Çanga Boğa, D. (2024). Creating a Lactation Model for 305-Day Milk Yield with Different Resampling Techniques (Bagging Mars) in Mars Modeling. *The Black Sea Journal of Sciences*, 14(2), 522-539.

The Black Sea Journal of Sciences, 14(2), 522-539, 2024. DOI: 10.31466/kfbd.1383458

KFBD

Karadeniz Fen Bilimleri Dergisi
The Black Sea Journal of Sciences
ISSN (Online): 2564-7377 https://dergipark.org.tr/tr/pub/kfbd

Araştırma Makalesi / Research Article

Creating a Lactation Model for 305-Day Milk Yield with Different Resampling Techniques (Bagging Mars) in Mars Modeling

Demet ÇANGA BOĞA[1*]

Abstract

The main purpose of this research is to obtain a prediction model for milk yield by using Multivariate Adaptive Regression Splines (MARS) and Bagging MARS algorithms as a non-parametric regression technique. For this purpose, the effects on milk yield of 305 days were investigated by using lactation parameters in dairy cattle. In the study, 9337 lactation milk yield records belonging to 37 animals belonging to the 2022-2023 period were used and the data set was created by randomly ordering the animals. Data on milk yield results were analyzed with MARS and Bagging MARS algorithms. For dairy cattle; it was modeled with explanatory variables such as lactation month (month), service period (SP), last 7 days average milk yield (L7DMMY), animal's first birth age (FP), animal's age (Age), number of lactations (LN).Correlation coefficient (r), coefficient of determination (R^2), Adjusted R^2, Root of Square Mean Error (RMSE), standard deviation ratio (SD ratio), mean absolute percent error (MAPE), mean absolute for MARS algorithm estimating total average milk yield deviation (MAD) and Akaike Information Criteria (AIC) values are 0.9986, 0.997, 0.977, 0.142, 0.052, 0.2389, 0.086 and -88, respectively. Similar statistics for the Bagging MARS algorithm are 0.754, 0.556, 0.453, 1.8, 0.666, 3.96, 1.47, and 115, respectively. It has been observed that MARS and Bagging MARS algorithms provide correct results according to the goodness of fit statistics. In this study, it was revealed that MARS algorithm gave better results in milk yield modeling of 305-day lactation.

Keywords: Lactation, Milk yield, MARS, Dairy Cattle, Bagging.

Süt Sığırlarında 305 Günlük SütVerimi için Mars Modellemesinde Farklım Yeniden Örnekleme Teknikleri(Bagging Mars) ile Laktasyon Modeli Oluşturma

Öz

Bu araştırmanın temel amacı, parametrik olmayan bir regresyon tekniği olarak Çok Değişkenli Uyarlanabilir Regresyon Splines (MARS) ve Bagging MARS algoritmalarını kullanarak süt verimi için bir tahmin modeli elde etmektir. Bu amaçla çalışmada süt sığırlarında laktasyon parametreleri kullanılarak 305 günlük sütverimi üzerine etkileri incelenmiştir. Çalışmada 37 tane hayvana ait 2022-2023 dönemine ait 9337 adet laktasyon süt verimi kaydı kullanılmış ve hayvanlar rastgele sıralanarak veri seti oluşturulmuştur. Süt verimi sonuçlarına ilişkin veriler MARS ve Bagging MARS algoritmaları ile analiz edilmiştir. Laktasyon ayı(month), Servis periyodu (SP), son 7 günlük ortalama süt verimi(L7DMMY), hayvanın ilk doğum yaşı(FP), hayvanın yaşı(Age), laktasyon sayısı(LN) gibi açıklayıcı değişkenler ile modellenmiştir.Toplam ortalama süt verimini tahmin eden MARS algoritması için korelasyon katsayısı (r), belirleme katsayısı (R^2), Düzeltilmiş R^2, Hata Kareler Ortalamasının Karekökü (RMSE), standart sapma oranı (SD oranı), ortalama mutlak yüzde hatası (MAPE), ortalama mutlak sapma (MAD) ve Akaike Bilgi Kriterleri (AIC) değerleri sırasıyla 0.9986, 0.997, 0.977, 0.142, 0.052, 0.2389, 0.086 ve -88'dir. Bagging MARS algoritması için benzer istatistikler sırasıyla 0.754, 0.556, 0.453, 1.8, 0.666, 3.96, 1.47 ve 115'tir. MARS ve Bagging MARS algoritmalarının uyum iyiliği istatistiklerine göre doğru sonuçlar ortaya koyduğu gözlemlenmiştir. Bu çalışmada, MARS algoritmasının 305 günlük laktasyona ait süt verimi modellemesinde daha iyi sonuçlar verdiği ortaya çıkmıştır.

Anahtar Kelimeler: Laktasyon, Süt verimi, MARS, Süt Sığırı, Bagging.

[1]Niğde Ömer Halisdemir University, Faculty of Economics and Administrative Sciences, Department of Business Administration Department of Numerical Methods, Niğde, Türkiye, demetcangaboga@ohu.edu.tr

[*]Sorumlu Yazar/Corresponding Author
Geliş/Received: 30.10.2023 **Kabul/Accepted:** 28.04.2024 **Yayın/Published:** 18.06.2024

1. Introduction

It is important to select animals with high milk yield in order to reach milk production in developed countries and to meet the milk needs of individuals in our country. The Holstein breed is very important for this purpose. Producers in Türkiye prefer farm breeds with higher meat and milk yields than local breeds, and the share of these breeds in the total cattle population is increasing every year. While 18.94% of the total number of Turkish cattle was 9.8 million in 2002, breeding, 44.45% crossbreed and 36.58% domestic cattle breeds, the total number of cattle in 2022 is approximately 17 million. The proportions of breeding cattle crossbred cattle and domestic cattle breeds in the total are 49.2%, 43.4% and 7.3%, respectively (TUIK, 2023).

Lactation period, on the other hand, is the period from starting to ending of the millking. The average lactation period of dairy cattle is 305 days. This period may vary depending on care and feeding (Özyurt and Özkan, 2009).It is imperative to take into account their reproductive performance, which affects the productivity of dairy cows and has a major impact on the profitability of dairy production establishments. To this end; In many studies, calving interval, pregnancy rate, insemination number for pregnancy, service period, first insemination and first calving age have been used to evaluate reproductive performance (Bayril and Yilmaz, 2017; Boğa and Boğa, 2022; Çanga and Boğa, 2022; Doğan, 2003; Mee, 2004; Şahin and Ulutaş, 2010). Lactation curves can be used to calculate lactation continuity, which indicates the animals' ability to maintain a constant milk yield. Various methods have been proposed for calculating lactation persistence, but there is still no standard method. In studies conducted in different years on the average milk yield of 305 days in Holstein cows in Türkiye, varying results have been reported over the years (Gürses and Bayraktar, 2012; Omar, 2022; Orhan, Çetin Teke, and Karcı, 2018; Şahin and Ulutaş, 2010; Sarar and Tapkı, 2017; Yaylak and Kumlu, 2005).

