BW prediction model using CHAID algorithm were least and R^2 and Adj-R^2 values were greater as compared to the Exhaustive CHAID, CART and ANN methods. These results depicted that CHAID was the most superior decision tree algorithm having better fitting performance in BW prediction of Thalli sheep as compared to the Exhaustive CHAID, CART and ANN algorithms. In line with our results, some earlier reports highlighted the biological advantage of CHAID algorithms in BW prediction [2, 4, 14, 16]. The worst performance in the current research work was recorded for the CART tree based algorithm. A study of Yakubu [25] also obtained low R^2 (62.0 %) for predicting BW of UDA sheep by using the CART algorithm.

Due to the information that CHAID algorithm was the most suitable algorithm according to its results of a performance criterion, a decision tree was constructed for CHAID algorithm. In the decision tree diagram generated for CHAID algorithm, the most influential predictor was BL. Then, HL and SPW were determined to be the second degree significant predictors in the BW prediction of Thalli sheep (Figure 1). The model explain an accuracy of % 67.38 R^2 and 60.29 Adj-R^2 %, the variation of the BW on Thali sheep (Adj.P-value =0.000, F = 31.054, df1 = 4, df2 = 80).

Average BW of all 85 sheep in Node 0 (root node) was found to be (6.378 kg, S=1.768 kg). Node 0 was divided by BL (the most effective variable in the prediction body weight) into 5 subsets named Nodes 1-5, respectively (Adj. P-value = 0.000). Node 1 was a subgroup of sheep with BL ≤ 18.500 inch (BW= 6.044 kg, S=1.453 kg). Node 2 was a subgroup of sheep with 18.5 < BL < 21.500 inch among all the sheep (BW= 4.850 kg, S= 0.825 kg). The subgroup of sheep with a 21.500 < BL < 22.200 inch was entered into Node 3 in the decision tree construction of CHAID algorithm (BW= 6.124 kg, S= 0.967 kg). The subgroup of those having 22.200< BL< 25.000 inch was included in Node 4 through CHAID algorithm (BW= 7.162 kg, S=0.811 kg). Node 5 was the subgroup of sheep that is BL > 25.000 inch (BW=9.596 kg, S=2.029 kg). As BL increased from Node 1 to Node 5, sheep weight was also increased.

Node 1, on the basis of HL, was subsequently divided into Node 6 (BL ≤ 18.500 inch and HL ≤ 7.00 inch) and Node 7 (BL ≤ 18.500 and HL > 7.00 inch), respectively. The sheep (HL < 7.000 inches and BL ≤ 18.500 inch) in Node 7 were found lighter in BW than those (with HL ≤ 7.000 inch and BL ≤ 18.500 inch) in Node 6 (Adjusted P= 0.041; 4.74 vs. 6.594 kg).

Node 2 was sub-divided into Node 8 (SPW ≤ 20.00 inch and 18.500 < BL < 21.500 inch) and Node 9 (SPW >20.00 inch and 18.500 < BL< 21.500 inch) based on the values of SPW. Node 9 appeared to offer a better prediction of (BW=5.097 kg, S=0.782 kg) than the corresponding values of (BW=4.068 kg, S=0.321 kg) recorded for Node 8, respectively. Since Nodes 3, 4, 5, 6, 7, 8 and 9 were not divided into subgroups, they could be said to be homogenous (terminal Nodes).

As mentioned in the introduction section, continual importance is available for BW estimation

Table 2. Predictive performance of CHAID, Exhaustive CHAID, CART and ANN algorithms.

