Computers and Electronics in Agriculture 182 (2021) 105995

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Machine learning algorithms for lamb survival

B.B. Odevci [a], E. Emsen [b,*], M.N. Aydin [a]

[a] Kadir Has University, Management Information Systems, Istanbul 34083, Turkey
[b] Ataturk University, Department of Animal Science, 25240 Erzurum, Turkey

ARTICLE INFO

Keywords:
Machine learning algorithms
Lamb survival

ABSTRACT

Lamb survival is influenced by the culmination of a sequence of often interrelated events including genetics, physiology, behaviour and nutrition, with the environment providing an overarching complication. Machine learning algorithms offer great flexibility with regard to problems of complex interactions among variables. The objective of this study was to use machine learning algorithms to identify factors affecting the lamb survival in high altitudes and cold climates. Lambing records were obtained from three native breed of sheep (Awassi = 50, Morkaraman = 50, Tuj = 50) managed in semi intensive systems. The data set included 193 spring born lambs out of which 106 lambs were sired by indigenous rams (n = 10), and 87 lambs were sired by Romanov Rams (n = 10).

Factors included were dam body weight at lambing, age of dam, litter size at birth, maternal and lamb behaviors, and lamb sex. Individual and cohort data were combined into an original dataset containing 1351 event records from 193 individual lambs and 750 event records from 150 individual ewes. Classification algorithms applied for lamb survival were Bayesian Methods, Artificial Neural Networks, Support Vector Machine and Decision Trees. Variables were categorized for lamb survival, lamb behavior, and mothering ability. RandomForest performed very well in their classification of the mothering ability while SMO was found best in predicting lamb behavior. REPtree tree visualization showed that grooming behavior is the first determinant for mothering ability. Classification Trees performed best in lamb survival. Our results showed that Classification Trees clearly outperform others in all traits included in this study.

1. Introduction

Lamb survival is a complex trait influenced by many different factors associated with management, climate, behavior of the ewe and lamb, and other environmental effects (Tomaszyk et al., 2014; Aktaş et al., 2015; and Moraes et al., 2016). Brien et al. (2010) suggested selecting related traits that are more reliably evaluated to make genetic improvement in lamb survival rather than improving it through genetic selection since heritability estimates of lamb survival are typically low (0.00–0.11; Safari et al., 2005). Correlated traits (recorded at lamb tagging) include birthweight (BWT), birth coat score (BCS), maternal behaviour score (MBS), lamb ease (LE), rectal temperature (RT), visually assessed lamb vigour (OBV), five timed lamb behaviours and three skeletal measures, crown–rump length (CRL), metacarpal bone length (ML) and thorax circumference (THO), (Fogarty et al., 2007).

Increased litter size is one of the biggest contributors to higher profits on lamb production. Crossbreeding with prolific sheep breeds is a way of increasing proportion of ewes having twins and triplets. However, lamb survival is an important issue in high litter size for sheep flocks. Davis et al. (1983) reported that as mean litter size increases above 1.7, the decline in single-bearing ewes is offset by an increase in triplet-bearing ewes. In studies investigating the survivability of lambs from mixed-age ewes, it was found that the lamb's birth weight is a strong driver of lamb survival (Yapi et al., 1990; Morel et al., 2008); and researchers reported that lambs weighing <3 kg at birth have a lower survival rate from birth to weaning (Nowak and Lindsay, 1992). Neonate survival is dependent on the coordinated expression of appropriate behaviors from both mother and lamb (Dwyer, 2003); and behavioral interactions are much more important for prolific sheep with higher litter size. Mora-Medina et al. (2016) have recently reviewed sensory recognition (olfaction, vision, vocalization, hearing and direct contact) in relation to the ewe-lamb bond and emphasized the study demonstrated by Dwyer et al. (2003) that malnutrition of pregnant ewes during the gestation period impairs attachment between ewes and lambs by affecting maternal behaviors expressed at birth.

Data mining and its application in animal husbandry was studied by

* Corresponding author.

https://doi.org/10.1016/j.compag.2021.105995
Received 20 February 2020; Received in revised form 7 January 2021; Accepted 8 January 2021
Available online 19 February 2021
0168-1699/

Wang et al. (2014). They underlined that animal husbandry management system structure is quite complex, with various problems faced by high volume and complex data; and some cannot establish a precise mathematical model.

