Selcuk Journal of Agriculture and Food Sciences

https://dergipark.org.tr/tr/pub/selcukjafsci

Research Article

(2025) 39(2), 476-486 DOI: 10.15316/selcukjafsci.1653578 e-ISSN: 2458-8377 Selcuk J Agr Food Sci

Prediction of Lactation Milk Yield in Simmental Cattle Milked with Robotic Milking System Using CHAID and CART Algorithms

Rabia Albayrak Delialioğlu[1,*], Ayşe Övgü Şen[2]

[1]Ankara University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetics Unit, Ankara, Türkiye

[2]Ankara University, Faculty of Agriculture, Department of Animal Science, Ankara, Türkiye

HIGHLIGHTS

- Robotic milking systems improve efficiency and hygiene while reducing the need for human labor.
- Days in Milk (DIM) is the most important factor affecting lactation milk yield (LMY).
- Cows on higher feed allocations achieve notably increased milk yields.
- Advanced algorithms support tailored strategies for improved milk yield and animal welfare in robotic milking farms.

Abstract

Animal husbandry has long been a key component of agriculture, fulfilling essential nutritional needs. Technological advancements have gradually replaced human labor with machines, particularly in dairy farms, where the milking process is vital for economic sustainability. Robotic milking systems have emerged as significant innovations, allowing for efficient, hygienic, and automated milking while reducing dependence on labor. This study aims to predict lactation milk yield (LMY) in Simmental cows during their first lactation period using the Classification and Regression Tree (CART) and Chi-squared Automatic Interaction Detector (CHAID) algorithms, both of which are widely applied in data mining. To achieve this, the independent variables included Days in Milk (DIM), Status (S), Number of Inseminations (NI), Milk Flow Rate (MFR), Number of Robot Rejection (NRR), Rumination Time (RT), Time Spent in the Robot (TSR), Feed Amount in the Robot (FAR), Feed Consumption Rate in the Robot (FCRR), and Milking Frequency (MF). The analysis identified DIM as the most influential predictor of LMY, while FAR, RT, and MF also played significant roles in the prediction model. Findings from the CHAID algorithm demonstrated that feed allocation based on milk yield in robotic systems directly impacted lactation milk yield. Especially, newly calved cows (DIM <30 days) produced an average of 5 692 L of milk; however, those receiving more than 5.09 kg of feed in the robotic system (FAR >5.09) exhibited an estimated lactation yield of 8 426 L. During the 30–81 days of lactation, increased feed availability was positively correlated with milk yield. Additionally, the CART algorithm supported these findings, reinforcing days in milk (DIM) as a key determinant in estimating lactation milk yield. In terms of predictive performance, the CHAID algorithm achieved an R^2 value of 78.5% and demonstrated superior prediction accuracy in most criteria, except for MAPE and MAD. Overall, robotic milking

Citation: Albayrak Delialioğlu R, Şen AÖ (2025). Prediction of lactation milk yield in Simmental cattle milked with robotic milking system using CHAID and CART Algorithms. *Selcuk Journal of Agriculture and Food Sciences*, 39(2), 476-486. https://doi.org/10.15316/selcukjafsci.1653578
***Correspondence:** ralbayrak@ankara.edu.tr
Received date: 07/03/2025
Accepted date: 26/06/2025

systems enhance the individualized management of dairy cows by optimizing factors such as feed allocation and milking frequency. Further research in this area could provide valuable insights, even for farms without robotic milking systems, guiding dairy farmers on strategies to maximize lactation milk yield effectively.

Keywords: Robotic milking; milk yield; cattle, data mining; CART; CHAID

1. Introduction

Animal husbandry has been an important and indispensable part of agriculture for generations, providing for individuals' fundamental nutritional requirements. For this purpose, the care, feeding, etc. needs of animals raised on farms have been met by human power for many years. However, thanks to technological developments, machines have started to replace human labor over time (Isık et al. 2021; Demir and Öztürk 2010)

In dairy farms, a significant portion of income is derived from milk, making the milking process a critical component of farm management (Britt et al. 2018). For this reason, the milking procedure must be performed quickly, hygienically, and with minimal labor input (Bach and Cabrera 2017). Technological advancements have enabled the automation of milking process using robotic systems, thereby significantly reducing reliance on human labor. Although the initial installation costs of robotic systems are high, their long-term benefits, including labor cost reduction, improved hygienic conditions, and providing continuous operation, make them advantageous over time. A study conducted by Mathijs (2004) in Belgium, Denmark, Germany, and the Netherlands revealed that robotic milking operations saved approximately 20% in labor costs compared to traditional systems, with this difference being statistically significant.