However, many researches made have been done on milk production modeling and various models have been developed to explain the relationship between control day and milk production (Bayril and Yilmaz, 2017; Eyduran, Yilmaz, Tariq, and Kaygisiz, 2013; Orhan et al., 2018). MARS (Multivariate Adaptive Regression Spline) decomposes multivariate nonlinear models and offers the ability to be explained by linear models (Grzesiak et al., 2010),described cows inseminated using statistical and machine learning methods (classification functions, logistic regression, neural networks and MARS). It was determined that the best results were obtained with artificial neural networks (ANN) and MARS methods. MARS, which will be used in this research, decomposes multivariate nonlinear models and offers the opportunity to be explained with linear models (Çanga, 2022; Faraz et al., 2021). Faraz et al., (2021), in their study, modeled with MARS, which is a non-linear method to estimate the body weight (BW, kg) of bovine animals. It has been reported that models with high predictive power were obtained with the sheep's body length (BL, cm) and chest circumference (CG, cm) as explanatory variables with the model they presented. In this study, MARS data mining algorithm, which performs the best estimation among statistical methods, was used. The MARS data mining algorithm, which is an embodiment of the CART (Classification and Regression Tree) algorithm, does not require any assumptions about the distribution of the variables. However, the fact that it does not require a functional hypothesis between dependent and independent variables distinguishes this algorithm from other previously used algorithms (Turhan, 2020; Iqbal et al., 2021).

Re-sampling clustering (Bagging=Bootstrap+Aggreating) is an effective method specified by various classifications and regressions to improve prediction accuracy in large data sets. Resampling clustering can be used as a tool to reduce the variance of a predictor and increase the stability and power of predictions (Çanga and Boğa, 2020; Celik and Yilmaz, 2021; Otok, Putra, Sutikno, and Yasmirullah, 2020; Şengül et al., 2022). When the literature is reviewed, many researches have been done on milk production modeling and various models have been developed to explain the relationship between control day and milk production (Alıç, 2007; Bilgiç, 1999; Çetin and Alkoyak, 2018; Doğan, 2003; Orhan and Kaygısız, 2002; Şahin and Ulutaş, 2010).

In this research, it is aimed to predict the future milk yields with the results obtained by making Bagging and MARS separately, with models with less errors. Milk yield was estimated by MARS method by using independent variables of lactation period, service period, average milk yield for the last 7 days, age at first birth, age of animal, number of lactations in dairy cattle in Holstein cows.

2. Materials and Methods

The research material was created from Holstein dairy cattle raised in a cattle farm in Niğde Bor district. There are approximately 100 cattle in the collection in the enterprise, and the primary target in the enterprise is milk production, and meat production is also contributed by fattening male animals in the region. For this reason, milk production, breeding heifer breeding and fattening activities are carried out in the enterprise. A total of 9561 milk data, 37 of which were milkers, were used.

2.1. Animal Material

The research material includes the 305-day milk yield of 37 Holstein dairy cattle calving between 2022 and 2023 in a cattle farm in Niğde-Bor district. Lactation month (month), service period (SP), last 7 days average milk yield (L7DMMY), age at first birth (FP), age of animal (Age), number of lactations (LN) were taken as basis in fertility records.

2.2. Statistical Analysis

Data on lactation results obtained from dairy cattle were analyzed with MARS and Bagging MARS algorithms. In the statistical modeling studies conducted in the last ten years, the prediction performances of decision trees, artificial neural networks and MARS algorithms have been comparatively examined. The MARS algorithm, which is a modified version of the CART algorithm, has been the focus of attention of researchers in terms of defining the high-level relationships between the variables under consideration (Akin et al., 2020; Altay et al., 2022; Çanga, 2022; Çanga and Boğa, 2022; Tirink et al., 2022).

The MARS model was developed using the R software "earth" package and "ehaGoF" (Milborrow 2011; R Core Team 2024; Milborrow 2018; Eyduran et al., 2019; Eyduran and Gulbe 2020; Eyduran and Duman 2020; Tırınk et al., 2023). Among the observed and predicted values of in the study, the smallest GCV, SDRATIO, RMSE, MAPE, MAD, AIC, AICc and the MARS model with the highest determination coefficient (R^2) and Pearson coefficient (r) were accepted as the best.

2.2.1. Multivariate Adaptive Regression Spline (MARS)

MARS algorithm used by Friedman (1991) to capture nonlinear relationships between predictors and response variable(s) is a powerful approach that does not require assumptions about functional relationships between dependent and input variables. The model that emerges as the weighted total basic function including the BFi(x) function is given by Equation 1 below (Akin et al., 2020; Eyduran et al., 2020; Çanga 2022; Çanga and Boğa 2020; Çelik et al., 2021).

$$y = \sum_{\mathrm{i=1}}^{\mathrm{k}} a_i \mathrm{BFi(x)} \qquad (1)$$

Mars algorithm is formed by the linear breakdown of the basic function of BFi (x) with the following equation.

$$BF_1 = \max(0, x - t) \begin{Bmatrix} x - t, x > t \\ 0, x \leq t \end{Bmatrix} \qquad (2a)$$

$$BF_2 = \max(0, x - t) \begin{Bmatrix} x - t, x > t \\ 0, x \leq t \end{Bmatrix} \qquad (2b)$$

Here, x is the variable range; t is the node. The linear combination of the basic functions obtained accordingly:

$$Y_i = a_0 + a_1 BF_1 + a_2 BF_2 + \cdots + a_k BF_k$$

and the estimation equation is obtained. Here Y_i dependent variable, a_0 intercept, and $a_1, \dots, a_k$ are coefficients of the related basic functions (Everingham and Sexton 2011, Emamgolizadeh et al., 2015; Çanga and Boğa 2021; Çanga et al., 2021; Akin et al., 2020; Akın et al., 2021).

The fact that it does not require assumptions not only about the distribution of the variables but also about the functional relationships between the variables, as well as providing an equation that shows the high-dimensional relationships between the variables makes the MARS algorithm very popular. In the MARS method, generalized cross-validation error (GCV) is the best criterion for choosing the best model. Because GCV takes into account both errors and model complexity (Grzesiak et al., 2010).

Bagging a method first introduced by (Breiman, 1994) is used to increase the stability and power of the estimator by reducing the variance of an estimator. The Bagging MARS algorithm is based on creating the desired number of Bootstrap samples based on the original data set (Çanga and Boğa, 2020; Kulekçi et al., 2022). Many generated examples are derived from datasets. If changes in the dataset cause significant changes, bagging can increase sensitivity. So the basic idea of Bagging is to use resembling to create an estimator with multiple versions; here, after combining the objectives, the result should be better than a single predictive index created to solve the same problem. The bootstrap example is the training set, and the group of observations that are not included in the bootstrap example serves as the test set (Kulekçi et al., 2022). In order to talk about the reliability of the established MARS model, it is important that the generalization ability is good. However, MARS analysis can be performed using Bootstrap, one of the resampling methods (Akin et al., 2020).