For these problems, the application of data mining techniques can reflect a higher superiority, which is a powerful tool to solve such problems. Machine Learning (ML) and breeding share important objectives like prediction; and not surprisingly, several works have applied ML algorithms to genomic prediction (e.g., review Gonzalez-Recio et al., 2014). Sheep breeding use data sets and statistical techniques that qualify it for ML scope. In terms of lamb survival and factors affecting this very important parameter, ML methods can support us in creating predictive models by analyzing a large amount of data; and these methods can help us in decision-making. Machine learning researchers have developed sophisticated and effective algorithms which either complement or compete with the traditional statistical methods (Zupan et al., 2000).

The ability to study animal behaviour is important in many fields of science; and behavioral data represents large or open-ended data volumes which require machine learning techniques to automatically classify these large datasets into behavioural classes (Le Roux et al., 2017). In the scope of sheep breeding, ML algorithms have previously been used to detect basic behaviours (Fogarty et al., 2020a) such as mutually-exclusive behaviours (grazing, lying, standing, walking), active (or inactive) behaviour or detection of body posture (upright or prostrate). They reported that most effectively performing ML algorithms were Linear Kernel Support Vector Machine (SVM), Classification Tree (CART), and Linear Discriminant Analysis (LDA). Further studies have applied behaviour classification machine learning (ML) algorithms to accelerometer data to monitor changes in sheep behaviour around the time of lambing. It was aimed to facilitate the future development of algorithms based on ear tag accelerometer data for the detection of behavioural changes around the time of lambing in real-time or near-real-time (Fogarty et al., 2020b). More specific applications have analysed the relationships between serum lactoferrin concentrations and serum IgG concentrations in lambs (Gökçe et al., 2014) and automated detection of lameness in sheep using machine learning approaches (Kaler et al., 2020). These studies vary in their approach, they have differences in study purpose, maternal and offspring behavioral interactions, and dam and lamb intrinsic factors.

The study is divided into three parts. In the first one, we analyzed the behaviors of dam and lamb for target output of mothering ability for dam and successful sucking time for lambs. A comparison of classifiers was given in Table 2. In the second part of the study, we applied the PCA to reduce the dimensionality of the problem in order to check whether the results obtained with the whole set of variables are improved or not. The objective is to find the best classifier to predict lamb survival. In the last section, we tested the suitability of the different classifiers for lamb survival rate until weaning by using mothering ability and successful sucking time for lambs and excluded other mothering and lamb behavior data sets. The best two performers of the six classifiers that are applied for the prediction of mothering ability, lamb behavior and lamb survival at weaning is presented in Table 2. Results with whole sets of attributes were not improved with PCA; and classifiers were used without PCA reduction.

This paper provides an approach to machine learning algorithms in the behavioral (dam and offspring) and productive traits (dam age, dam live weight at lambing, dam breed, litter size, lamb birth weight and lamb sex) affecting lamb survival of native and crossbreed lambs produced in high altitudes and cold climate regions of Turkey. With this information, appropriate animal breeding and management programs can be formulated to reduce lamb mortality rates.

2. Materials and methods

2.1. Data

The data set included 193 spring born lambs out of which 106 lambs were sired by indigenous rams (n = 10), and 87 lambs were sired by Romanov Rams (n = 10); and 150 indigenous ewes (Awassi = 50, Morkaraman = 50, Tuj = 50) managed in a semi intensive system at the Ataturk University Experiment Station, Erzurum, Turkey. Erzurum is a province in the north-eastern Turkey with a high altitude (1757 m above sea level) and defined as a winter city according to World Winter Cities Association for Mayors (WWCAM). According to the Turkey General Directorate of Meteorology report, the average temperature in the three months of winter (December, January, and February) is −7.63C. Flock was housed in a semi open shed (one side of the shed is open), and were offered hay and concentrates to meet their nutritional requirements according to international (NRC, 2007) estimation during the experimental period. Lambs were born in April-May (spring); and starting approximately 3 days before expected parturition dates, ewes were kept under 24-h observation by at least two researchers. Mean temperature and relative humidity (provided by the Turkish Meteorological Institute) were of 8.3 C (min: 0.4 and max: 15.7 C) and 71%, respectively in April and May, which are the prime lambing months in the region. As it is described by Dwyer (2003), parturition was observed with minimum interruption, and assistance for lambing and care of the lambs was applied when needed. Sources of data on dam behavioral factors were used as stated by Emsen et al. (2012). Behavioral data on lambs were recorded in minutes from birth to knees, stands to udder, and to successful suck as previously described by Dwyer (2003). Lambs were weaned at 60 days of age. Factors included were dam body weight at lambing, age of dam, litter size at birth, maternal and lamb behaviors, and lamb sex.