Robotic milking systems began to be developed in the late 20th century, with significant contributions emerging from the Netherlands, Sweden, and the United States in the early 1990s (Bach and Cabrera 2017; Simões Filho et al. 2020). One of the first robotic milking systems was introduced in the Netherlands in 1992, capable of automatically identifying cows in the herd, performing the milking process autonomously, and monitoring both the quality and quantity of milk during the procedure. As a result, robotic milking systems gradually gained popularity and were widely adopted in modern dairy farming operations worldwide, with the growing demand for this technology evident in its rapid rate of adoption (Bach and Cabrera 2017; Ji et al. 2022).

Robots are automatic programmable devices that perform all operations of the milking cycle with maximum accuracy and minimal variation. They handle numerous functions, such as preparing the udder, connecting the milking machine, and disinfecting teat liners. Robotic milking allows for individualized milking frequency based on daily milk yield and lactation phase. Furthermore, robotic milking systems give priority to animal welfare by providing a consistent and efficient milking process, promoting long-term improvements in farm productivity, milk yield, and animal health, while also enhancing farmers' working conditions and facilitating better monitoring and management of the animals (Jacobs and Siegford 2012; Rodenburg 2017; Tse et al. 2018; Coşkun et al. 2023; Himu and Raihan 2024; Johansen et al. 2025).

The aim of this study is to predict lactation milk yield (LMY) in Simmental cows during the first lactation period in robotic milking farms using the variables Days in Milk (DIM), Status (S), Number of Inseminations (NI), Milk Flow Rate (MFR), Number of Robot Rejection (NRR), Rumination Time (RT), Time Spent in the Robot (TSR), Feed Amount in the Robot (FAR), Feed Consumption Rate in the Robot (FCRR), and Milking Frequency in a Day (MF) through the application of CART (Classification and Regression Trees) and CHAID (Chi-squared Automatic Interaction Detector) algorithms.

2. Materials and Methods

2.1 Animal Material

The material of this study consists of 106 Simmental cows in the first lactation period, obtained from a dairy farm in Denizli utilizing a robotic milking system. While the roughage sources for feeding the animals were hay, alfalfa, and meadow grass given ad libitum in the paddock, concentrated feed was given in the automatic weighing feeders in the milking robot while the animals were milked. Lactation Milk Yield (LMY), Days in Milk (DIM), Status (S), Number of Inseminations (NI), Milk Flow Rate (MFR), Number of Robot Rejection (NRR), Rumination Time (RT), Time Spent in the Robot (TSR), Feed Amount in the Robot (FAR), Feed Consumption Rate in the Robot (FCRR), and Milking Frequency (MF) variables were obtained from these cows.

2.2. Method

Lactation milk yield of the cows was predicted using the independent variables DIM, S, NI, MFR, NRR, RT, SR, FAR, FCRR, and MF via CART and CHAID algorithms.

2.2.1. Chi-squared Automatic Interaction Detector (CHAID) Algorithm

The CHAID (Chi-squared Automatic Interaction Detection) algorithm is a data mining algorithm developed by Gordon V. Kass in 1980. The algorithm is designed to explore relationships between dependent and independent variables and to identify the groups that best explain the dependent variable. Unlike other decision tree algorithms such as CART, which uses splitting points based on certain criteria, CHAID determines split points based on the significance level of the chi-square test. This approach allows for the evaluation of whether the differences between the resulting classes are statistically significant, making the splits based on hypothesis testing. Additionally, the CHAID algorithm permits more than two splits, unlike many other algorithms. Moreover, CHAID combines and splits categories of independent variables to prevent overfitting, thus reducing the risk of overlearning the model (Kass, 1980; Maimon and Rokach, 2014; Altay et al., 2021).

2.2.2. Classification and Regression Trees (CART) Algorithm

The CART algorithm, developed by Breiman et al. in 1984, is an extension of the AID (Automatic Interaction Detection) decision tree algorithm introduced by Morgan and Sonquist (1963a, b). CART is a non-parametric statistical method used for solving classification and regression problems with both categorical and continuous data. This algorithm transforms the complex structures within a heterogeneous dataset into simple decision structures by dividing the dataset into homogeneous groups. CART examines both the interactions between independent variables and the relationships between dependent and independent variables, with branching occurring in a binary tree structure. (Altay and Albayrak Delialioğlu 2022; Altay 2022b). In the process of selecting the best independent variable using the CART algorithm, the Gini index is used as the branching criterion (Zhang et al. 2018). The Gini index defines the purity of a category after partitioning based on a specific feature. The split leading to the formation of the purest classes is considered the best split. A situation where all individuals belong to a single class is considered pure. The Gini index varies between 0 and 1; 0 indicates perfect classification purity, where all individuals belong to the same class, resulting in only one class. A value of 1 indicates that individuals are randomly distributed across different classes. A Gini index value of 0.5 means that individuals are equally distributed among the classes. When constructing a regression tree, features with the lowest Gini index value are preferred (Maimon and Rokach 2014). The Gini Index is computed by taking one minus the sum of the squared class probabilities, as illustrated in Equation (1) (Tangirala 2020).