2.2.2. A MARS Model Application

Earth and ehaGoF packages were used for the MARS model in the study. With the earth package, the same basic functions are generated for MARS prediction models created simultaneously for more than one dependent variable. As can be seen, MARS models produced with the *earth* package have different coefficients. To make this estimate, the Generalized cross-validation (GCV) method, a computational solution for linear models that provide an estimated exclusion cross-validation error metric, is used. According to the GCV criterion, MARS generalizes the model by eliminating the terms. GCV is given by Equation 2, a form of regulation that balances model complexity with the goodness of fit (Eyduran et al., 2019; Akin et al., 2020; Akin et al., 2021; Çanga et al., 2021)

$$\mathrm{GCV} = \sum_{i=1}^{N} \frac{(y_i - \hat{y})^2}{\left(1 + \frac{C}{N}\right)^2} \tag{3}$$

Here, C = 1 + cd, is the number of items in the N dataset; d is a degree of freedom; c is the basic function addition penalty. y_i Is an independent variable and $\hat{y}y_i$is an estimated value (Eyduran 2020; Akin et al., 2020).

2.2.3. Model validity:

The most common model fit criteria to be used in measuring the predictive accuracy of the MARS algorithm (Goodness of Fit Criteria) are the goodness of fit criteria such as R-square, RMSE and MAE mentioned below (Akin et al., 2020; Eyduran et al., 2019; Eyduran and Zaborski, 2017; Çanga et al., 2021; Nayana,et al., Chesneau, 2022). The model was evaluated according to these values.

Determination coefficient (R^2):

It is the percentage of the total variation in the response variable explained by the regression line. The equation is expressed by X.

$$\mathrm{R}^2 = 1 - \frac{\mathrm{SSE}}{\mathrm{SST}} \qquad \text{(a)}$$

Where SSE $= (\mathrm{yi} - \hat{y})^2$ is the sum of the squares of the differences between the predicted and the observed value, and SSE $= (y_i - \bar{y})^2$is the sum of the squares of the differences between the observed and the overall average value.

The adjusted determination coefficient ($Adj.R^2$) is calculated with the formula Equation 4b.

$$Adj.R^2 = 1 - \frac{R^2(n-1)}{n-k-1} \qquad \text{(4b)}$$

It is preferred to be close to the R^2 value. Average square error (RMSE), average estimation error (is the square root of the average square error). The formula is stated as follows:

$$\mathrm{RMSE} = \sqrt{\sum_{\mathrm{i=1}}^{\mathrm{n}} (y_i - \hat{\mathrm{y}})^2} \qquad (5)$$

Average error (ME) is the average estimation error. It is less sensitive to outliers. It is given by the formula as follows

$$\mathrm{ME} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i=1}}^{\mathrm{n}} (y_i - \hat{\mathrm{y}}) \qquad (6)$$

Mean absolute deviation (MAD) is the mean absolute estimate error. It is less sensitive to outliers. The formula is given as follows:

$$\mathrm{MAD} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}} |y_i - \hat{\mathrm{y}}| \tag{7}$$

Pearson correlation coefficient between actual values and estimated values in terms of a dependent variable (r),

$$\mathrm{PC} = \mathrm{r}_{\mathrm{y_i}\hat{\mathrm{y}}} = \frac{\mathrm{Cov}(\mathrm{y_i}, \hat{\mathrm{y}})}{\mathrm{S_{y_i}} \mathrm{S}_{\hat{\mathrm{y}}}} \tag{8}$$

$\mathrm{Cov}(\mathrm{y_i}, \hat{\mathrm{y}})$: The covariance between actual and predicted values in terms of a dependent variable,

S_{y_i} : The standard deviation of the actual values of the dependent variable and

$S_{\hat{\mathrm{y}}}$: It refers to the standard deviation of the mined values of the dependent variable. Akaike information criterion (AIC)

$$\mathrm{AIC} = \mathrm{n}\ln\left[\frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}} (y_i - \hat{\mathrm{y}})^2\right] + 2\mathrm{k};\ \ \mathrm{Eğer}\frac{\mathrm{n}}{\mathrm{k}} > 4 \tag{9a}$$

$$AIC_C = \mathrm{n}\ln\left[\frac{1}{\mathrm{n}}\sum_{\mathrm{i=1}}^{\mathrm{n}} (y_i - \hat{\mathrm{y}})^2\right] + 2\mathrm{k} + \frac{2\mathrm{k}(\mathrm{k}+1)}{\mathrm{n}-\mathrm{k}-1};\ otherwise \tag{9b}$$

Standard deviation ratio (SD_{ratio}):

$$SDratio = \frac{\mathrm{S_m}}{\mathrm{S_d}} \tag{10}$$

$\mathrm{S_m}$: Standard deviation of model error terms,

$\mathrm{S_d}$: The standard deviation of the dependent variable,

The standard deviation ratio is calculated with the SD ratio formula, and values less than 0.20 are preferred for biometric studies. The smaller it is, the more acceptable it is (Eyduran et al. 2017).:

$\widehat{F_e}$ (x) being the Mars prediction model; The bagging MARS model is as expressed in equation 10 (Breiman, 1994):

$$\hat{\mathrm{f}}_{\mathrm{Bagging\ MARS}} = \frac{1}{\mathrm{E}}\sum_{\mathrm{i=1}}^{\mathrm{E}} \widehat{\mathrm{F_e}}\,(\mathrm{x}) \tag{ 10}$$

3. Findings and Discussion

3.1. Findings

Introductory statistics of continuous variables related to lactation yield used in the study are given in Table 1.

Table 1. Descriptive statistics of the studied explanatory variables.

	Month	SP	Total_yield	L7DMMY	LN	Age	FPA
Vars	1	2	3	4	5	6	7
n	9553	9552	9552	9533	9452	9312	9312
Mean	6.72	211.69	30.87	31.26	2.68	56.99	23.30
Standart Deviation	3.44	150.29	9.77	8.39	1.02	14.18	1.76
Median	7.00	193.00	30.40	30.91	3.00	65.00	23.00
MAD	4.45	149.74	8.64	7.12	1.48	2.97	1.48
Min	1	0	8.77	9.01	1.00	24.00	20.00
Max	12	691.00	79.18	116.78	5.00	71.00	27.00
Standard Error	0.04	1.54	0.10	0.09	0.01	0.15	0.02

Total yield: Total average milk yield of lactation; SP: Service Period, L7DMMY: mean milk yield for the last 7 days; LN: number of lactations; Age: Age of the animal, FPA: Age at first birth of the animal.

Introductory statistics related to independent variables such as milk yield and lactation period (month), age at first birth (FPA), age of animal (Age), number of lactations (LN) used in the study are given in Table2, Table3, Table4, respectively.

Table2. Descriptive statistics of monthly average milk yield and number of lactations

LN/Total yield	Mean± SE	N	Minimum	Maximum
1	24.43±0.249	1235	5	62
2	33.26±0.176	2885	1	71
3	28.98±0.166	3618	4	79
4	34.96±0.238	1383	3	68
5	31.99±0.248	432	6	66
Total	30.63±0.102	9653	1	79

In the study, a sample R script file was created for the data set. The data set of the studied lactation was first divided into two subsets, 70% training and 30% testing. Then, 50 bootstrap samples were created with the resampling method to obtain 3 MARS models using different sampling methods for the training set in order to determine the appropriate number of terms and the degree of interaction(Akin, Eyduran, and Eyduran, 2020).

The ehaGoF package was used to measure the prediction performance and generalization capabilities of the established models (Eyduran,2020).The data set obtained from approximately 10,000 data used in the research is important in terms of the reliability of the results obtained, working on large samples in data mining studies. Goodness of fit statistics calculated for MARS and Bagging MARS algorithms are given in Table 3. The prediction performances of MARS and Bagging MARS in predicting Total_MY were evaluated comparatively. The results of the fit criteria are summarized in Table 3.

