The Journal of Animal & Plant Sciences, 27(1): 2017, Page: 57-64
ISSN: 1018-7081

COMPARISON OF DIFFERENT DATA MINING ALGORITHMS FOR PREDICTION OF BODY WEIGHT FROM SEVERAL MORPHOLOGICAL MEASUREMENTS IN DOGS

S. Celik*[1] and O. Yilmaz[2]

[1]Bingol University, Faculty of Agriculture, Department of Animal Science, Bingol, Turkey
[2] Ardahan University, Vocational High School of Technical Science, Ardahan, Turkey
*Corresponding Author: senolcelik@bingol.edu.tr

ABSTRACT

The aim of this study was to find the best one among CHAID **(**Chi-square Automatic Interaction Detector), Exhaustive CHAID, and CART (Classification and Regression Tree) data mining algorithms in the prediction of body weight (BW) from several body measurements (abdominal width (AW), body length (BL), chest circumference (CC), chest depth (CD), face length (FL), front shank circumference (FSC), head circumference (HC), head length (HL), head width (HW), leg length (LL), tail length (TL), rear chest width (RCW), rump elevation (RE), rump width (RW), withers height (WH)) measured easily from three Kangal (Karabash) dog color varieties (Dun/Fawn, Grizzle, and Ashy) maintained in Sivas and Konya provinces, Turkey. Several goodness-of-fit criteria (coefficient of determination (R^2%), adjusted coefficient of determination (Adj.R^2%), coefficient of variation (CV%), SD ratio, Root Mean Square Error (RMSE), Relative Approximation Error (RAE), Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE), and Pearson correlation between actual and predicted values were estimated for describing the most suitable algorithm in terms of the predictive performance. r values are 0.846, 0.838 and 0.732 for CHAID, Exhaustive CHAID and CART algorithms, respectively. RMSE values are 4.966, 5.083 and 6.349 for CHAID, Exhaustive CHAID and CART algorithms, respectively. The most important predictors are BE of BW for all algorithms. Among the algorithms, CHAID provided the most appropriate predictive capability in the prediction of the BW characteristic. The heaviest average BW of 61.375 kg was obtained from the subgroup of those having FSC > 14 cm and RE > 80 cm. The secondly heaviest average BW (53.455kg) was found for the subgroup of those having FSC > 13 cm and 74.000 $<$ RE $\leq$ 80 cm in Sivas province of Turkey. Consequently, it is hoped that the results of the study on the morphological characterization of Kangal dog varieties might be a good reference for next dog breeding studies.

Keywords: CHAID; Exhaustive CHAID; CART; Karabash dog; body weight.

INTRODUCTION

Usability of some morphological measurements in absence of the scale is very essential in the BW prediction (Valdez and Valencia, 2004). The BW prediction is needful for providing knowledge on appropriate drug dose and feed amount for an animal (Khan *et al.,* 2014). Emehelu *et al.* (2012) estimated BW from body measurements in Nigerian local dogs and recorded very strongly correlations between the BW and the measurements. Valdez and Valencia (2004) predicted BW using morphological measurements (54 females and 46 males, total of 100) in adult Philippine native dogs through regression and correlation analyses for each gender and recorded positive relationships between BW and the body measurements, regardless of gender. The BW has been predicted by non-linear growth functions (Logistic, Brody, Gompertz and Von-Bertalanffy) in Kangal dogs (Coban *et al.,* 2011). In addition, Yildirim (2012) used general linear model to determine the effect of birth season, sex, dam age, and number of baby dogs per birth on BW at 1st, 2nd, and 3rd months in Kangal dogs. Atasoy *et al.* (2011) reported averages of body weight, body and head measurements taken at different age and gender groups of Akbas dogs. However, no declared document is obtainable on implementing data mining algorithms in the BW prediction from morphological characteristics in especially the dogs. This means that further studies on the BW prediction of the dogs should be performed.