$$\text{Gini Index}\,(L) = 1 - \sum_{i=1}^{j} p_i^2 \tag{1}$$

j : Number of class

L : A dataset with j different class labels

Pi : Relative frequency if class i in T′

2.2.3. Data Analysis

In the study, regression tree algorithms, CHAID and CART, were used to predict lactation milk yield in 106 Simental dairy cows. The algorithms were performed with a parent node size of 6, a child node size of 3 and 5-fold cross-validation. The regression tree analyses were conducted with IBM SPSS Statistics 23 software package (IBM Corp. Released 2015). Model fit criteria were obtained with the "ehaGoF" package (Eyduran 2020; R Core Team 2020). The Pearson correlation coefficients between variables were calculated and correlogram plots were generated using the 'corrplot' package in R (Wei and Simko 2021).

2.2.4. The Performance Criteria Used to Evaluate the Goodness of Fit of Data Mining Algorithms

The goodness-of-fit criteria used to assess the predictive performance of the CART and CHAID algorithms are provided in Table 1 (Grzesiak and Zaborski 2012; Zaborski et al. 2019). In the evaluation of the prediction performance of tree-based data mining algorithms used in LMY prediction, the goodness-of-fit criteria, including the global relative error of approximation (RAE), standard deviation ratio (SDratio), root-mean-square error (RMSE), mean absolute percentage error (MAPE), and mean absolute deviation (MAD), should be very close to 0, while the (adjusted) coefficient of determination (AdjR^2 and R^2) values should be close to 1 (Altay 2022a; Topal et al. 2010).

Table 1. Goodness-of-fit criteria considered in the study.

Equation	Equation (continued)
$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	$MRAE = \sqrt{\frac{1}{n}RAE}$
$RRMSE = \frac{RMSE}{\bar{y}} * 100$	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left\|\frac{y_i - \hat{y}_i}{y_i}\right\| * 100$
$SD_{ratio} = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}}$	$MAD = \frac{1}{n}\sum_{i=1}^{n}\|y_i - \hat{y}_i\|$
$CV(\%) = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}}{\bar{y}} * 100$	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$
$r = \frac{cov(y_i, \hat{y}_i)}{S_{y_i}S_{\hat{y}_i}}$	$Adj - R^2 = 1 - \frac{(1 - R^2)(n-1)}{(n-k-1)}$
$PI = \frac{RRMSE}{1 + r}$	$AIC = n\,In\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right] + 2k$
$ME = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$	
$RAE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}y_i^2}}$	if n/k>40 $AIC_c = AIC + \frac{2k(k+1)}{n-k-1}$

RMSE: root mean square error, RRMSE: relative root mean square error, SDR: standard deviation ratio, CV: coefficient of variation, r: Pearson's correlation coefficients, PI: performance index, ME: mean error, RAE: relative approximation error, MRAE: mean relative approximation error, MAPE: mean absolute percentage error, MAD: mean absolute deviation, R2: coefficient of determination, Adj-R2: adjusted coefficient of determination, AIC: Akaike's information criterion, CAIC: corrected Akaike's information criterion

3. Results and Discussion

Descriptive statistics for lactation milk yield, days in milk, milk flow rate, rumination time, time spent in the robot, feed amount in the robot, feed consumption rate in the robot, and milking frequency, obtained from 106 Simmental cows in their first lactation, are presented in Table 2.

Table 2. Descriptive statistics for the continuous variables considered in the study.

Variables	Mean± SEM	Min	Max	CV
LMY (L)	8 902.00 ±191.000	915.00	12 673.00	% 22.10
DIM (days)	168.30 ±11.100	1.00	362.00	% 67.77
MFR (kg/min)	2.82 ±0.082	1.02	5.82	% 29.80
RT (min)	525.68 ±8.840	144.00	688.00	% 17.31
TSR (sec)	361.85 ±8.950	192.00	705.00	% 25.47
FAR (kg)	3.41 ± 0.129	1.00	5.63	% 39.07
FCRR (%)	0.96 ± 0.007	0.63	1.00	% 7.05
MF (times/day)	2.98 ±0.049	1.00	4.10	% 16.77

LMY: Lactation Milk Yield, DIM: Days in Milk, MFR: Milk Flow Rate, RT: Rumination Time, TSR: Time Spent in the Robot, FAR: Feed Amount in the Robot, FCRR: Feed Consumption Rate in the Robot, MF: Milking Frequency

Upon examining Table 2, it can be observed that the average lactation milk yield for the cows is 8 902 liters, with most of the cows being in the mid-lactation period (DIM; 168 days). The average daily milking frequency is 3 times, and nearly all of the feed provided in the robot is consumed by the cows. The average rumination time for the cows is 526 minutes, while the time spent in the milking robot is approximately 362 seconds (around 6 minutes).