Table 3. Predictive performance of MARS and Bagging Algorithms

Methods	r	R^2	Adj.R^2	RMSE	SDratio	MAPE	MAD	AIC
MARS	0.988	0.968	0.968	10.204	0.178	1.374	4.846	143.073
Bagging MARS	0.762	0.436	0.435	7.364	0.751	4.515	5.405	735.927

RMSE: Root-mean-square error, **SD ratio:** Standard deviation ratio, **MAPE:** Mean absolute percentage error, **MAD:** Mean absolute deviation, **AIC:** Akaike Information Criteria

When Table 3 is examined, the order of superiority in the prediction accuracy of the mentioned algorithms is MARS > Bagging MARS according to the estimated model evaluation criteria. As can be seen, the prediction performance of the MARS algorithm was found to be better than Bagging MARS. The results of the MARS algorithm for cattle are presented in Table 4.

The model equation of the MARS algorithm is as follows. Total_yield = 41.87 + 4.016 * label13 + 3.385 * label21- 6.265 * label4 + 2.017 * label40 + 2.8 * label45 - 1.392 * Monthjune_22 + 0.02707 * max (0, 75 - SP) - 0.0177 * max(0, SP - 75) + 0.02061 * max(0, SP - 417)- 0.2831 * max (0, L7DMMY - 32.32) - 0.7361 * max(0, 41.5 - L7DMMY) + 0.4114 * max(0, L7DMMY - 41.5) - 0.1251 * max(0, 66 - Age) - 0.4353 * max(0, Age - 66) + 0.9208 * max(0, 22 - FPA) + 0.3967 * max(0, FPA - 22)

Table 4. Coefficients of the MARS model and results of MARS analysis

Terms	Basis Functions (BF_i)	Coefficients
	Intercept	41.87
BF1	Label13	4.016
BF2	Label21	3.385
BF3	Label4	- 6.265
BF4	Label40	2.017
BF5	Label45	2.8
BF6	Monthjune_22	-1.392
BF7	max(0, 75 - SP)	0.02707
BF8	max(0, SP - 75)	-0.0177
BF9	max(0, SP - 417)	0.02061
BF10	max(0, 41.5 - L7DMMY)	-0.7361
BF11	max(0, L7DMMY - 41.5)	0.4114
BF12	max(0, 66- Age)	-0.1251
BF13	max(0, Age - 66)	-0.4353
BF14	max(0, 22 -FPA)	0.9208
BF15	max(0, FPA - 22)	0.3967

Among the independent variables, the most important and highest positive effects are the variables belonging to Label13, Label21, Label40 cattle, respectively. When the variable numbered Label 13 is selected, that is, label 13=1, an increase of 4.016 units is expected in the total milk yield (Total_MY)(Table3). Similarly, when Table 4 is examined; When the variable label 4 is selected, that is, when Label 4=1, a decrease of 6.265 units is expected in the total milk yield (Total_MY).In addition, when the seventh term max(0, 75 - SP) and positive coefficient(0.02707) for Total_MY of the service period (SP) are examined; For a variable with SP= 75 months or SP > 75, the positive effect of the service period (due to the coefficient) on Total_MY will be masked. However, for a variable with SP<75, the time to last calving (SP) is expected to have an increasing effect (due to the coefficient) on Total_MY. Although the SP value is 1 unit less than 75 months, an increase of 0.02707 liters is expected in the Total_MY amount. The relative importance values of the independent variables are presented in Table 5. As seen in Table 5, the highest importance values are respectively; L7DMMY (100%), SP (28.6%), Label13 (22.8%), Age (20.3%), label21 (15.3%), Label4 (12.3%), Monthjune_22 (12.3%), Label4 (12.3%), Label4 (Obtained for 9.6%, label40 (8.3%), label45 (7.1%), FPA (5.3%)

Table 5. Relative importance of independent variables in the model

Variables	Nsubsets	GCV	RSS
L7DMMY	16	100.00	100.00
SP	15	28.6	29.9
Label13	14	22.8	24.3
Age	13	20.3	21.9
label21	11	15.3	17.1
label4	9	12.3	14.1
Monthjune_22	6	9.6	11.2
label40	5	8.3	9.8
label45	4	7.1	8.5
FPA	3	5.3	6.7

(Akin, Eyduran, Eyduran, and Reed, 2020)and some previous authors (Boğa and Boğa, 2022; Çanga and Boga, 2019; Çanga and Boğa, 2022; Çelik et al., 2021; Turhan, 2020) and, as highlighted by the results of the current study, predicted that MARS captures linear, nonlinear and interaction effects of important factors in regression-type problems..This difference can be attributed to the overfitting problem that may occur in MARS. The problem is that the MARS model contains redundant terms that reduce the predictive quality. Therefore, redundancies should be removed from the MARS model, which has the maximum complexity in the forward transition, by using the backward pruning method (Eyduran et al., 2019).The best way to understand the overfitting problem is to estimate the R^2 values for Cross-validation, training and test sets (Grzesiak and Zaborski, 2012),emphasized that if the standard ratio value of an established regression model is 0.10 or 0.40, there is good or very good fit. When Table 6 is examined, all coefficients for milk yield were found to be statistically very significant.

Tablo 6. Results of the MARS algorithm

Coefficients:	***Estimate Std***	***Std. Error***	***t value***	***Pr(>|t|)***
(Intercept)	41.872269	0.347857	120.372	< 2e-16 ***
bx[, -1]max (L7DMMY-41.5014)	0.411366	0.062027	6.632	3.50e-11 ***
bx[, -1]max (41.5014-L7DMMY)	-0.736077	0.024382	-30.189	< 2e-16 ***
bx[, -1]max (SP-75)	-0.017704	0.001196	-14.800	< 2e-16 ***
bx[, -1]max (75-SP)	0.027067	0.005402	5.010	5.53e-07 ***
bx[, -1] max(Age-66)	-0.435297	0.107400	-4.053	5.10e-05 ***
bx[, -1] max(66-Age)	-0.125130	0.010561	-11.848	< 2e-16 ***
bx[, -1] label13	4.016018	0.453965	8.847	< 2e-16 ***
bx[, -1]Label21	3.385252	0.471059	7.186	7.16e-13 ***
bx[, -1]max (SP-417)	0.020615	0.002866	7.192	6.88e-13***
bx[, -1] label4	-6.265041	0.939251	-6.670	2.70e-11***
bx[, -1]max (FPA-22)	0.396662	0.087689	4.524	6.16e-06 ***
bx[, -1] max(22-FPA)	-1.391727	0.195930	4.700	2.64e-06***
bx[, -1] Monthjune_22	-1.922e+01	0.263221	-5.287	1.27e-07***
bx[, -1] max(L7DMMY-32.3214)	-0.283084	0.049394	-5.731	1.03e-08***

bx[, -1]label45	2.799577	0.557498	5.022	5.22e-07 ***
bx[, -1]label40	2.016650	0.472066	4.272	1.96e-05 ***

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The prediction equation of the Bagging MARS algorithm is as follows:

34.07 -2.30 * label11 + 2.39* label13 - 5.68 * label37 - 6.63 * label4 + 2.04* label40 - 3.33 * label43 - 1.78* LN3 + 0.01 * h(410-SP) - 0.60 * h(39.31-L7DMMY) + 0.24* h(L7DMMY-39.31) - 0.11 * h(66-Age)

+43.83+ 2.35* label13 - 6.63 * label37 - 8.32 * label4 +2.70 * label45+ 1.44 * Monthfeb_23 + 2.26 * Monthmarch_23 + 1.60 * LN2 + 0.03 * h(71-SP) - 0.02 * h(SP-71) + 0.02 * h(SP-384) - 0.28*h(L7DMMY-33.35)+11.67 * h(L7DMMY-45.03) -0.69 * h(45.30-L7DMMY) -11.65 * h(L7DMMY-45.30) + 0.765 * h(L7DMMY-78.55) - 0.12 * h(66-Age)

+ 37.69+ 2.94 * label13 + 3.39 * label21 + 4.20 * label24 -8.21 * label4 + 2.13 * label40 +5.70 * label45 - 2.32 * label9 - 2.09 * Monthjune_22 - 2.08 * Monthmay_22 - 1.13* Monthsept_22 + 1.22* LN2 + 0.02 * h(379-SP) - 0.02 * h(SP-379) - 0.28* h(L7DMMY-34.18) - 1.22 * h(L7DMMY-41.02) + 7.31 * h(L7DMMY-44.53)- 0.64 * h(45.64-L7DMMY) - 6.08 * h(L7DMMY-45.64) + 0.76 * h(L7DMMY-77.37) - 0.12 * h(66-Age)) / 3

According to this obtained equation, in the first bootstrap equation, there is a negative decrease of 1.78266 units when lactation order = 3. The highest negative effect (- 6.634773) was caused by the variable number 4 on the label, while an increase of 0.0168345 units was observed when SP<410, while a negative effect of -0.1068442 and -0.5982774 units was observed when L7DMMY<39.31, respectively. In the second bootstrap equation, Monthfeb_23 (1.444099), Monthmarch_23 (2.268667), LN2(1.603694) contribute positively, while minor reduction is expected when Age≤ 66 in all three bootstrap equations. The relationship between the predicted and observed predictive values of the Bagging MARS algorithm is shown in Figure 1.

In the Bagging MARS model, there is a binary interaction between the variables.

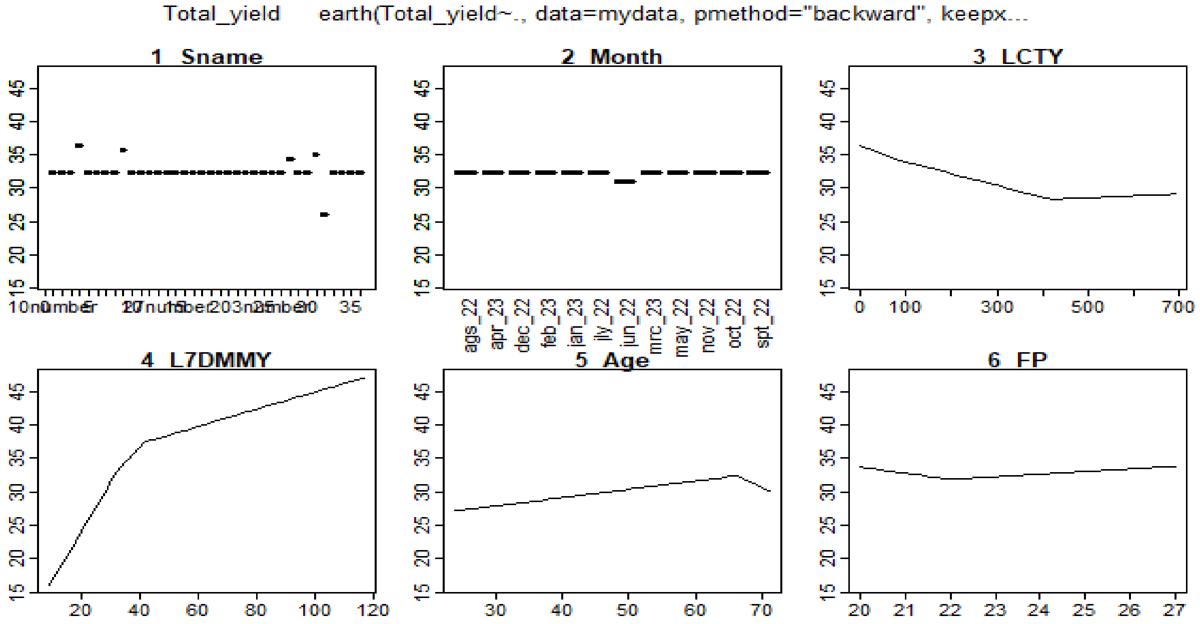

Figure 1. Relationship graphs between dependent variable and independent variables in Bagging Mars

Model selection, cumulative distribution of absolute errors, error values, distribution of predicted values and error QQ graphs of normal distribution of error values are given below for the MARS model selected according to cross validation instead of GCV(Figure2).