The tree-based CART, CHAID, and Exhaustive CHAID algorithms are applied for scale, nominal and ordinal response variables in order to obtain homogenous subgroups as soon as possible, depending upon sample size, structures of independent variables (nominal, ordinal and scale), non-linear and the interaction effects of the independent variables (Ali *et al.,* 2015). CART algorithm constructs a binary decision tree by partitioning a node into two new child nodes, recursively, whereas both CHAID algorithms with three essential stages (merging, splitting, and stopping) form a decision tree consisting multiple splits repeatedly, and have a difference merging stage from each other in the construction of the regression tree.

Phenotypic characterization of Karabash dog varieties is a very important tool for further breeding investigations in the conservation of domestic gene sources in Turkey. An accurate characterization depends on the selection of proper statistical techniques. As part of regression analysis, CART, CHAID, and Exhaustive CHAID data mining algorithms have been applied for BW prediction in sheep breeding (Ali *et al.*, 2015), and cattle breeding (Aksahan and Keskin, 2015); however, no previous work was recorded on application of the algorithms for predicting BW by means of morphological traits in dogs. Hence, the aim of the current study were to compare predictive capabilities of CHAID, EXHAUSTIVE CHAID, and CART data mining algorithms in the prediction of body weight (BW) from several body measurements from three dog skins, respectively on the basis of several goodness-of-fit criteria.

MATERIALS AND METHODS

Subjects and data set. In this study, it was examined some morphological characteristics of Kangal (Karabash) Dogs raised in countryside private breeders, in civil and state farms. In the study, the morphological data of 208 Kangal (Karabash) dogs (101 males and 107 females) belonging to three color varieties (Dun/Fawn (184), Ashy (14), and Grizzle (10) in Konya and Sivas provinces, Turkey were used. The present data were taken from Yilmaz (2007) to initially apply the data mining algorithms for the dog data. Names and abbreviations of independent (scale=continuous) variables used in the study were summarized in Table 1. The number of dogs in ages of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11 were 52, 46, 43, 28, 15, 11, 8, 1, 2, 1, and 1 respectively (Table 2). Frequency statistics (frequency and percent values) for flock, sex and color are given in Table 3.

Table 1. Morphological measurements used in the study and descriptive statistics

Variable	Body measurements (cm)	Mean	SD
AW	Abdominal width	15.507	1.994
BL	Body length	81.630	9.633
BW	Body weight	45.471	7.714
CC	Chest circumference	86.471	6.659
CD	Chest depth	31.264	2.899
FL	Face length	12.739	1.036
FSC	Front shank circumference	13.313	0.884
HC	Head circumference	52.899	4.519
HL	Head length	30.317	2.211
HW	Head width	13.877	1.650
LL	Leg length	35.598	3.483
QL	Tail length	47.389	3.248
RCW	Rear chest width	19.320	2.085
RE	Rump elevation	73.423	5.196
RW	Rump width	22.356	2.106
WH	Withers height	74.514	5.106

SD (Standard deviation)

Table 2. Frequency statistics for age

Age	Frequency	Percent
1	52	25.00
2	46	22.12
3	43	20.67
4	28	13.46
5	15	7.21
6	11	5.29
7	8	3.85
8	1	0.48
9	2	0.96
10	1	0.48
11	1	0.48
Total	208	100.00

Mean: 3.024, SD 1.940

Table 3. Frequency statistics for flock, sex and color

Flock	Frequency	Percent
1	117	56.250
2	45	21.635
3	46	22.115
Total	208	100
Sex	Frequency	Percent
Female	107	51.442
Male	101	48.558
Total	208	100
Color	Frequency	Percent
Dun/Fawn	184	88.462
Grizzle	10	4.808
Ashy	14	6.731
Total	208	100

In the study, flock (1, 2, and 3), color (Dun/Fawn, Ashy and Grizzle) and sex (male and female) are nominal (independent) variables, but age and other (independent) variables in Table 1 are continuous variables. In the study, AW, BL, CC, CD, FL, FSC, HC, HL, HW, LL, QL, RCW, RE, RW, WH, age, sex, skin color and farm were included as independent variables in the prediction of BW, as a dependent (continuous=scale) variable.