The frequency distribution table and percentages for the values of Status, Number of Inseminations, and Robot Rejection Number for 106 Simental cows are provided in Table 3.

Table 3. The frequency distribution table for the discrete/categorical variables considered in the study.

Variables	Levels	n	%
Status (S)	Estrus	9	8.5
	Inseminated	39	36.8
	Non-pregnant	20	18.9
	Pregnant	38	35.9
Number of Inseminations (NI)	1	62	58.5
	2	24	22.6
	3	11	10.4
	4	6	5.7
	5	1	0.9
	6	2	1.9
Number of Robot Rejection (NRR)	0	61	57.6
	1	32	30.2
	2	7	6.6
	3	5	4.7
	4	1	0.9

The Pearson correlation coefficients among the continuous variables measured from the cows in the study are presented in Figure 1 using a correlogram plot.

The correlogram in Figure 1 demonstrates a generally positive but weak relationship among the variables. The strongest correlations with lactation milk yield are observed for milking frequency (0.61*) and rumination time (0.58**) ($p<0.01$). Johnston and DeVries (2018) indicated that cows with longer rumination times had higher milk yields. Additionally, negative correlations were observed between days in milk and both the feed amount in the robot and the feed consumption rate in the robot (-0.63** $p<0.01$, -0.21* $p<0.05$). Therefore, in the early stages of lactation, cows have higher feed requirements, leading to a greater amount of feed being allocated in the robot. However, while the feed consumption rate is relatively lower in the initial days of lactation, it increases as lactation progresses. Considering that feed demand is high during the first 2.5–3 months post-calving, but not all the allocated feed is consumed, implementing appropriate adjustments could help reduce overall feed costs and improve farm efficiency.

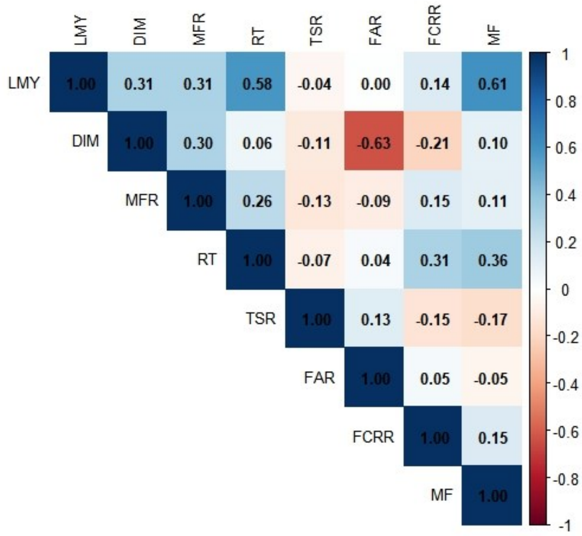

Figure 1. Correlogram graph of data for dairy cows [LMY: Lactation Milk Yield, DIM: Days in Milk (DIM), MFR: Milk Flow Rate, RT: Rumination Time, TSR: Time Spent in the Robot, FAR: Feed Amount in the Robot, FCRR: Feed Consumption Rate in the Robot, MF: Milking Frequency]

3.1. CHAID Algorithm

The tree diagram for the CHAID algorithm used to predict lactation milk yield is presented in Figure 2. In predicting lactation milk yield, the most critical variable is days in milk. In the herd, the average lactation milk yield (LMY) of newly calved cows (n=10, DIM<30) is 5 692 L. Given that feed allocation in robotic milking systems is based on milk yield, the 10 cows in Node-1 are divided into three groups according to the feed amount in the robot. Among these groups, those receiving more than 5.09 kg of feed exhibit the highest average LMY, reaching 8 426 L.