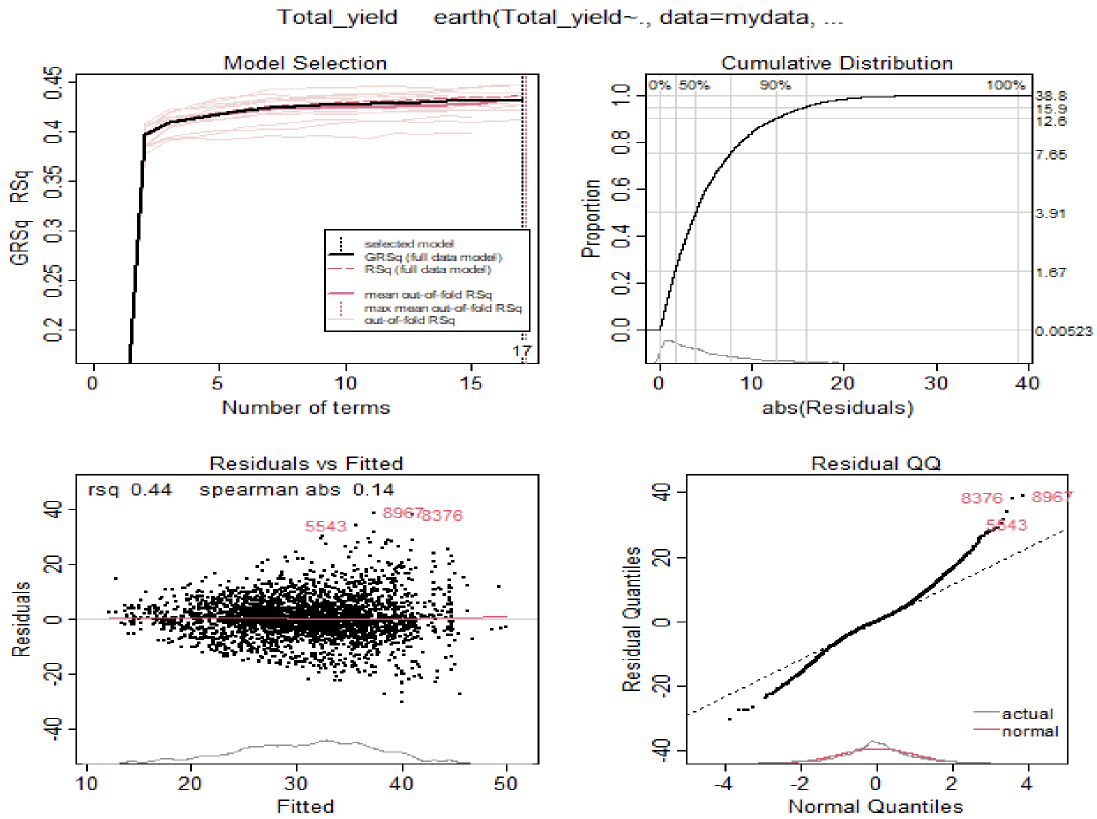

Figure2. Bagging MARS Model selectionand graphs of error terms

3.2. Discussions

In the study, the effect of 4 different parameters (average milk yield in the last 7 days, service period, animal number 13, Age) on milk yield in cattle was found to be significant (Table 5).In this study, when the reproductive characteristics of cattle are examined; For the MARS model estimation equation for optimum milk yield, the corresponding optimum values for the independent variables Label, Month, SP, L7DMMY, LN, Age, FPA, respectively; 20, feb_23, 194(day), 30,96286(lt), 3, 65(month), 23(month). These threshold values, which are created by the multi-response MARS model equation that constitute the basic functions, are the most common(Akin et al., 2020; Akin et al., 2020).

First gestational age, starting the productive life of heifers as early as possible is one of the factors that directly affect the profitability of the enterprises and this value varies according to the breeds. In this study, when we subtract the gestational age from the average gestational age (FPA), which was calculated as 23 months for the Holstein breed, the gestational dates for the first of the animals can be determined as an average of 14 months. These values are in agreement with the averages reported by the mean for the same breed (16.5, 18.3 months, and 16.8 months, 14.9, respectively (Alkoyak, 2016; Asan, 2021; Çetin and Alkoyak, 2018; Eyduran et al., 2008; Omar, 2022).The found service period obtained in this study was found to be 123 days and this value was 127 days in the study conducted by (Gürses and Bayraktar, 2012). Gürses and Bayraktar (2012) in TİGEM Ceylanpınar, Dalaman, Koçaş and Tahirova Agricultural Enterprises, on the same breed, while it was 127 days in Aydın province In the study conducted by the company Gürses (2019), (Gürses, 2019) it is shorter as 144 days; In the study conducted by Sarar and Tapkı (2017) (Sarar and Tapkı, 2017) in Koçaş Agricultural Enterprise, 106 days; In a study conducted by Arı (2019) in a private enterprise in Aydın province, 109 days, (Arı, 2019, Kopuzlu et al., 2008). Kopuzlu et al. (2008) in the Eastern Anatolia Agricultural Research Institute, it was found to be 119.9 days, (Omar, 2022) as 73.19 days by Omar (2022). Omar (2022) which is higher than the averages reported in the same breed.

Goodness-of-fit statistics are important in comparing data mining and other statistical methods used to predict any trait in cattle as well as in all animals. In recent years, many studieshave used MAE, mean square error (MSE), AdjR^2 to compare artificial neural network, MARS, correlation coefficient (r), RMSE and mean absolute models. and Akaike information criteria (AIC) statistics were used (Akin et al., 2020; Altay et al., 2022; Boğa and Boğa, 2022; Çanga, 2022; Çanga and Boga, 2019; Çanga and Boğa, 2022; Eyduran et al., 2019; Eyduran et al., 2020; Kulekçi et al., 2022; Şengül et al., 2022; Tirink et al., 2022).Although the goodness of fit statistics used by the authors in their studies are similar, the conditions for using the MARS method and the calculation conditions, correlation coefficient and RMSE statistics differ. In terms of the results obtained with other methods and several different goodness-of-fit statistics used, the dependent variable was estimated (Şengül et al., 2022). Şengül et al. (2022) reported that the final body weights of Kıvırcık lambs were evaluated on the basis of MARS and Bagging MARS algorithms, and showed excellent performance as a robust algorithm without overfitting (Akin, Eyduran, Eyduran, et al., 2020). Akin et al. (2020) predicted that MARS captures linear, nonlinear and interaction effects of important factors in regression-type problems, as emphasized by some previous authors (Eyduran et al., 2020; Öztürk, 2022) and those obtained in the current study.

In the current study, the average milk yield according to the lactation order in Table 2 varies between 24.43 and 34.96 (Omar, 2022). In the study of Omar (2022) the milk yield of Holstein and Simmental cattle was calculated as 305 days milk yield as 9690.02. When this result is divided by 305 days to calculate the average daily milk yield, it is 9690.02/305= 31.71, which is consistent with this value. In different studies conducted between 2005-2021(Asan, 2021; Çetin and Alkoyak, 2018; Genç and Soysal, 2018; Keskin and Boztepe, 2011; Şahin and Ulutaş, 2010; Sarar and Tapkı, 2017; Tırınk, 2021; Yaylak and Kumlu, 2005)was reported to vary between 5395.11-8264.70 kg and these values are among the values of the current study.

4. Conclusions and Recommendations

Considering the results of this research, it was revealed that MARS algorithm gave better results in milk yield modeling of 305-day lactation. Many of the generated samples are derived from datasets, increasing the bagging precision. Thus, it is possible to compare the goodness of fit criteria by using resampling to create an estimator with more than one version. It is expected that this study will inspire data mining research that want to discuss the results using MARS and Bagging MARS algorithms in the future.

Acknowledgements

The author would like to thank Dr. Mustafa Boğa for his animal experiments and his contributions to making measurements during the experiment.

Funding Support

There is no funding source.

Competing of Interest

The author declared no competing interests.

Author Contributions

DÇB conceived and supervised the study. DÇB statistical analysis and writing the manuscript. DÇB contributed to the critical revision of the manuscript and have read and approved the final version.

References

Akin, M., Eyduran, S. P., and Eyduran, E. (2020). *R Yazılımı ile Tarım Bilimlerinde Regresyon ve Sınıflandırma Tipi Problemlerin Çözümünde Mars Algoritması* (1st Editio). Nobel Academic Publishing.

Akin, M., Eyduran, S. P., Eyduran, E., and Reed, B. M. (2020). Analysis of macro nutrient related growth responses using multivariate adaptive regression splines. *Plant Cell, Tissue and Organ Culture*, *140*(3), 661–670. https://doi.org/10.1007/S11240-019-01763-8

Alıç, D. (2007). *Siyah Alaca İneklerde Dış Yapı Özellikleri, Sürüde Kalma Süresi ve Süt Verimi Üzerine Araştırmalar*. Ankara University.

Alkoyak, K. (2016). *Farklı orijinli holştaynların döl ve süt verimi özellikleri*. Selçuk Üniversitesi.

Altay, Y., Aytekin, İ., and Eyduran, E. (2022). *Use of Multivariate Adaptive Regression Splines , Classification Tree and Roc Curve in Diagnosis of Subclinical Mastitis in Dairy Cattle Journal of the Hellenic Veterinary Medical Society*. *April*. https://doi.org/10.12681/jhvms.25864

Arı, Ç. (2019). *Aydın İlinde özel bir süt sığırı işletmesinde yetiştirilen Simmental ve Kirmizi Alaca sığırların süt ve döl verimi ile süt kalite özellikleri.* Aydın Adnan Menderes Üniversitesi.