Algorithm	r	SD ratio	CV (%)	R^2 (%)	Adj.R^2 (%)	RAE	RMSE	MAD	MSE	MAPE	ME	AIC
CHAID	0.821	0.571	15.85	67.38	60.29	0.151	1.003	0.708	1.007	11.32	0.0008	37.59
Exhaustive CHAID	0.802	0.596	16.56	64.37	56.62	0.158	1.049	0.758	1.100	12.43	0.0016	45.09
CART	0.768	0.640	17.76	59.02	50.11	0.170	1.125	0.869	1.265	17.86	-0.0014	56.98
ANN	0.784	0.616	17.26	61.45	53.52	0.162	1.093	0.794	1.192	15.01	0.094	49.47

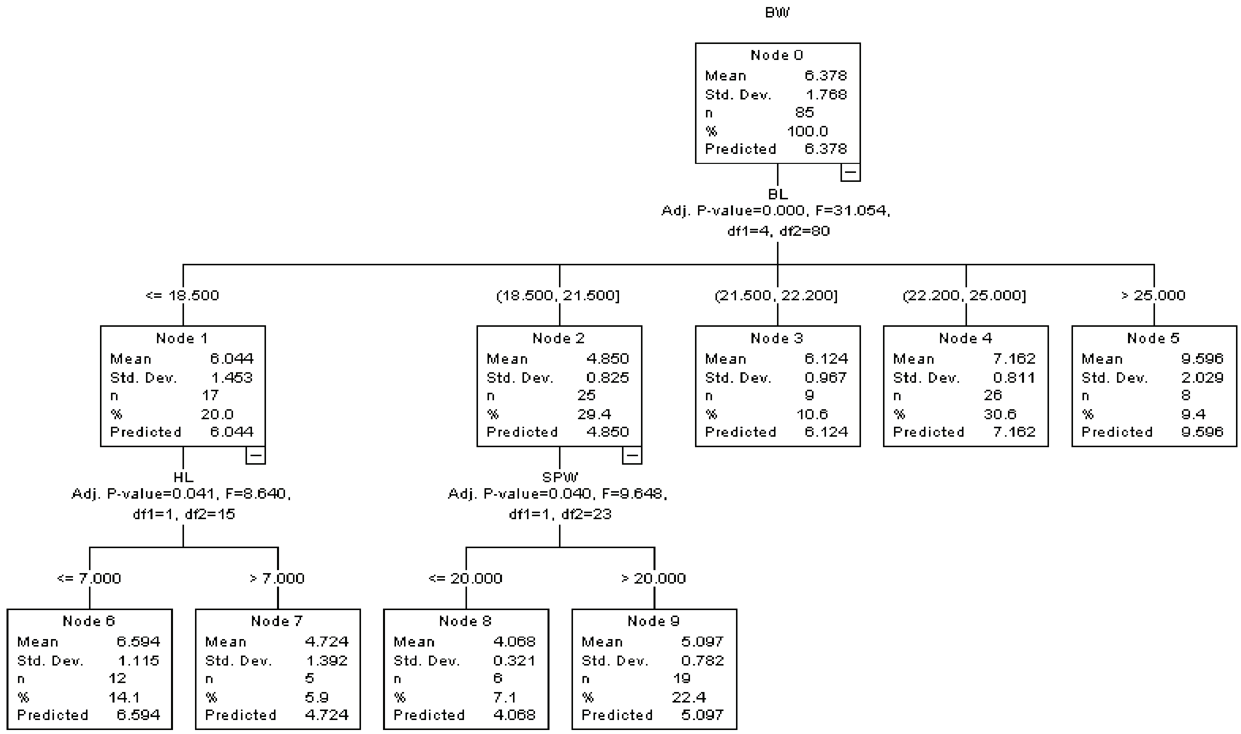

Fig. 1. Regression tree diagram for BW in sheep using CHAID algorithm

with the help of body measurements by using one variable, two variables regression analysis, factor scores and principle component analysis in multiple regression analysis [2]. Karabacak et al. [16] predicted the BW by means of several body measurements. In their study, they also used the different data mining algorithms and found that CHAID algorithm determined the better results for body weight prediction i.e., r = 0.805, R^2 = 64.8, SD ratio = 0.594 and RMSE =1.181) than Exhaustive CHAID (r =0.787, R^2 = 62.0, SD ratio =0.616 and RMSE =1.226) and CART (r =0.754, R^2 =56.9 %, SD ratio =0.657, and RMSE =1.306). These results are almost consistent to our study results. Another study by Mohammad et al. [4] also predicted the BW from chest girth and height at withers of indigenous sheep breeds of Pakistan. They used different methods but 72 % variability in BW was explained due to chest girth by using CHAID algorithm. Khan et al. [2] recorded (R^2 = 0.844) for Exhaustive CHAID algorithm in the estimation of BW of Harnai sheep from significant predictors, withers height, chest girth, face length, and body length. Their estimates were also better as compared to the results obtained from the current study. R^2 estimate of the CHAID algorithm in the present study was higher than those reported by Tyasi et al. [20]. Altay et al. [21] reported R^2 (0.956), adjusted R^2 (0.955), RMSE (0.377), RAE (0.018) and SD ratio (0.210) for CHAID algorithm in the description of factors affecting wool fineness in Karacabey Merino sheep, which was much higher than the current CHAID algorithms estimates. The estimation of BW in indigenous beetal goat of Pakistan through head girth, neck length, diagonal body length, belly sprung, shank circumference and rump height input variables was reported by Eyduran et al. [22] in the scope of CHAID, CART, RBF, MLP1, MLP2, MR modelling, the correlation coefficient (0.8475, 0.8212, 0.8643, 0.8199, 0.8339, 0.8620), AIC (594.16, 