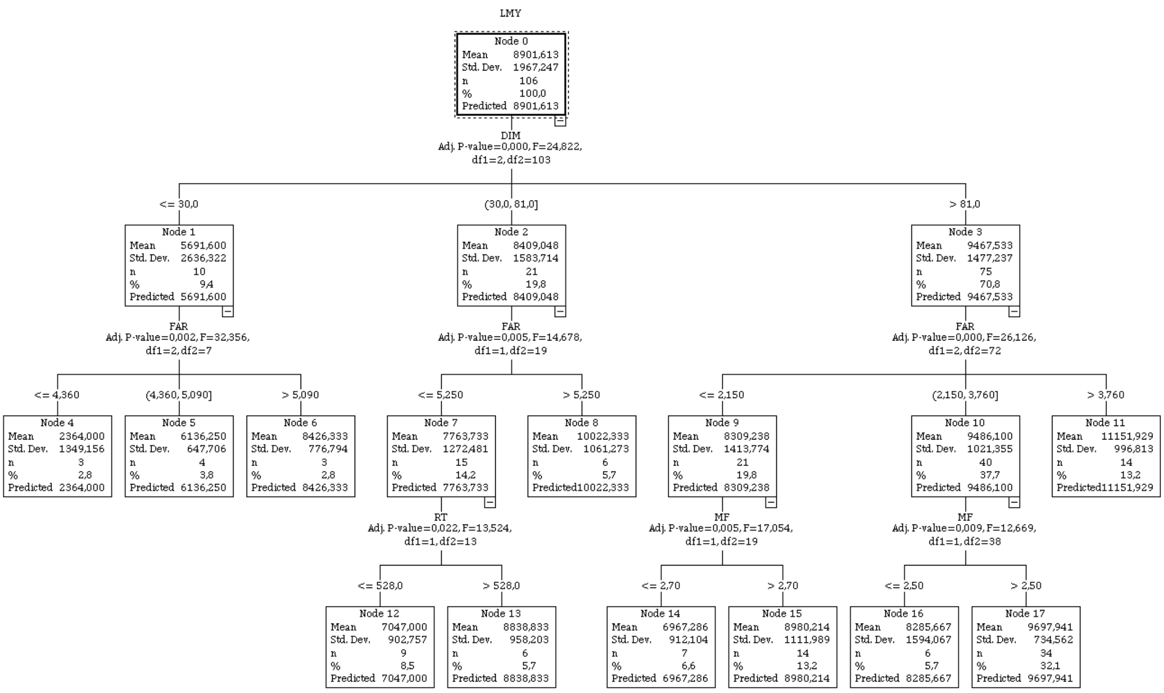

Figure 2. The results of CHAID algorithm (LMY: Lactation Milk Yield, DIM: Days in Milk, FAR: Feed Amount in the Robot, RT: Rumination Time, MF: Milking Frequency)

For cows in days 30 to 81 of lactation (n=21, 30<DIM<81), when FAR is less than 5.25 kg, the average LMY is 7 764 L. Conversely, when FAR exceeds 5.25 kg, the average LMY reaches 10 022 L. The cows in the first group (n=15, FAR<5.25 kg) are further divided into two subgroups based on rumination time. In this subgroup, cows with a rumination time exceeding 528 minutes have a significantly higher average LMY (8 839 L) compared to those with a rumination time of less than 528 minutes. Johnston and DeVries (2018) observed in their study that for every 1-hour increase in rumination time, there was an average increase of 1.26 kg day^{-1} in milk yield. Similarly, Dado and Allen (1994) observed that high-producing dairy cattle tended to have fewer, but longer, rumination time.

The group with the highest average lactation milk yield (Node-3) consists of cows that are at least 81 days post-calving. Within this group, cows are further divided into three sub-categories based on feed allocation in the robot. Among these sub-categories, the group with the highest average lactation milk yield (Node-11, n=14, mean LMY=11 152 L) is associated with a feed allocation (FAR) of 3.76 kg or more. Notably, milk yield in cows during their first lactation increases after the 81st day of lactation (i.e., after the service period), while the feed amounts allocated in the robot are relatively lower compared to those in the earlier stages of lactation.

For cows with a days in milk yield greater than 81, those in subgroups with feed amounts in the robot below 2.15 kg or between 2.15–3.76 kg had a higher average lactation milk yield when the milking frequency was approximately three times per day. It has been observed that when the milking frequency is three times daily, milk yield increases compared to systems with a twice-daily milking frequency (Erdman and Varner 1995; Amos et al. 1985; Gisi et al. 1986). Furthermore, Hart et al. (2013) emphasized that increasing the milking frequency from two to three times per day resulted in an approximate increase of 2.9 kg/day in daily milk yield.

When lactation milk yield was predicted using the CHAID algorithm with days in milk, reproductive status, number of inseminations, milk flow rate, number of robot rejections, rumination time, time spent in the robot, feed amount in the robot, feed consumption rate in the robot, and milking frequency as predictor variables, the coefficient of determination was found to be 78.5%, which was statistically significant ($p<0.01$).

3.2. CART Algorithm

In the CART algorithm given in Figure 3, which allows for a binary tree structure, days in milk was identified as the most influential variable on lactation milk yield, similar to the CHAID algorithm In this model, the average lactation milk yield was found to be 9 211 L at least 18 days post-calving. During the second period (the first 18 days post-calving), the cow's milk yield increased approximately 2.5 times compared to the first 7 days.

At least 18 days after calving, when the milking frequency of high-yielding cows was 2.85 or higher, the average lactation milk yield was found to be approximately 9 786 L. In the group with a milking frequency below 2.85, cows with a rumination time exceeding 603 minutes had a higher average milk yield (9 761 L). However, in cows milked more than 2.85 times per day, those with a rumination time longer than 527 minutes had an average lactation milk yield of 10 161 L.