2.2. Machine learning algorithms

Individual and cohort data were combined into an original dataset containing 1351 event records from 193 individual lambs and 750 event records from 150 individual ewes. This data set contained 15 unique variables, and several combinations, derived, and redundant variables. We categorized the variables used in this study into three groups: 1. Dam intrinsic variables; such as dam breed, age, weight at lambing, etc. 2. Lamb intrinsic variables; such as lamb sex, genotype, time to stand etc. 3. Dam to lamb interaction intrinsic variables; such as mothering ability, time to first touch to udder, etc. We chose Waikato Environment for Knowledge Analysis (WEKA®3.9.3) in order to analyse and explore the available data and to induce the data mining models (DMM) for mothering ability, lamb behavior and lamb survival. In WEKA, we used six machine learning algorithms for classification, namely: BayesNet, NaiveBayes, Multilayer Perceptron, Sequential Minimal Optimization (SMO), RandomForest and REPtree. Regarding model evaluation; we divided each of these datasets into two subsets. We used one subset to construct the classifier. This data set is called the train-set and used 70% of data set. The other set was used to evaluate the classifier using remaining (30%) part of data set. We also used 10-Fold Cross Validation technique to evaluate the model. With this technique we have one data set which we divided randomly into 10 parts. 9 of those parts are used for training and reserved one tenth for testing. We repeated this procedure 10 times, each time reserving a different tenth for testing.

Lamb survival was a binary classifying variable "survival until 60 days of age" ewe raised lambs. The attributes used for classification have been described (Table 1); and the target attribute was the lamb survival rate, which was obtained by summing the scores of mothering ability and lamb behavior.

Mothering ability was determined by considering a set of maternal care just after lambing and post lambing. Maternal care was defined as Superior (A): if ewe gets up within 3 min of lamb expulsion, vigorously

Table 1
Cohort animal level variables used in this study and their data descriptions.

Categories	Attributes	Descriptions	Response
Dam intrinsic	Dam Breed	Three breeds; Awassi, Morkaraman, Tuj	Mothering Ability, Lamb survival
Dam intrinsic	Grooming	Duration of licking lamb; A: >30 min, B: 15–30 min, C < 15 min	Mothering Ability, Lamb survival
Dam intrinsic	Dam age	Year; 1,2,3,4,5,6,7,8	Mothering Ability, Lamb survival
Dam intrinsic	Dam weight at lambing	38–79 kg	Mothering Ability, Lamb survival
Dam intrinsic	Litter size	Single, twin, triplets	Mothering Ability, Lamb behavior, Lamb survival
Dam intrinsic	Birth assistance	No assistance, minor assistance, pulled manually	Mothering Ability, Lamb behavior, Lamb survival
Lamb intrinsic	Lamb sex	Female, male	Mothering Ability, Lamb behaviour, Lamb survival
Lamb intrinsic	Lamb birth weight	Weight taken 24 h after lambing in order to allow bonding and reduce the risk of mismothering due to human intervention.	Lamb behavior, Lamb survival
Lamb intrinsic	Time to lamb on knee*	Lamb on chest, pushes up on knees, supporting part of body off the ground A < 10 min, B: 10–20 min, C > 20 min	Successful suck by lamb (Lamb behavior), Lamb survival
Lamb intrinsic	Time to lamb stands*	Lamb supports itself on all four feet for at least 5 s A < 15 min, B: 15–30 min, C > 30 min	Successful suck by lamb (Lamb behavior), Lamb survival
Lamb intrinsic	Lamb Genotype	Indigenous, crossbreed	Mothering Ability, Lamb behavior, Lamb survival
Dam and Lamb interaction intrinsic	Time to first touch to udder*	Lamb in parallel inverse position with head nudging ewe in udder region and first touch A < 20 min, B: 20–40 min, C > 40 min	Successful suck by lamb (Lamb behavior), Lamb survival
Dam and Lamb interaction intrinsic	Mothering ability	A: Superior mother, B: moderate mother, C: Poor mothering	Lamb survival
Dam and Lamb interaction intrinsic	Successful suck	Lamb has teat in its mouth, in correct position, appears to be sucking for at least 5 s A < 40 min, B: 40–80 min, C > 80 min, M: Colostrum given manually	Lamb survival