Asan, H. (2021). *Burdur' da pard programı kapsamında bulunan holştayn ve simental işletmelerinin verimler ve karlılık bakımından değerlendirilmesi*. Mehmet Akif Ersoy Üniversitesi.

Bayril, T., and Yilmaz, O. (2017). Holştayn Sütçü İneklerde Süt Verim Performanslarına Buzağı Cinsiyeti , Servis Periyodu , Doğum Sayısı ve Buzağılama Mevsiminin Etkisi. *Dicle Üniversitesi Veteriner Fakültesi Dergisi*, 89–94.

Bilgiç, N. (1999). *Ankara Üniversitesi Ziraat Fakültesi Zootekni Bölümü S ığı rc ı l ı k i ş letmesinde Yeti ş tirilen Siyah Alaca İ neklerde Baz ı Süt ve Döl Verimi Özellikleri Some Milk and Breeding Characteristics of Hoistein Friesian Cows Raised at the Dairy Farm of the*. *5*(2), 81–84.

Boğa, D. Ç., and Boğa, M. (2022). Investıgatıon of Estımatıon of Lactatıon Mılk Yıeld in Daıry Cattle Usıng Nonlınear Regressıon Models Cart And Mars. In K. N. ÇELİK Şenol and F. ÇEMREK (Eds.), *Veri madenciliği ve makine öğrenmesi ile farkli alanlarda uygulamalar* (pp. 73–98). Holıstence Publıcatıons.

Breiman, L. (1994). *Bagging Predictor*.

Çanga, D. (2022). Use of Mars Data Mining Algorithm Based on Training and Test Sets in Determining Carcass Weight of Cattle in Different Breeds. *Tarım Bilimleri Dergisi*, *28*(2), 259–268. https://doi.org/10.15832/ankutbd.818397

Çanga, D., and Boga, M. (2019). Use of MARS in livestock and An application. *Use of MARS in Livestock and an Application*, 31–37.

Çanga, D., and Boğa, M. (2020). Determination of the Effect of Some Properties on Egg Yield with Regression Analysis Met-hod Bagging Mars and R Application Yumurta Verimi Üzerine Bazı Özelliklerin Etkisinin Regresyon Analiz Yöntemlerinden Bagging Mars ile Belirlenmesi ve R Uygulaması. *Turkish Journal of Agriculture - Food Science and Technology*, *8*(8), 1705–1712.

Çanga, D., and Boğa, M. (2022). Detection of correct pregnancy status in lactating dairy cattle using MARS data mining algorithm. *Turkish Journal of Veterinary and Animal Sciences*, *46*(6), 809–819.

https://doi.org/10.55730/1300-0128.4257

Çanga, D., Yavuz, E. and Efe, E. (2021). Prediction of Egg Weight Using MARS data mining Algorithm through R. Kahramanmaraş Sütçü İmam Üniversitesi Tarım ve Doğa Dergisi, 24 (1) , 242-251 . DOI: 10.18016/ksutarimdoga.vi.716880

Çelik, Ş., Eyduran, E., Şengül, A. Y., and Şengül, T. (2021). Relationship among egg quality traits in Japanese quails and prediction of egg weight and color using data mining algorithms. *Tropical Animal Health and Production*, *53*(3). https://doi.org/10.1007/s11250-021-02811-2

Celik, Ş., and Yilmaz, O. (2021). The Relationship Between the Coat Colors of Kars Shepherd Dog and its Morphological Characteristics Using Some Data Mining Methods. *International Journal of Livestock Research*, *11*(1), 53–61. https://doi.org/10.5455/ijlr.20200604

Çetin, O., and Alkoyak, K. (2018). Farklı orijinli holştaynların döl ve süt verimi özellikleri 2. süt verimi özellikleri. *Eurasian Journal of Veterinary Sciences*, *34*(2), 123–130. https://doi.org/10.15312/eurasianjvetsci.2018.191

Doğan, İ. (2003). Holştayn ırkı ineklerde süt verimine etki eden faktörlerin CHAID analizi ile incelenmesi. *Ankara Üniversitesi Veteriner Fakültesi Dergisi*, *50*(1), 065–070. https://doi.org/10.1501/vetfak_0000002231

Eyduran, E. (2020). *Package ' ehaGoF . '* https://cran.r-project.org/package=ehaGoF.

Eyduran, E. and Gulbe, A. (2020). ehaGoF: Calculates Goodness of Fit Statistics. R package version 0.1.0 2019. https://CRAN.R-project.org/package=ehaGoF

Eyduran, E., Akın, M., and Eyduran, S. P. (2019). *Application of Multivariate Adaptive Regression Splines through R Software* (1st edition). Nobel Academic Publishing,.

Eyduran, E., Karakus, K., Keskin, S., and Cengiz, F. (2008). Determination of factors influencing birth weight using regression tree (Rt) method. *Journal of Applied Animal Research*, *34*(2), 109–112. https://doi.org/10.1080/09712119.2008.9706952

Eyduran, E., Yakubu, A., Duman, H., Aliyev, P., and Tırınk, C. (2020). Predictive modeling of multivariate adaptive regression splines: An R Tutorial. In *Veri Madenciliği Yöntemleri: Tarım Alanında Uygulamaları* (pp. 25–48). Rating Academy Ar-Ge Yazılım Yayıncılık Eğitim Danışmanlık ve Organizasyon Ticaret Limited Şirketi.

Eyduran, E., Yilmaz, I., Tariq, M. M., and Kaygisiz, A. (2013). Estimation of 305-d milk yield using regression tree method in brown Swiss cattle. *Journal of Animal and Plant Sciences*, *23*(3), 731–735.

Eyduran, E., and Zaborski, D. (2017). 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 Ind... *Pakistan Journal of Zoology*, *49*(1), 257–265. https://doi.org/10.17582/journal.pjz/2017.49.1.257.265

Faraz, A., Tirink, C., Eyduran, E., Waheed, A., Tauqir, N. A., Nabeel, M. S., and Tariq, M. M. (2021). Prediction of live body weight based on body measurements in Thalli sheep under tropical conditions of Pakistan using cart and mars. In *Tropical Animal Health and Production* (Vol. 53, Issue 2). https://doi.org/10.1007/s11250-021-02748-6

Genç, S., and Soysal, M. İ. (2018). Türkiye Siyah Alaca Sığır Populasyonlarında Süt ve Döl Verimi * Milk Yield and Reproductive Traits of Holstein Cattle Population in Turkey * Giriş Materyal ve Yöntem Materyal. *Journal of Tekirdag Agricultural Faculty*, *15*(01).

Grzesiak, W., and Zaborski, D. (2012). Examples of the Use of Data Mining Methods in Animal Breeding. In *Data Mining Applications in Engineering and Medicine* (pp. 304–321). IntechOpen. https://doi.org/10.5772/50893

Grzesiak, W., Zaborski, D., Sablik, P., Zukiewicz, A., Dybus, A., and Szatkowska, I. (2010). Detection of cows with insemination problems using selected classification models. *Computers and Electronics in Agriculture*, *74*(2), 265–273. https://doi.org/10.1016/j.compag.2010.09.001

Gürses, M., and Bayraktar, M. (2012). Some milk production and reproductive traits of Holstein cattle raised in different regions of Turkey. *Kafkas Universitesi Veteriner Fakultesi Dergisi*, *18*(2), 273–280. https://doi.org/10.9775/kvfd.2011.5424

Gürses, R. (2019). *İlk laktasyondaki kırmızı-alaca ve siyah-alaca sığırlarda süt ve döl verimi ile süt kalite özellikleri üzerine bir araştırma*. Aydın Adnan Menderes Üniversitesi.