Cows milked less than 2.85 times per day and with a rumination time below 603 minutes, but with a time spent in the robot exceeding 419 seconds, had a higher lactation milk yield (8 923 L). At least 18 days after calving, in cows with a daily milking frequency greater than 2.85, an increase in rumination time and time spent in the robot was associated with higher milk yield. The highest milk yield was observed in the group in Node-20, where the milk flow rate was above 2.95. Similarly, studies conducted in this context have shown that increasing milking frequency from twice to three times per day improves milk yield without exhibiting any adverse effects on reproductive performance indicators (Kruip et al., 2002).

When lactation milk yield was predicted using the CART algorithm with days in milk, status, number of inseminations, milk flow rate, number of robot rejection, rumination time, time spent in the robot, feed amount in the robot, feed consumption rate in the robot, and milking frequency as predictor variables, the determination coefficient was found to be 77.3% ($p<0.01$).

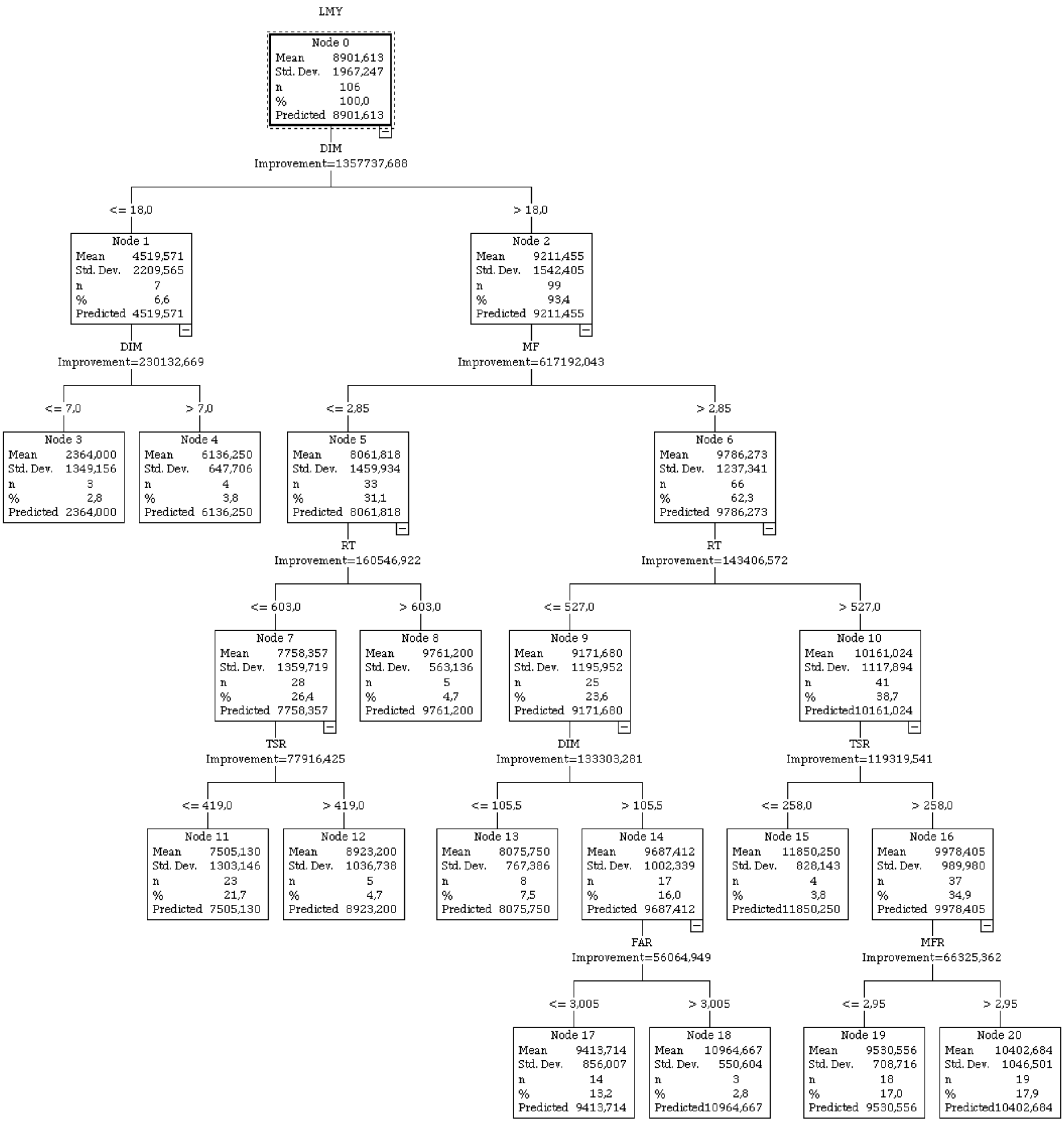

Figure 3. The results of CART algorithm (LMY: Lactation Milk Yield, DIM: Days in Milk, MF: Milking Frequency, RT: Rumination Time, TSR: Time Spent in the Robot, FAR: Feed Amount in the Robot, MFR: Milk Flow Rate)

3.3. *Goodness-of-Fit Criterias for Data Mining Algorithms*

When considering the goodness-of-fit criteria, it was observed that the prediction performances of both the CHAID and CART algorithms were relatively similar (Table 4). However, in all criteria except for MAPE and MAD, the CHAID algorithm demonstrated relatively better prediction performance compared to CART. In this context, the use of the CHAID algorithm instead of the CART algorithm for predicting lactation milk yield in robotic milking systems, based on automatically measured parameters, would allow for more reliable results.