619.81, 1172.16, 890.97, -30.12, 582.34), RMSE (4.1569, 4.4687, 3.9398, 4.4860, 4.3267, 3.9731), SD ratio (0.5308, 0.5706, 0.5030, 0.5727, 0.5522, 0.5072) respectively. Their estimates were better compared to the results obtained from the present study. Ali et al. [14] used different data mining algorithms in the prediction of BW of Harnai sheep by using the body measurements such as body length (BL), withers height (WH), chest girth (CG), paunch girth (PG), face length (FL), length between ears (LBE), ear length (EARL), fat tail length (FTL) and fat tail width (FTW), respectively. They obtained R^2 estimates (83.77, 84.210, 82.644, 81.999), correlation coefficient (r) between observed and predicted BW values (0.915, 0.918, 0.909, 0.906), RMS (1.509, 1.488, 1.560, 1.589), RAE (0.0564, 0.0556, 0.0583, 0.0594), SD ratio (0.403, 0.397, 0.417, 0.423) for CHAID, Exhaustive CHAID, CART, and ANN algorithms respectively. Their model selection criteria estimates were better as compared to the estimates of the current study. The difference in the results may be due to the wide variation in ages, ecological conditions, breed, rearing systems, managerial factors, use of different body measurements and their interface degrees, and the statistical tools employed in the study. However, it is recommended for further investigators that the predictive performances of the evaluated data mining methods should be used for different sheep breeds and studies with a large population, large number of sheep breeds and efficient factors in generalization of the results obtained from the current study.

The main limitation of the study was that the male Thalli sheep wasn't considered in the present research work. For body weight prediction, this work should be carried out in the future.

4. CONCLUSION

In conclusion, we found that CHAID algorithm is a best algorithm (R^2= 67.38, SD ratio = 0.571 and RMSE = 1.003) for determining the BW predictions of Thalli sheep. The most significant predictors in the model are BL, HL and SPW on BW. The predicted BW equation of Thalli sheep would be helpful for researchers, breeders and veterinary doctors of southern Punjab to determine the body weight of Thalli sheep. The researchers may also use these results for comparison purposes and may be used as a reference for next studies.

5. ACKNOWLEDGEMENTS

We wish to acknowledge and thank all the farmers who allowed us to measure the different morphological measurements of Thalli sheep.

6. CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

7. REFERENCES

1. K. Karadas., M. Tariq, M.M. Tariq, and E. Eyduran. Measuring Predictive Performance of Data Mining and Artificial Neural Network algorithms for predicting lactation milk yield in Indigenous Akkaraman Sheep. *Pakistan Journal of Zoology.* 49: 1-7 (2017).
2. M.A. Khan., M.M. Tariq, E. Eyduran, A. Tatliyer, M. Rafeeq, F. Abbas, N. Rashid, M.A. Awan, and K. Javed. Estimating body weight from several body measurements in Harnai sheep without a multicollinearity problem. *The Journal of Animal & Plant Sciences.* 24(1): 2014, Page: 120-126. ISSN: 1018-7081 (2014).
3. S. Kumar., S.P. Dahiya, Z.S. Malik, and C.S. Patil. Prediction of body weight from linear body measurements in sheep. *Indian Journal of Animal Research.* ISSN: 0976-0555 (2017).
4. M.T. Mohammad., M. Rafeeq, M.A. Bajwa, M.A. Awan, F. Abbas, A. Waheed, F.A. Bukhari, and P. Akhtar. Prediction of Body Weight from Body Measurements Using Regression Tree (RT) Method for Indigenous Sheep Breeds in Balochistan, Pakistan. *The Journal of Animal & Plant Sciences*. 22(1): 2012, Page: 20-24. ISSN: 1018-7081 (2012).
5. C.A. Valdez and J.P.L. Valencia. Weight Prediction in Medium to large sized adult phillipine native dogs using external body measurements. *Phillipine Journal of Animal & Plant Sciences.* 30: 161-169 (2004).
6. S. Berge. On the Estimation of Weight and Increase of Weight by means of the Chest Girth in Norwegain Red Cattle at the Agricultural University at As, Norway in the years 1972 and 1974. *Acta Agriculturae Scandinavica.* 27: 65-66 (1977).
7. V. Buvanendran., J.E. Umoh, and B.Y. Abubakar. An evaluation of body size as related to weight of three West African breeds of cattle in Nigeria. *The Journal of Agricultural Science.* 95: 219-224 (1980).
8. L.A. Goonerwaedene, and P. Sahaayuruban. Analysis of body measurements and prediction of body weight in crossbred Lanka bulls. Proc. 5th World Conf. *Prod*. 2: 27-28
9. M. Topal., N. Yildiz, N. Esenbuga, V. Akaskal, M. Macit, and M. Ozdemir. Determination of Best

Fitted Regression Model for Estimation of Body Weight in Awassi Sheep. *Journal of Applied Animal Research.* 23: 201-208 (2003).