* Dwyer (2003).

grooms lamb and stands and facilitates suckling; Moderate (B): if ewe approaches lamb but does not initiate physical contact, circles when lamb attempts to suckle; and Poor (C): if ewe kicks/butts lamb and will not allow lamb to suckle. Lamb behaviour was the speed of reaching the udder and successful suckling.

Machine Learning (ML) and breeding share important objectives like prediction; and not surprisingly, several works have applied ML algorithms to genomic prediction (e.g., review Gonzalez-Recio et al., 2014). Furthermore, the data sets in Animal Breeding have traditionally been bigger than in many of its contemporary biological sciences; and developing efficient algorithms have been an important and relevant activity in our field. Animal Breeding does use big data sets and statistical techniques that fall within the ML scope, so Animal Breeding is machine learning, or at least a subset of the machine learning area. These methods support us in creating assumptions on the future by analysing a large amount of data for any practice and they help us in decision-making (Pérez-Enciso, 2017).

2.3. Decision trees

Decision tree classification is a learning process that recursively partitions a training dataset, and is then used to determine the appropriate class for each example within a test dataset (Zhang, 2012). Hutchinson and Gigerenzer (2005) described decision trees as simple and intuitive predictive models which make them a popular choice when decision rules are required. For example, the activity status of an animal can be classified by asking a series of simple yes/no questions (e.g. is the velocity > 0.01 m/s2? Nadimi et al., 2008). We evaluated two variations of decision trees classification algorithms: Random forests and REPTree.

2.4. Bayesian methods

In general, Bayesian classifiers estimate the conditional probability distributions of each attribute within the training dataset, and then assign cases within the test datasets to the class with the highest posterior probability using Bayes' Theorem (Sebastiani et al., 2005). Mcnamara et al. (2006) gave an overview of the theoretical concepts and biological contexts in Bayesian decision theory, which can be used to model animal behaviour; and they outlined some directions. We used BayesNet (BN) and Naïve Bayes (NB) in our study.

2.5. Artificial neural networks (ANN)/support vector machine

Artificial Neural Networks are defined by Yegnanarayana (1994) as a computational model inspired by the structure and performance of biological neural network. In ANN, there is a computer representation of knowledge that attempts to mimic the neural networks of the human body. Sanz et al. (2016) compared ML algorithms and used Multi Layer Perceptron (MLP), which is an ANN composed of a certain number of layers, conformed by a set of neurons. They described that MLP at least has an input layer, a hidden layer and an output layer. These layers compose a directed graph, since the neurons of each layer are connected to neurons of the next layer by means of weights. Artificial Neural networks predict outcomes based on relationships between variables that may be complex and multidimensional, and are well suited to our data structure, as they do not require a priori assumptions about the underlying data structure (Zhang, 2005). The neural network survival analysis has been employed to predict the survival time of a subject directly from the given inputs. Multilayer Perceptron method is proposed in the literature which employs the neural network method to solve survival analysis problems. Sanz et al. (2016) reviewed the classification techniques and used Support Vector Machines (SVM), reported by Cortes and Vapnik (1995). That is a classifier which constructs a hyperplane or set of hyperplanes in a high-dimensional space. In this paper, we considered only one SVM, which is Platt's Sequential Minimal Optimization Algorithm (SMO). Sanz et al. (2016) summarized SMO as a common method for solving the quadratic problem arising from SVMs. Platt's Sequential Minimal Optimization algorithm breaks the problem down into 2-dimensional sub-problems that may be solved analytically, eliminating the need for a numerical optimization algorithm.