Table 4. Predictive performance criteria for CHAID and CART algorithms.

Prediction Performance Criteria	CHAID	CART
RMSE	907.739	933.599
RRMSE	10.197	10.488
SDR	0.464	0.477
CV	10.250	10.540
r	0.886	0.879
PI	5.407	5.582
ME	0.028	0.132
RAE	0.010	0.010
MRAE	0.010	0.010
MAPE	10.143	9.652
MAD	742.877	698.736
R^2	0.785	0.773
Adj-R^2	0.781	0.769
AIC	1447.923	1453.878
AIC_c	1448.039	1453.994

4. Conclusions

The dairy industry is increasingly influenced by technological advancements and continues to evolve. The development of milking robots is expected to become more important in the future due to their potential to reduce labor and ensure reliable food supply. The rapid development of technological tools and algorithms will greatly contribute to the production and improvement of new products and systems.

This study focuses on predicting lactation milk yield in first-lactation Simmental cows milked by robotic milking systems, utilizing data mining algorithms such as CHAID and CART, based on routinely collected data within the system. The findings obtained from the algorithms will assist the researcher in determining the additional feed amount to be provided during milking in small enterprises where robotic milking systems are not available and are quite costly.

Furthermore, by evaluating the effects of parameters such as milk yield, feed consumption during milking, time spent in the robot, rumination time, and milk flow rate on lactation milk yield, this study will facilitate adjustments aimed at increasing farm profitability. It is recommended that further comprehensive studies be conducted on other characteristics that are thought to be related to lactation milk yield. The results of this study provide practical insights into dairy cattle farms.

Author Contributions: All authors have contributed equally to the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Ethical Statement: This study was conducted using data obtained from a voluntary dairy farm operating a robotic milking system. The data were collected automatically by the system as part of routine herd management practices, without any experimental intervention or harm to the animals. Therefore, ethical approval was not required for this research.

References

Altay Y (2022a). Prediction of the live weight at breeding age from morphological measurements taken at weaning in indigenous Honamli kids using data mining algorithms. *Tropical Animal Health and Production* 54(3), 172.

Altay Y (2022b). Phenotypic characterization of hair and Honamli goats by using classification trees algorithms and multivariate adaptive regression splines (MARS). *Kafkas Üniversitesi Veteriner Fakültesi Dergisi* 28(3), 401-410. https://doi.org/10.9775/kvfd.2022.27163

Altay Y and Albayrak Delialioğlu R (2022). Diagnosing Lameness with the Random Forest Classification Algorithm Using Thermal Cameras and Digital Colour Parameters. *Mediterranean Agricultural Sciences* 35, 47-54.

Altay Y, Boztepe S, Eyduran E, Keskin I, Tariq MM, Bukhari FA, Ali I (2021). Description of factors affecting wool fineness in Karacabey Merino sheep using CHAID and MARS algorithms. *Pakistan Journal of Zoology* 53(2), 691. https://doi.org/10.17582/journal.pjz/20190329150359

Amos HE, Kiser T, Loewenstein M. (1985). Influence of milking frequency on productive and reproductive efficiencies of dairy cows. *Journal of Dairy Science* 68(3), 732-739.

Bach A, Cabrera V (2017). Robotic milking: Feeding strategies and economic returns. *Journal of Dairy Science* 100(9), 7720-7728. https://doi.org/10.3168/jds.2016-11694

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984). Classification and regression trees. Chapman and Hall, Wadsworth Inc., New York, NY, USA.

Britt J, Cushman R, Dechow C, Dobson H, Humblot P, Hutjens M, Jones G, Ruegg P, Sheldon I, Stevenson J (2018). Invited review: Learning from the future—A vision for dairy farms and cows in 2067. *Journal of Dairy Science* 101(5), 3722-3741. https://doi.org/10.3168/jds.2017-14025

Coşkun G, Şahin Ö, Aytekin İ (2023). Robotik Sağımda Sürü Yönetimi. *Turkish Journal of Agricultural Research* 10(3), 361-371. https://doi.org/10.19159/tutad.1339586

Dado RG, Allen MS. (1995). Intake limitations, feeding behavior, and rumen function of cows challenged with rumen fill from dietary fiber or inert bulk. *Journal of Dairy Science* 78(1), 118-133.