10. M. Topal, and M. Macit. Prediction Body Weight from Body Measurements in Morkaraman Sheep. *Journal of Applied Animal Research.* 25: 97-100 (2004).
11. M.A. Cam., M. Olfaz, and E. Soydan. Possibilities of Using Morphometrics Characteristics as a Tool for Body Weight Prediction in Turkish Hair Goats (Kilkeci). *Asian Journal of Animal and Veterinary Advances.* 5: 52-59 (2010).
12. U. Younas., M. Abdullah, J.A. Bhatti, T.N. Pasha, N. Ahmad, M. Nisar, and A. Hussain. Interrelationship of body weight with linear body measurements in Hissardale sheep in different stages of life. *Journal of Animal & Plant Sciences.* 23: 40-44 (2013).
13. A. Yakubu. Fixing Collinearity Instability in the Estimation Body Weight from Morpho-biometrical Traits of West African Dwarf Goats. *Trakia Journal of Sciences.* 7: 61-66 (2009).
14. M. Ali., E. Eyduran, M.M. Tariq, C. Trink, F. Abbas, M.A. Bajwa, M.H. Baloch, A.H. Nizamani, A. Waheed, M.A. Awan, S.H. Shaw, Z. Ahmad, and S. Jan. Comparison of Artificial Neural Network and Decision Tree Algorithms used for Predicting Live Weight at Post Weaning Period from Some Biometrical Characteristics in Harnai Sheep. *Pakistan Journal of Zoology.* 47(6) 1579-1585 (2015).
15. M. Akin., E. Eyduran, and B.M. Reed. Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. *Plant Cell Tissue and Organ Culture.* 128: 303-316 (2017).
16. A. Karabacak., S. Celik, A. Tatliyer, I. Keskin, Y.E. Erturk, E. Eyduran, Y. Javed, and M.M. Tariq. Estimation of Cold Carcass Weight and Body Weight from Several Body Measurements in Sheep through Various Data Mining Algorithms. *Pakistan Journal of Zoology.* 49(5): 1731-1738 (2017).
17. M. Akin., C. Hand, E. Eyduran, and B.M. Reed. Predicting minor nutrient requirements of hazelnut shoot cultures using regression trees. Plant Cell, *Tissue and Organ Culture.* 132: 545-559 (2018)
18. C. Gozuacik., E. Eyduran, H. Cam, and M.K. Kara. Detection of infection preferences of the alfalfa seed chalcid, Bruchophagus roddi Gussakovskiy, 1933 (Hymenoptera: Eurytomidae) in alfalfa (Medicago sativa L.) fields of Igdir, Turkey. *Legume Research.* 41(1): 150-154 (2018)
19. E. Zaborski., M. Ali, E. Eyduran, W. Grzesiak, M.M. Tariq, F. Abbas, A. Waheed, and C. Tirink. Prediction of selected reproductive traits of Indigenous Harnai sheep under the farm management system via various data mining algorithms. *Pakistan Journal of Zoology.* 51(2): 421-431 (2019)
20. T.L. Tyasi., E. Eyduran, and S. Celik. Comparison of tree-based regression tree methods for predicting live body weight from morphological traits in Hy-line silver brown commercial layer and indigenous Potchefstroom koekoek breeds raised in South Africa. *Tropical Animal Health and Production.* 53:7. https://doi.org/10.1007/s11250-020-02443-y (2021)
21. Y. Altay., S. Boztepe, E. Eyduran, I. Keskin, M.M. Tariq, F.A. Bukhari, and I. Ali. Description of factors affecting wool fineness in Karacabey Merino sheep using chaid and Mars Algorithms. 53(2): 691-697 (2021)
22. E. Eyduran., D. Zaborski, A. Waheed, S. Celik, K. Karadas, and W. Grzesiak. Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of body Measurements in the Indigenous Beetal Goat of Pakistan. *Pakistan Journal of Zoology.* 49(1) 257-265 (2017).
23. Y. Koc., E. Eyduran, and O. Akbulut. Application of Regression Tree for Different Data from Animal Science. *Pakistan Journal of Zoology*. 49: 599-607 (2017).
24. S. Celik, and O. Yilmaz. Comparison of different data mining algorithms for prediction of Body weight from several morphological measures in dogs. *The Journal of Animal and Plant Sciences.* 27 (1): 57-64 (2017).
25. A. Yakubu. Application of Regression Tree Methodology in Predicting the Body Weight of Uda Sheep. *Scientific Papers: Animal Science and Biotechnologies*. 45(2) (2012).
26. I.Y. Kovalchuk., Z. Mukhitdinova, T. Turdiyv, G. Madiyeva, M. Akin, E. Eyduran, and B.M. Reed. Modeling some mineral nutrient requirements for micropropagated wild apricot shoot cultures. *Plant Cell, Tissue and Organ Culture.* 129(2): 325-335 (2017).
27. S. Celik., E. Eyduran, K. Karadas, and M.M. Tariq. 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 (2017).
28. S. Celik. Comparing Predictive Performances of