2.6. Statistical methods

Statistical analyses were performed using SPSS (SPSS PC Ver.22; IBM© SPSS Inc., New York, U.S.). Data were presented as mean ± SD and two-tailed probabilities <0.05 with 95% CI were considered as significant. General Linear Model and Bayesian Anova were performed to show the possible effects of parameters on survival. Automatic Linear Modelling procedure in SPSS was used for data mining approach like regression trees, which utilizes a machine learning approach to find the best predictive model using the available data (Yang, 2013). This modelling is different from the traditional linear regression modelling approaches that require the user to find the best fitting model. The procedure accelerates the data analysis process through several automatic mechanisms such as automatic variable selection and automatic data preparation.

3. Results

We investigated the best indicators of mothering ability and time to successful suck by lambs. Classification algorithms for mothering ability performed better by using a different split of the database into training and test sets. While lamb behavior and lamb survival were best predicted with 10-fold cross-validation. The classifier's performance was relatively lower in predicting lamb behavior for successful sucking event time. RandomForest performed very well in their classification of the mothering ability while SMO was found best in predicting lamb behavior. Time to teeth touch was the first determinant and lambs having average time to first teeth touch was sub-classified into time to stands (ON FOOT) as shown in REPtree tree visualization (Fig. 1). Litter size was indicative for triplet born lambs that are manually given colostrum. Twins surprisingly performed better than single ones for their time to successful suck compared to those observed in single born litter. It is probably due to stimulatory effect of twins for sucking behaviors. Single born lambs with heavier birth weight (>5 kg) were slower to receive first milk from their mother.

Although RandomForest was found the best performer among decision trees, tree visualization is not available in WEKA software. Therefore, we presented REPtree tree visualization for mothering ability with 80.31% accuracy rate and 0.18 mean absolute error in Fig. 2. It can be clearly seen from tree visualization that grooming behavior is the first determinant for mothering ability; and it directly defines mothering ability if the duration is longer than 15 min. Dam breed played an important role when time for grooming is <15 min; and Awassi breed with shorter grooming was still classified as superior while the opposite case was observed for Morkaraman ewes. Tuj, known for its aggressiveness, showed dependency of their mothering ability on lamb birth weight. Lambs with lower birth weight (<2.85 kg) were taken good care of by Tuj ewes.

As Table 2 clearly shows, the best solutions were found by SMO and RandomForest for lamb survival; and NaiveBayes followed with 92.19% accuracy rate.

Moreover, statistical analysis (GLM) indicates that litter size, mothering ability and time to successful suck by lambs were a source of significant ($P < 0.05$) variation of lamb survival. Triplets had lower (66.4%) survival rates than those recorded for single (95.2%) and twin (89.5%) born lambs. In terms of mothering ability, recorded with at least two observers, it played an important role on lamb survival in which the best scored mother (91.8%) weaned significantly more lambs than average (81.8%), and poor mothering ability dams (77.5%). Coefficient estimate model with Automatic Linear Modelling is given in Fig. 3. It can be seen from the Figure that litter size, sucking time and dam age had a negative effect on lamb survival, while birth assistance and mothering ability had the opposite effect.

4. Discussion

Machine learning algorithms used in this study for predicting dam and lamb behavior led to meaningful conclusions and interpreted results correctly. Postnatal lambs' behaviors showed that twin born lambs are quicker to stand and suck than those born as singles or triplets. These results agree with observations reported by Dwyer and Morgan (2006), who indicated that both triplet and single lambs were slower to stand than twin lambs. However, O'Connor, et al. (1989) found that singles were more active, lying for less time and having a greater number of suckling attempts than twins. It has been documented that birth weight, gender and litter size may influence the time lambs take to stand and find the teat. Twins were reported slower than singletons in the expression of early behaviour, although some authors suggest that this twin effect is a function of reduced birth weight (Nowak and Poindron, 2006). On the other hand, REPtree shows that birth weight is a determinant of vigour of lambs that single lambs with lower that 5.05 kg birth weight are quicker to suck successfully than those above this birth weight. It has been stated that (Matheson et al., 2011) selection increases lamb birth weight which is phenotypically and genetically related to improved lamb vigour. The ideal birthweight range appears to be between 3.5 and 6 kg, with a slight variation between breeds, and maximum survival at approximately 4.5 kg (Oldham et al., 2011). Dwyer (2003) found that reduced lamb survival in heavier birthweight lambs was a direct result of dystocia and prolonged parturition. In this study, we found that lambs above 5 kg birth weight slow down in a