Data Mining Algorithms: Classification and Regression Tree (CART) method is a recursive partitioning method used both for regression and classification problems. The best predictor is chosen using a variety of impurity or diversity measures. The aim is to produce subsets of the data which are as homogeneous as possible with respect to the target variable (Breiman *et al.*, 1984). Chi-squared Automatic Interaction Detector (CHAID) method is used based on the chi-square test of association. A CHAID tree is a decision tree that is constructed by repeatedly splitting subsets of the space into two or more child nodes, beginning with the entire data set (Michael and

Gordon, 1997). Exhaustive CHAID has the same splitting and stopping steps 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 (Biggs *et al*., 1991).

For regression and classification problems, the tree-based CART, CHAID, and Exhaustive CHAID data mining algorithms, which may provide ones to characterize morphological traits for detecting standards of Kangal (Karabash) dogs, are available in SPSS statistical package program. But, in the study, as part of general linear model for regression type problem, we have used the tree-based data mining algorithms to construct optimal decision tree in the prediction of BW from continuous variables (age and morphological characteristics) and categorical (nominal) variables (flock, color, province and sex), respectively (Khan *et al.*, 2014; Ali *et al.,* 2015). In this study, flock, color, province and sex are nominal variable. Pruning operation was made automatically for both CHAID algorithms, but not activated in CART data mining algorithm giving binary splitting nodes recursively in the decision tree structure. V fold cross validation is set at 10 (Mendes and Akkartal, 2009). CHAID algorithm is effectively implemented for ordinal, nominal, and continuous variables. Also, the CHAID algorithm examines non-linear and interaction effects of independent variables. The CHAID algorithm employs merging, splitting, and stopping stages for constructing a regression tree diagram, and converts continuous variables into ordinal variables. It yields homogenous subgroups (nodes) by splitting nodes, repeatedly for maximizing variance in dependent variable among nodes (Nisbet *et al*., 2009; Orhan *et al*., 2016). Bonferroni adjustment was performed on CHAID algorithm in order to calculate Adjusted P values of F values (Ali *et al*., 2015; Eyduran *et al*., 2016; Akin *et al*., 2016). CHAID data mining algorithm automatically pruning insignificant nodes in a decision tree constructed through IBM SPSS program is worked on the basis of F test if a continuous dependent variable is used as in our study (Orhan *et al*., 2016). A ten-fold cross validation was activated in the study. In CHAID and Exhaustive CHAID data mining algorithms, pruning operation was automatically performed in IBM SPSS 22 statistical package program, but pruning operation in CART algorithm must be activated by analysis. The aim of CHAID algorithm is to minimize variance within nodes in the dependent variable during constructing regression tree diagram.

Minimum dog numbers for parent and child nodes were fixed at 10 and 5 for constructing optimal decision tree structure and improving predictive performance of the algorithms. SPSS automatically made Bonferroni adjustment to calculate adjusted P values for both CHAID algorithms with multiple splitting nodes. But, the adjustment is unavailable in CART algorithm.

The main target of the algorithms is to minimize the variation within nodes in order to construct homogenous subgroups in the optimal decision tree diagram with significant independent variables.

2.3. Goodness-of-fit criteria: To determine the best one among the data mining algorithms, we calculated several goodness-of-fit criteria described by Takma *et al.* (2012), Grzesiak and Zaborski (2012) and Ali *et al.* (2015), respectively. The related goodness-of-fit criteria can be formulated as follows:

Coefficient of Determination (%)

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

Adjusted Coefficient of Determination (%)

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

Coefficient of Variation (%)

$$CV(\%) = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\pounds_i - \bar{\pounds}\right)^2}}{\bar{Y}}*100$$

Standard Deviation Ratio

$$SD_{ratio} = \sqrt{\frac{\frac{1}{n-1}\sum_{i=1}^{n}\left(\pounds_i - \bar{\pounds}\right)^2}{\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2}}$$

Relative Approximation Error

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

Root Mean Square Error

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

Mean Absolute Deviation

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

Mean Absolute Percentage Error

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

where, Y_i: actual BW value of ith dog, $\hat{Y}_i$: the predicted BW value of ith dog, $\bar{Y}$: mean of the actual BW values of ith dog, ε_i: the residual value of ith dog associated with BW, $\bar{\varepsilon}$: mean of the residual values associated with BW, *k*: number of independent variables included significantly in the model, and n: total sample size.