Demir B, Öztürk İ (2010). Robotlu sağım sistemleri. *Alinteri Journal of Agriculture Science* 19(2), 21-27.

Erdman RA, Varner M (1995). Fixed yield responses to increased milking frequency. *Journal of Dairy Science* 78(5), 1199-1203.

Eyduran, E. (2020). ehaGoF: Calculates Goodness of Fit Statistics. R package version 0.1.0. https://CRAN.Rproject.org/package=ehaGoF

Gisi DD, DePeters EJ, Pelissier CL. (1986). Three times daily milking of cows in California dairy herds. *Journal of Dairy Science* 69(3), 863-868.

Grzesiak W, Zaborsk D (2012). Examples of the use of data mining methods in animal breeding. Data mining applications in engineering and medicine 303-324.

Hart KD, McBride BW, Duffield TF, DeVries TJ. (2013). Effect of milking frequency on the behavior and productivity of lactating dairy cows. *Journal of Dairy Science* 96(11), 6973-6985.

Himu HA, Raihan A (2024). An overview of precision livestock farming (PLF) technologies for digitalizing animal husbandry toward sustainability. *Global Sustainability Research* 3(3), 1-14.

IBM Corp. Released (2015). IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.

Isık AH, Alakus F, Eskicioğlu ÖC (2021). Hayvancılıkta robotik sistemler ve yapay zeka uygulamaları. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi* 9(6), 370-382.

Jacobs JA, Siegford JM (2012). Invited review: The impact of automatic milking systems on dairy cow management, behavior, health, and welfare. *Journal of Dairy Science* 95(5), 2227-2247.

Ji B, Banhazi T, Phillips CJ, Wang C, Li B (2022). A machine learning framework to predict the next month's daily milk yield, milk composition and milking frequency for cows in a robotic dairy farm. *Biosystems Engineering* 216, 186-197. https://doi.org/10.1016/j.biosystemseng.2022.02.013

Johansen F, Buijs S, Arnott G (2025). Social rank and personality are associated with visit frequency in dairy cows learning to use an automatic milking system. *Animal* 19(3), 101446. https://doi.org/10.1016/j.animal.2025.101446

Johnston C, DeVries TJ. (2018). Associations of feeding behavior and milk production in dairy cows. *Journal of Dairy Science* 101(4), 3367-3373.

Kass GV (1980). An exploratory technique for investigating large quantities of categorical data. *Journal of the Royal Statistical Society: Series C (Applied Statistics*) 29(2), 119-127.

Kruip TAM, Morice H, Robert M, Ouweltjes W. (2002). Robotic milking and its effect on fertility and cell counts. *Journal of Dairy Science* 85(10), 2576-2581.

Maimon OZ, Rokach L (2014). Data mining with decision trees: theory and applications (Vol. 81). World Scientific.

Mathijs E (2004). Socio-economic aspects of automatic milking. In Automatic milking, a better understanding. Wageningen Academic. (pp. 46-55).

Morgan JN, Sonquist JA (1963a). Problems in the analysis of survey data, and a proposal. *Journal of the American Statistical Association* 58, 415–434.

Morgan JN, Sonquist JA (1963b). Some Results from a Non-Symmetrical Branching Processthat Looks for Interaction Effects. *Proceedings of the Social Statistics Section, American Statistical Association* 8, 40–53.

R Core Team (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/ (access date: 25.11.2024).

Rodenburg J (2017). Robotic milking: Technology, farm design, and effects on work flow. *Journal of Dairy Science* 100(9), 7729-7738. https://doi.org/jds.2016-11715

Simões Filho LM, Lopes MA, Brito SC, Rossi G, Conti L, Barbari M (2020). Robotic milking of dairy cows: a review. *Semina: Ciências Agrárias* 41(6), 2833-2850.

Tangirala S (2020). Evaluating the impact of Gini index and information gain on classification using decision tree classifier algorithm. *International Journal of Advanced Computer Science and Applications* 11(2), 612-619.

Topal M, Aksakal V, Bayram B, Yaganoglu AM (2010). An analysis of the factors affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis. *The Journal of Animal and Plant Sciences* 20(2), 63-6.9.

Tse C, Barkema HW, DeVries TJ, Rushen J, Pajor EA (2018). Impact of automatic milking systems on dairy cattle producers' reports of milking labour management, milk production and milk quality. *Animal* 12(12), 2649-2656. https://doi.org/10.1017/S1751731118000654

Wei T, Simko VR (2021). Package "corrplot": Visualization of a Correlation Matrix (Version 0.92). Package Corrplot for R Software.

Zaborski D, Ali M, Eyduran E, Grzesiak W, Tariq MM, Abbas F, Waheed A, Tirink C. (2019). 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.

Zhang Y, Chi G, Zhang Z (2018). Decision tree for credit scoring and discovery of significant features: an empirical analysis based on Chinese microfinance for farmers. *Filomat* 32(5), 1513-1521.