Tree Based Data Mining Algorithms and Mars Algorithm in the Prediction of Live Body Weight from Body Traits in Pakistan goats. *Pakistan Journal of Zoology.* 51(4) 1447-1456 (2019)

29. M. Mendes, and E. Akkartal. Regression tree analysis for predicting slaughter weight in broilers. *Italian Journal of Animal Science.* 8: 615-624 (2009)
30. A. Aksoy., Y.E. Erturk, E. Eyduran, and M.M. Tariq. Utility of MARS Algorithm for describing non-genetic factors affecting pasture revenue of Morkaraman Breed and Romanov × Morkaraman F1Crossbred sheep under semi intensive conditions. *Pakistan Journal of Zoology.* 51(1): 235-240 (2019)
31. E. Eyduran., M. Akin, and S.P. Eyduran. Application of multivariate adaptive regression splines through R software. *Nobel Academic Publishing, Ankara. FAO,* 2019. http://www.fao.org/faostat/en/#data/QA Access date:30.10.2020
32. F. Iqbal., E. Eyduran, A. Raziq, M. Ali, Z.E. Huma, C. Tirink, and H. Sevgenler. Modeling and predicting the growth of indigenous Harnai sheep in Pakistan: non-linear functions and MARS algorithm. *Tropical Animal Health and Production.* 53(2):248. doi: 10.1007/ s11250-021-02700-8
33. A. Fatih., S. Celik, E. Eyduran, C. Tirink, M.M. Tariq, I.S. Sheikh, A. Faraz, and A. Waheed. Use of MARS Algorithm for predicting mature weight of different camel (Camelus dromedaries) breeds reared in Pakistan and morphological characterization via cluster analysis. *Trpical Animal Health and Production.* 53:191 (2021)
34. I. Aytekin., E. Eyduran, K. Karadas, R. Aksahan, and I. Keskin. 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 algorithms. *Pakistan Journal of Zoology.* 50: 189-195 (2018)
35. G.V. Kass. An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Journal of the Royal Statistical Society.* 29: 119-127 (1980).
36. J.A. Michael and S.L. Gordon. Data Mining Technique: For Marketing, Sales, and Customer support. *New York: Wiley* (1997).
37. E. Eyduran., K. Karakus, S. Keskin, and F Cengiz. Determination of Factors Influencing Birth Weight Using Regression Tree (RT) method. *Journal of Applied Animal Research* 34: 109-112 (2008).
38. V. Jarosik. CART and related methods. In, Simberloff and D, Rejmanek M (Eds): Encyclopaedia of Biological Invasions. *University of California Press, Berkeley and Los Angeles.* 104-108 (2011).
39. H. Orhan., E. Eyduran, A. Tatliyer, and H. Saygici. Prediction of egg weight from egg quality characteristics via ridge regression and regression tree methods. *Revista Bras. de Zootecnia.* 45: 380-385 (2016).
40. D. Biggs., B.D. Ville, and E. Suen. A method of choosing multiway partitions for classification and decision trees. *Journal of Applied Statistics.* 18: 49-62 (1991).
41. L. Breiman., J.H. Friedman, R.A. Olshen, and C.J. Stone. Classification and regression trees. *Chapman and Hall, Wadsworth Inc., New York, NY, USA.*
42. I.K. Omurlu., M. Ture, and A.T. Kurum. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Systems with applications.* 34(1): 366-374 (2008).
43. O. Gorgulu. Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *South African Journal of Animal Science.* 42(3): 280-287 (2012).
44. E. Eyduran., I. Keskin, Y.E. Erturk, B. Dag, A. Tatliyer, C. Tirink, R. Aksahan, and M.M. Tariq. Prediction of fleece weight from wool characteristics of sheep using regression tree method (Chaid Algorithm). *Pakistan Journal of Zoology.* 48: 957-960 (2016)
45. W. Grzesiak and D. Zaborski. Examples of the Use of Data Mining Methods in Animal Breeding. ISBN