In addition, Pearson correlation coefficient (r) between actual and predicted BW values was estimated. In the present survey, we selected the best algorithm having the greatest R^2 (%), Adj. R^2 (%), and r values, but the lowest CV(%), RAE, RMSE, SD ratio, MAPE, and MAD values, respectively. All statistical notations were received from a paper (by Ali *et al*., (2015) and a review written by Grzesiak and Zaborski, (2012). More detailed information is available in Grzesiak and Zaborski, (2012) with URL: http://cdn.intechopen.com/pdfs-wm/385.pdf. All the statistical evaluations were carried out by using IBM SPSS 22.0 software program.

RESULTS

Summary results of goodness-of-fit criteria estimated for data mining algorithms in BW prediction are presented in Table 4. For CHAID and Exhaustive CHAID algorithms, BW was significantly affected by RE, CC, FSC, farm, province and RCW independent variables. For CART algorithm, BW was significantly affected by RE, FSC and CC independent variables. We selected the CHAID as the ideal data mining algorithm according to its values of goodness-of-fit criteria. In this regard; CHAID algorithm constructed the tree-based decision tree structure. Its decision diagram is presented in Figure 1.

Table 4. Predictive performance results of goodness-of-fit criteria for data mining algorithms for BW trait.

Algorithm	r	SD ratio	CV%	R^2(%)	Adj-R^2 (%)	RAE	RMSE	MAD	MAPE
CHAID	0.846	0.533	10.720	71.575	70.727	0.089	4.966	4.472	9.858
EX. CHAID	0.838	0.546	10.973	70.214	69.325	0.091	5.083	4.584	10.152
CART	0.732	0.682	13.705	53.537	52.854	0.114	6.349	5.830	13.085

The overall BW average of 45.471 (S=7.714) kg was predicted for Node 0 (208 Kangal (Karabash) dogs) at the top of the regression tree diagram constructed by the CHAID algorithm. Node 0, a root node, was divided into five new child nodes (Nodes 1- 5) according to RE trait, respectively (Adj-P=0.000, F=50.119, df1=4, and df2=203). The visual result reflected that BW increased as RE increased from Node 1 through Node 5. In the tree diagram, terminal nodes were Nodes numbered 6, 7, 14, 15, 16, 17, 18, 19, 20, 21, 22, and 23, respectively.

As a subgroup of Kangal (Karabash) dogs with RE $\leq$ 67 cm, Node 1 had the average body weight (BW) of 37.545 (S=3.233) kg in the decision tree diagram. CC had a considerable effect on BW of the dogs in Node 1, and divided Node 1 into two new child Nodes numbered 6 and 7, respectively (34.500 vs. 39.286 kg). Node 6 was determined to be the subgroup of those having CC $\leq$ 77 cm and RE$\leq$67 cm, and Node 7 was the subgroup of those having CC > 77 cm and RE$\leq$67 cm.

Node 2, the subgroup of those with 67 < RE $\leq$ 72, was branched into Nodes 8 and 9, with respect to FSC, respectively (39.000 vs. 42.431 kg). Node 8, a terminal node, is the subgroup of those having FSC $\leq$ 12.50 cm and 67 < RE $\leq$ 72, and was not affected by any morphological measurement. However, Node 9, the subgroup of those having FSC > 12.50 cm and 67 < RE $\leq$ 72, was divided into two child nodes numbered 16 and 17 according to farm factor, respectively (41.632 vs. 44.769 kg). Node 16 was found lighter in BW than Node 17, statistically (Adj-P=0.039, F=6.626, df1=1, and df2=49).