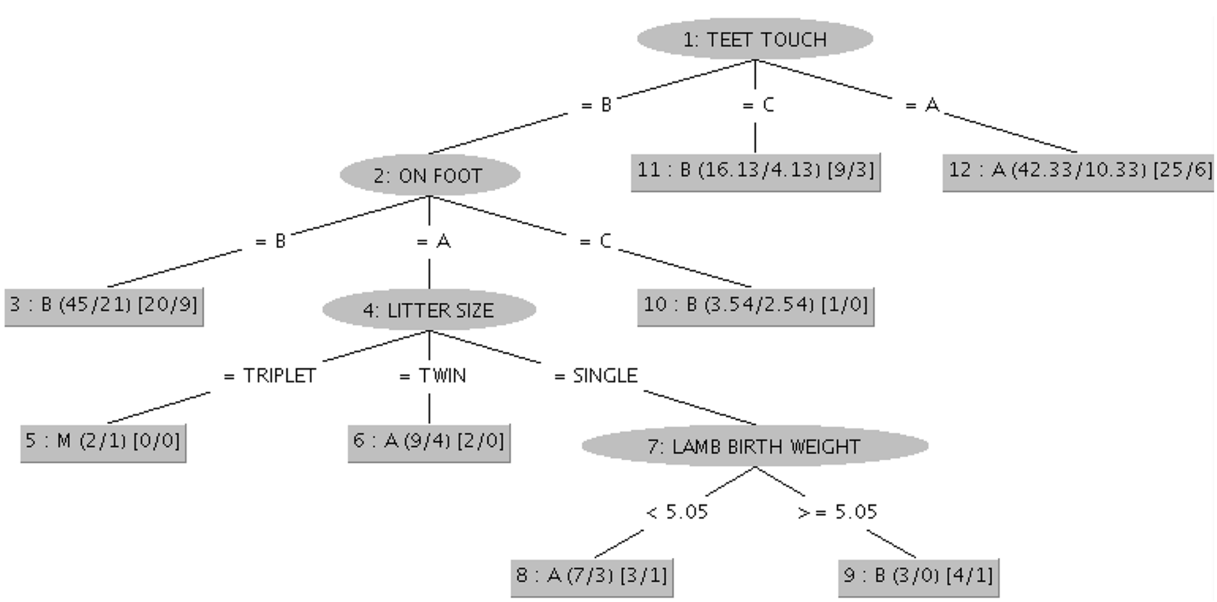

Fig. 1. Lamb behavioral REPtree classification for successful suck event.