Iqbal, F., Eyduran, E., Raziq, A., Ali, M., Zil-e-Huma, Tirink, C., and Sevgenler, H. (2021). Modeling and predicting the growth of indigenous Harnai sheep in Pakistan: non-linear functions and MARS algorithm. *Tropical Animal Health and Production 2021 53:2*, *53*(2), 1–12. https://doi.org/10.1007/S11250-021-02700-8

Keskin, İ., and Boztepe, S. (2011). Siyah Alaca Sığırlarda Kısmi Süt Verimlerinden Yararlanılarak 305 Günlük

Süt Veriminin Tahmini. *Tekirdağ Ziraat Fakültesi Dergisi*, *8*(1), 1–8.

Kopuzlu, S., Emsen, H., Özlütürk, A., and Küçüközdemir, A. (2008). Esmer ve Siyah Alaca Irkı Sığırların Doğu Anadolu Tarımsal Araştirma Enstitüsü Şartlarında Döl Verim Özellikleri. *Lalahan Hay. Araşt. Enst. Derg.*, *48*(1), 13–24.

Kulekçi, M., Eyduran, E., Altın, A. Y., and Tariq, M. M. (2022). Usefulness of MARS and Bagging MARS Algorithms in Prediction of Honey Production in Beekeeping Enterprises from Elazig Province of Turkey. *Pakistan Journal of Zoology*, *54*(3), 1087–1093.

Mee, J. F. (2004). peer reviewed Correspondence : *Irish Veterinary Journal*, *57*(4), 226–231.

Milborrow, S. (2018). Milborrow. Derived from mda:mars by T. Hastie and R. Tibshirani. İnternet url: https://CRAN.R-project.org/package=earth (08.10.2018)

Milborrow, S. (2011). Derived from MDA: MARS by T. Hastie and Tibshirani earth: Multivariate adaptive regression splines, R package.

Nayana, B. M., Kumar, K. R., and Chesneau, C. (2022). Wheat Yield Prediction in India Using Principal Component Analysis-Multivariate Adaptive Regression Splines (PCA-MARS). *AgriEngineering 2022, Vol. 4, Pages 461-474*, *4*(2), 461–474. https://doi.org/10.3390/AGRIENGINEERING4020030

Omar, M. Y. (2022). *Holstein ve Simmental İneklerde İlk Laktasyonda Süt Verimi, Pike Çıkma-Pikte Kalma Süreleri ve Süt Verimi Persistensleri ile Üreme Performanslarının Karşılaştırılması*. Bursa Uludağ Üniversitesi.

Orhan, H., Çetin Teke, E., and Karcı, Z. (2018). Laktasyon Eğrileri Modellemesinde Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (MARS) Yöntemi Uygulaması. *Kahramanmaraş Sütçü İmam Üniversitesi Doğa Bilimleri Dergisi*, *21*(3), 363–373. https://doi.org/10.18016/ksudobil.334237

Orhan, H., and Kaygısız, A. (2002). Siyah Alaca Sığırlarda Farklı Laktasyon Eğrisi Modellerinin Karşılaştırılması. *Hayvansal Üretim*, *43*(1), 94–99.

Otok, B. W., Putra, R. Y., Sutikno, and Yasmirullah, S. D. P. (2020). Bootstrap aggregating multivariate adaptive regression spline for observational studies in diabetes cases. *Systematic Reviews in Pharmacy*, *11*(8), 406–413. https://doi.org/10.31838/srp.2020.8.59

Öztürk, S. T. B. (2022). Evaluation of - PM 10 concentration by using Mars and XGBOOST algorithms in Iğdır Province of Türkiye. *International Journal of Environmental Science and Technology*. https://doi.org/10.1007/s13762-022-04511-2

Özyurt, A., and Özkan, M. (2009). Orta Anadolu'da Yetiştirilen Siyah-Alaca Sığırlarda Laktasyon Eğri Şekli ve Eğriye Etkili Olan Faktörler. *Hayvansal Üretim*, *50*(1), 31–37.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, http://www.R-project.org

Şahin, A., and Ulutaş, Z. (2010). Polatlı Tarım İşletmesinde Yetiştirilen Siyah Alaca İneklerde Süt ve Döl Verim Özellikleri. *Anadolu Tarım Bilimleri Dergisi*, *25*(3), 202–212.

Sarar, A. D., and Tapkı, İ. (2017). Türk Tarım - Gıda Bilim ve Teknoloji Dergisi Türkiye ' de Yetiştirilen Holştayn İneklerde Döl Verim Özelliklerine Ait. *Türk Tarım - Gıda Bilim ve Teknoloji Dergisi*, *5*(12), 1476–1481.

Şengül, Ö., Çelik, Ş., and İbrahim, A. K. (2022). Determination of the Effects of Silage Type, Silage Consumption, Birth Type and Birth Weight on Fattening Final Live Weight in Kıvırcık Lambs with MARS and Bagging MARS Algorithms. *Kafkas Universitesi Veteriner Fakultesi Dergisi*, *28*(3), 379–389. https://doi.org/10.9775/kvfd.2022.27149

Tirink, C., Sariçiçek, Z., and Önder, H. (2022). *Determination of The Direct and Indirect Effects of Feed Characteristics on the Metabolizable Energy of Natural Pasture Hay Mera Samanı İçin Yem Özelliklerinin Metabolize Edilebilir Enerjisi Üzerine Path Analizi*. *37*(June), 331–340.

Tırınk, C. (2021). Analysis of Continuous Proportional Data: Case Study of Determination of Milk Protein. In Y. Keskin, İsmail; Mikail, Nazire; Altay (Ed.), *Different Statistical Applications in Agriculture* (1st ed., Issue January, pp. 211–229). Iksad Publications.

Tırınk, C., Piwczyński, D., Kolenda, M., and Önder, H. (2023). Estimation of Body Weight Based on Biometric Measurements by Using Random Forest Regression, Support Vector Regression and CART Algorithms. *Animals*, *13*(5), 798.

Turhan, M. (2020). Mars Algoritmasi Uygulanarak Farkli Gruplardaki Ari Kolonilerinde Bulunan Varroa'nin Bulaşiklik Tespiti. In Ş. Çelik (Ed.), *Veri Madenciliği Yöntemleri: Tarim Alanında Uygulamaları* (1st Editio, pp. 1–23). Rating Academy Ar-Ge Yazılım Yayıncılık Eğitim Danışmanlık ve Organizasyon Ticaret Limited Şirketi,.

Yaylak, E., and Kumlu, S. (2005). Siyah Alaca Sığırların 305 Günlük Süt Verimine Vücut Kondisyon Puanı ve Bazı Çevre Faktörlerinin Etkisi. *Ege Üniv. Ziraat Fak Derg*, *42*(3), 55–66.