Node 3, the subgroup of those having 72 < RE $\leq$ 74 cm, was partitioned at first stage into the child nodes numbered 10 and 11 in terms of farm factor, respectively (47.762 vs. 37.857 kg). Node 10, the subgroup of those with 72 < RE $\leq$ 74 cm reared in peasant and personal farms, was heavier in BW than Node 11, the subgroup of those with same RE values in formal farm, (Adj-P=0.000, F=25.426, df1=1, and df2=26). Also, CC divided Node 10 into the new child nodes numbered 18 and 19, respectively (45.200 vs.50.091 kg). Node 18 was the subgroup of the subgroup of those with CC $\leq$84 cm and 72 < RE $\leq$ 74 cm reared in peasant and personal farms, and Node 19 was assigned as the subgroup of those with CC >84 cm and 72 < RE $\leq$ 74 cm reared in peasant and personal farms in the decision tree diagram.

Province factor split Node 4, the subgroup of those having 74.000< RE$\leq$ 80 cm, into two new child nodes numbered 12 and 13 (47.500 vs. 51.977 kg), respectively. Node 12, the subgroup of those having 74.000< RE≤ 80 cm in Konya province of Turkey, was obtained lighter in BW than Node 13, the subgroup of those having 74.000 < RE ≤ 80 cm in Sivas province of Turkey, (Adj-P=0.011, F=6.922, df1=1 and df2=59). Node 12 in BW was affected by RCW, and was divided into two child nodes numbered 20 and 21 (41.750 vs. 52.100 kg), respectively (Adj-P=0.000, F=40.008, df1=1 and df2=16); However, Node 13 influenced by FSC was split into two child nodes numbered 22 and 23 (47.100 vs. 53.455 kg), (Adj-P=0.011, F=10.667, df1=1 and df2=41). In the tree diagram, Node 20 was identified as the subgroup of those having RCW ≤ 21 cm and 74.000 < RE ≤ 80 cm in Konya province of Turkey, while Node 21 was the subgroup of those having RCW > 21 cm and 74.000 < RE ≤ 80 cm in Konya province of Turkey. Node 22 was detected as the subgroup of those having FSC ≤ 13 cm and 74.000 < RE≤ 80 cm in Sivas province of Turkey, but Node 23 presented the subgroup of those having FSC > 13 cm and 74.000< RE≤ 80 cm in Sivas province located in the Central Anatolia Region of Turkey.

Node 5, the subgroup of those having RE > 80 cm, was divided into two child nodes numbered 14 and 15, (50.417 vs. 61.375 kg), respectively (Adj-P=0.000, F=22.714, df1=1 and df2=18). Node 14, the subgroup of those having FSC ≤ 14 cm and RE > 80 cm, was recorded lighter in BW in comparison to Node 15, the subgroup of those having FSC > 14 cm and RE > 80 cm, statistically.

In conclusion, Nodes with the BW > 50.000 kg was terminal nodes numbered 14, 15, 19, 21, 23, respectively, but Node 15 provided the heaviest BW among them.

DISCUSSION

This document is the first data mining modeling study in the prediction of BW by means of several morphological characteristics and categorical variables. This current finding is in agreement with those reported by some investigators (Khan *et al.,* 2014; Ali *et al.,* 2015).

In this study, predictive performance of the tree-based data mining algorithms used for predicting body weight by means of some morphological measurements in dogs has been evaluated comparatively. However, Valdez and Valencia (2004) predicted BW from external body measurements in adult Philippine native dogs by using regression and correlation analyses for each gender and they found positive relationships between BW and the body measurements. But, better statistical techniques like data mining algorithms can be adopted to detect the relationship between BW and other morphological traits. In literature, a good chapter was published recently on the application of several data mining algorithms like CART, MARS, and Naïve Bayes for both classification and regression tasks in animal data. Using several goodness-of-fit criteria, Ali *et al.* (2015) reported the data on the predictive accuracy of CART, CHAID, Exhaustive CHAID, and Artificial Neural Network (ANN) in the BW prediction from morphological traits in indigenous Harnai sheep of Pakistan.