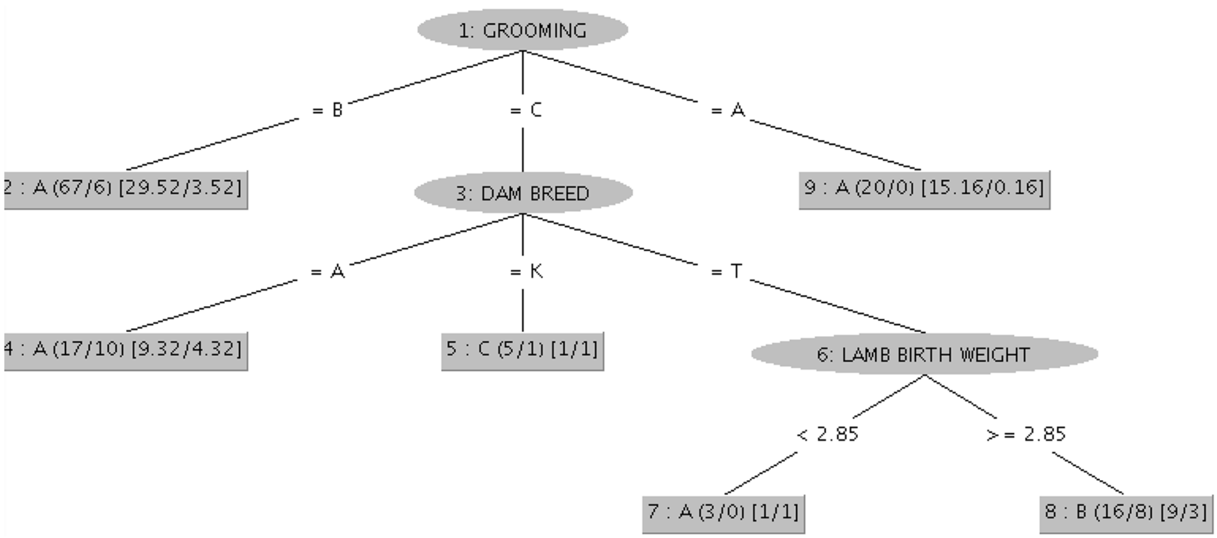

Fig. 2. Mothering ability REPtree classification tree.

Table 2
Classification algorithms and accuracy rates for mothering ability (MA) and lamb behavior (LB).

Traits	Algorithms	Correctly Classified Instances (%)	Mean absolute error
Mothering ability	RandomForest*	82.76	0.18
	REPTree*	80.31	0.18
Lamb behavior	SMO**	63.02	0.29
	REPTree**	61.45	0.24
Lamb survival rate	SMO**	93.75	0.06
	RandomForest**	93.23	0.10

* Test mode: split 70% train, remainder test.
** Test mode: 10-fold cross-validation.

successful suck event. Bangar et al. (2016) studied lamb mortality and its associated factors using survival analysis. The data records of 2168 lambs obtained from inventory and death registers for 10 years were subjected to the Cox proportional hazard model to determine the potential risk factors affecting lamb mortality. The Cox model pointed out the significant effect of birth weight on survival of lambs. It indicates that birth weight was priority criteria for survival of lambs during initial days. The survival analysis showed that the implementation of improved farm practices according to seasonal variation, flock structures (pregnant ewes, newly born lambs etc.) can potentially reduce economic losses due to lamb mortality. Southey et al. (2001) have compared survival analysis and logistic analysis for lamb mortality and concluded that the estimates due to survival analysis had lower standard errors than the logistic analysis. Our results showed that mean standard error was lower than those computed for mothering ability and lamb behaviour.

Lamb vigour is also associated with the ewe factor, which is called mothering ability. We tested factors influencing mothering ability and found that grooming of the lamb is the main determinant of mothering ability with REPtree classification algorithm. Maternal behaviour is described by Goursaud and Nowak (1999) as several characteristics of mother intense grooming of the lamb, low pitch bleating and standing to suckle within the first 6 h, which in turn facilitate sucking and