For regression tasks, Yakubu (2012), Mohammad *et al.* (2012), Khan *et al.* (2014) and Ali *et al.* (2015) used data mining algorithms CART and CHAID algorithms to predict BW from morphological traits in sheep with the %R^2 estimates of 61.8, 72.0, 84.4, respectively. Ali *et al.* (2015) used data mining algorithms ANN and CHAID algorithms to predict BW from morphological traits in sheep with the %R^2 estimates of 82 and 83.77, respectively.

As well as the %R^2 estimates varying between 82 (ANN)-83.77 (CHAID), respectively. In a similar study, Aksahan and Keskin (2015) found the %R^2 of 87.82 and Adj. %R^2 of 87.32 in the BW prediction from morphological traits in young bulls, which was found lower than the present %R^2 estimates for all the algorithms under the study. The current %R^2 estimates for both CHAID algorithms were obtained lower in comparison with those recorded by Mohammad *et al.* (2012), Khan *et al.* (2014) and Ali *et al.* (2015). The difference may be ascribed to animal number, animal age, animal species, number and structures of the used variables, and interaction between the variables.

Valdez and Valencia (2004) recorded significant correlations between BW and some external morphological traits (thoracic girth, midriff girth, flank girth, body length, height at point of the shoulder, and height at withers), but the highest significant correlation of 0.684 for BW-thoracic girth. Emehelu *et al.* (2012) found very strongly relationship (92.7 %R^2) between the BW and CC in Nigerian local dogs. Until now, no reported publication was found on the application of data mining algorithms in the BW prediction on the dogs. As part of data mining applications, the present survey is a pioneer study conducted for the first time on the dogs in the BW prediction from differential body measurements, and therefore, the present results could not be made a comparison with previous published studies on sheep and cattle species.

Both CHAID algorithm and Exhaustive CHAID algorithm almost equally good performance (judging from the figures in Table 4). It has been tried with limited sample (208 dogs).

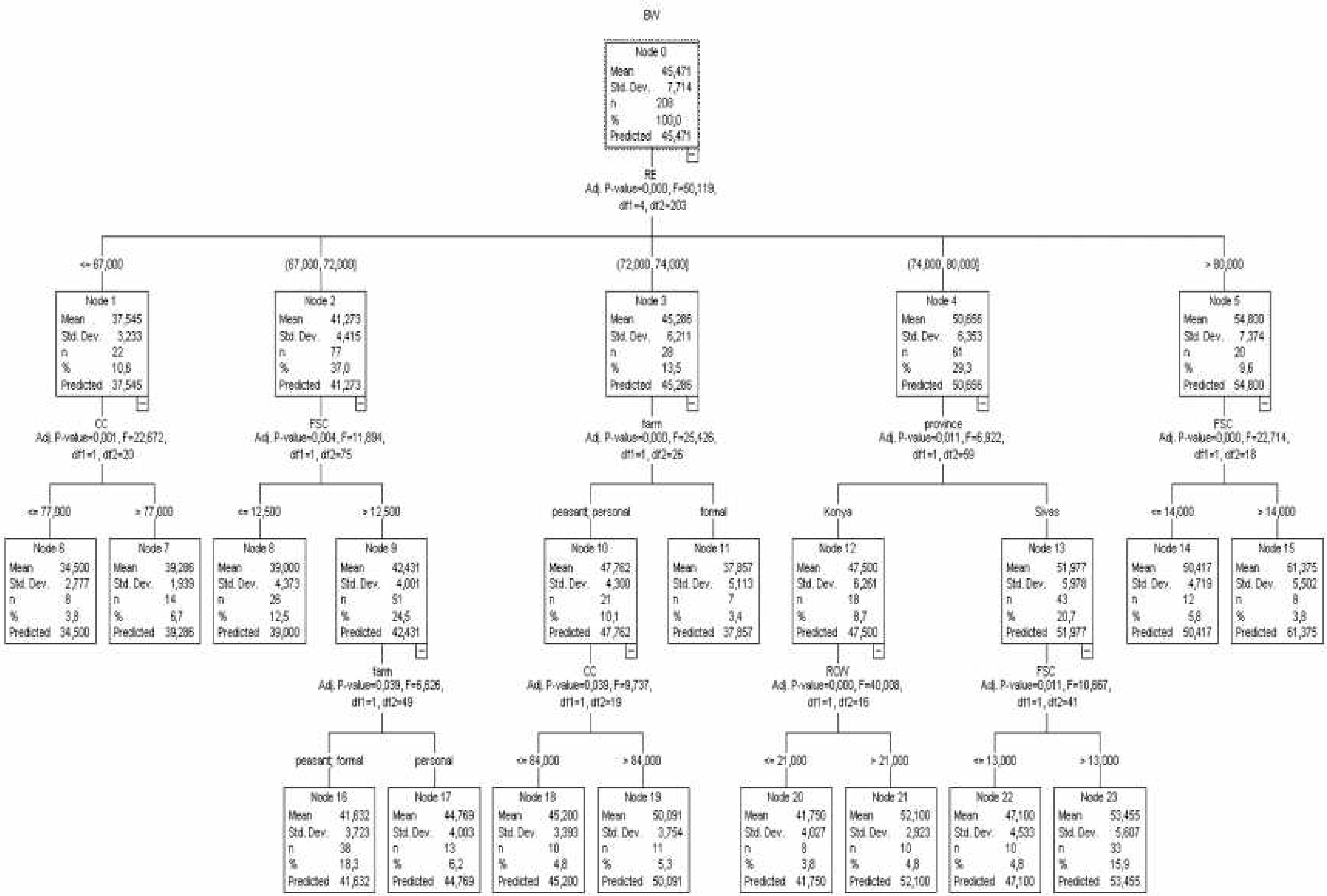

Figure 1. The decision tree diagram constructed by CHAID algorithm for BW prediction

Conclusion: The survey presented the initial document to measure predictive accuracy of CART, CHAID, and Exhaustive CHAID data mining algorithms in predicting the BW from various morphological traits of Kangal (Karabash) dogs, which are important gene sources for Turkey. The algorithms would be a very useful tool for dog breeders to classify the best dog taking after in the examined traits and to obtain knowledge about breed standards and morphological traits linked positively with BW for dogs. To generalize the available outcomes, further surveys should be carried out on different dog populations. The best model is CHAID algorithm. The most influential predictors are RE, CC, FSC, farm, province and RCW on BW.

The significant results of CHAID algorithm obtained in this study are summarized below:

1. The tree-based CHAID algorithm was selected as the ideal data mining algorithm.
2. Node 15, the subgroup of those having FSC > 14 cm and RE > 80 cm, yielded the heaviest average BW with 61.375 kg.
3. Node 14, the subgroup of those having FSC $\leq$ 14 cm and RE > 80 cm, had the average 50.417 kg BW.
4. An average BW of 50.091 kg was obtained by Node 19, the subgroup of those with CC > 84 cm and 72 < RE $\leq$ 74 cm reared in peasant and personal farms in the decision tree diagram.
5. An average BW of 52.100 kg was provided by Node 21, the subgroup of those having RCW > 21 cm and 74.000< RE $\leq$ 80 cm in Konya province of Turkey.
6. An average of 53.455 kg was found for Node 23, the subgroup of those having FSC > 13 cm and 74.000< RE $\leq$ 80 cm in Sivas province of Turkey.

As a result, we expect that the application of the tree-based CHAID algorithm in the BW prediction will be a respectable reference for next studies.

REFERENCES

Akin, M., E. Eyduran, and B. M. Reed (2016). Use of RSM and CHAID data mining algorithm for predicting mineral nutrition of hazelnut. Plant Cell, Tissue and Organ Culture (PCTOC), DOI 10.1007/s11240-016-1110-6.

Aksahan R. and I. Keskin (2015). Determination of the Some Body Measurements Effecting Fattening Final Live Weight of Cattle by the Regression Tree Analysis. Selcuk J. Agricultural Sciences 2(1): 53-59.

Ali, M., E., Eyduran, M. M., Tariq, C., Tirink, F., Abbas, M. A. Bajwa, M. H. Baloch, A. H. Nizamani, A. Waheed, M. A. Awan, S. H. Shah, Z. Ahmad, and S. Jan (2015). Comparison of artificial neural network and decision tree algorithms used for predicting live weight at post weaning period from some biometrical characteristics in Harnai sheep. Pakistan J. Zool. 47:1579-1585.

Atasoy, F., M. Ugurlu, B. Ozarslan and A. Yakan (2011). Body weight and measurements of Akbas dogs in its nature work condition. Ankara University, Faculty of Vet J. 58: 213-215.

Biggs D., B. De Ville and E. Suen (1991). A method of choosing multiway partitions for classification and decision trees. J. Applied Statistics 18: 49-62.

Breiman, L., J. Friedman, R. Olshen, and C. Stone (1984). Classification and Regression Trees. Chapman and Hall, Belmont, CA.

Coban O., A. Yildiz, N. Sabuncuoglu, E. Lacin and F. Yildirim (2011). Use of Non-linear Growth Curves to Describe the Body Weight Changes in Kangal Puppies. Atatürk Universitesi Vet. Bil. Derg. 6(1): 17-22.

Emehelu C. O., J. I. Eze, A. Akune and K. F. Chah (2012). Estimation of Live Body Weight From Body Measurements in Nigerian Local Dogs. African Journals Online 30(2): 65-73.

Eyduran, E., I. Keskin, Y. E. Erturk, B. Dag, A. Tatliyer, C. Tirink, R. Aksahan and M. M. Tariq (2016). Prediction of Fleece Weight from Wool Characteristics of Sheep Using Regression Tree Method (Chaid Algorithm). Pakistan J. Zool. 48(4): 957-960.

Grzesiak, W. and D. Zaborski (2012). Examples of the use of data mining methods in animal breeding. (Book) ISBN 978-953-51-0720-0.

IBM SPSS (2013). Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Khan, M. A., M. M. Tariq, E. Eyduran, A. Tatliyer, M. Rafeeq, F. Abbas, N. Rashid, M. A. Awan, and K. Javed (2014). Estimating body weight from several body measurements in Harnai sheep without multicollinearity problem. The J. Anim. Plant Sci. 24: 120-126.

Mendes, M. and E. Akkartal (2009). Regression tree analysis for predicting slaughter weight in broilers. Italian J. Anim. Sci. 8: 615-624.

Michael, J. A. and S. L. Gordon (1997). Data mining technique for marketing, sales and customer support. New York: Wiley.

Mohammad, M. T., M. Rafeeq, M. A. Bajwa, M. A. Awan, F. Abbas, A. Waheed, F. A. Bukhari, and P. Akhtar (2012). Prediction of body weight from body measurements using regression tree (RT) method for indigenous sheep breeds in Balochistan, Pakistan. The J. Anim. Plant. Sci. 22: 20-24.

Nisbet, R., J. Elder, and G. Miner (2009). Handbook of statistical analysis and data mining applications. Canada.

Orhan, H., E. Eyduran, A. Tatliyer and H. Saygici (2016). Prediction of egg weight from egg quality characteristics via ridge regression and regression tree methods. Revista Bras. Zootec. 45(7): 380-385.

Takma, Ç., H. Atil and V. Aksakal (2012). Comparison of Multiple Linear Regression and Artificial Neural Network Models Goodness of Fit to Lactation Milk Yields. Kafkas Univ Vet Fak Derg. 18(6): 941-644.

Valdez C. A. and J. P. L. Valencia (2004). Weight prediction in medium to large sized adult Philippine native dogs using external body measurements. Philippine J. Vet. Anim. Sci. 30(1): 161-169.

Yakubu, A (2012). Application of regression tree methodology in predicting the body weight of Udasheep. Anim. Sci. Biotech. 45: 484-490.

Yildirim, A. (2012). Characteristics of the Growth and Body Measurements of the Kangal Dogs. Igdir Univ. J. Inst. Sci. and Tech. 2(3): 117-126.

Yilmaz, O. (2007). Some morphological characterıstics of Kangal dogs raised in various regions of Turkey. PhD. Thesis, Ankara University Graduate School of Natural and Applied Science, Ankara.