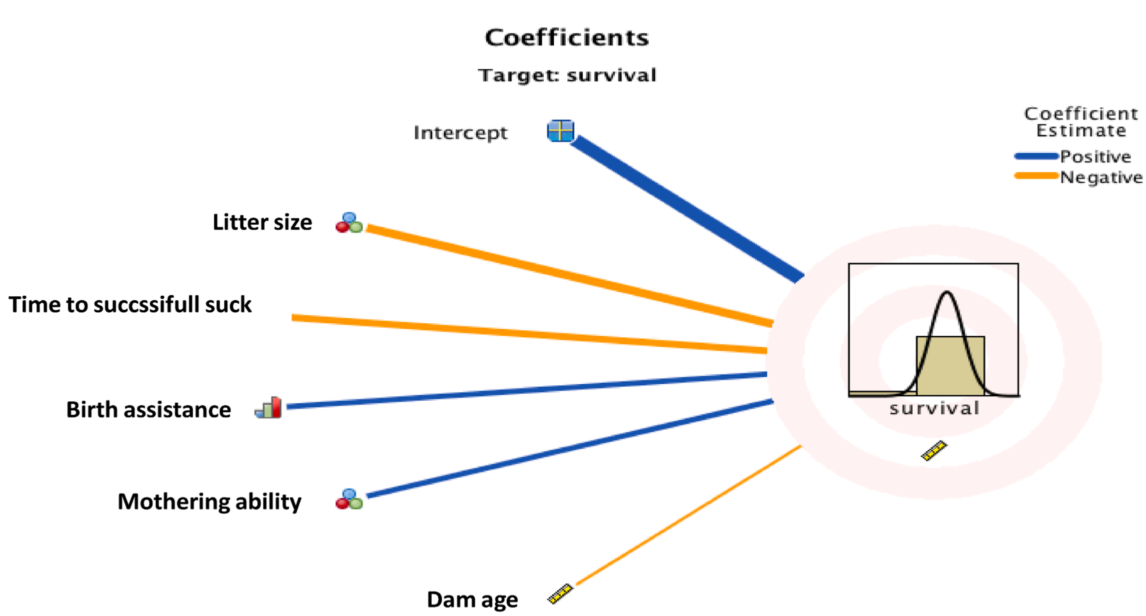

Fig. 3. Coefficient estimate of lamb survival with Automatic Linear Modelling.

recognition by the lamb and formation of the ewe-lamb bond. Mothering ability scored in this study was highly correlated with high levels of licking and grooming by mother right after birth. Duration of grooming was found to be a strong indicator of mothering ability as time spent for grooming longer than 15 min indicated the ideal dam. However, it needs to be considered as a series of dam behaviour as described by Dwyer (2013). Grooming should be accompanied by a specific maternal vocalisation (low-pitched bleating) and cooperation with the lamb attempting to find the udder and suckle, and an absence of rejection or avoidance of the lamb. As it is indicated by Cloete and Scholtz (1998), there are behavioural differences between breeds; and some of the maternal behaviours which play a part in lamb survival are likely to be under genetic control. Dam breed plays an important role if grooming behaviour is weak; and mothering ability of Awassi ewes was found unrelated to the duration of grooming behaviour for maternal care of its new-borns.

Automatic linear modelling used in this study to estimate coefficient of lamb survival and dam age, litter size and successful sucking event by lambs were found negatively correlated with lamb survival, while mothering ability and birth assistance had positive correlations. Lambs born from mature ewes (2 to 4 yrs. old) had higher survival rates than lambs born from aged ewes, which agrees with Lopez-Villalobos and Garrick (1999) who reported that ideal dam age for lamb survival was 3–5 years old in prolific Romney flock.

A range of machine learning algorithms have been used in animal behaviour studies such as linear discriminant analysis, k-means clustering, support vector machines, classification and regression trees, artificial neural networks and Random Forest (Alvarenga et al., 2016). Automated behavioural classification has the potential to improve health and welfare of the animals. Yet, there are no studies in precision livestock farming that have evaluated the effect of all these factors simultaneously. We observed a high percentage of correctly classified instances and Kappa for Random Forest in mothering ability. However, the Random Forest algorithm doesn't compute a P value, confidence intervals or regression coefficients. Therefore, it is not appropriate for hypothesis testing or ANOVA. The relative importance of variables measured by Random Forest is used to subjectively identify an indicator of general importance. Random Forest is an ensemble classifier in which each tree depends on the values of a random vector sampled independently, and all trees in the forest have the same distribution (Breiman, 2001). Nathan et al. (2011) used tri-axial acceleration data to identify behavioral modes of free-ranging animals and analysed the accuracies generated from linear discriminant analysis, support vector machines, classification and regression trees, Random Forest and artificial neural networks in both training and testing data sets. They also reported that Random Forest performed the best compared with the other methods based on Tukey's test.

The results of the BayesNet and SMO algorithms for lamb behaviour, compared to the other classifications, are very encouraging. SMO is a carefully organized algorithm, which has excellent computational efficiency. On the other hand, Bayes net theory provides one answer for a question by formulating a computational and graphical representation based on conditional probability between observed events. However, because of its way of computing, and use of a single threshold value in SMO computing, it can become inefficient. In the future, multiple threshold parameters can be used to improve the performance.

Ishwaran et al. (2008) reported that a promising method with respect to the statistical analysis of right censored survival data represents the machine learning method Random Forest. However, applications of RF in animal biology studies with complex data to identify biological markers promoting the development of lamb survival are still a rarity.

5. Conclusion

The goal of our investigation was to refine the set of criteria that could lead to better risk stratification in lamb mortality. To reach this goal, we started from the well-known factors affecting lamb survival including behavioral interactions and proceeded with the application of several machine learning schemes in order to perform a comparison between them. Our results showed that while all the machine learning algorithms we used do have predictive power in classifying lamb mortality into risk classes, classification trees clearly outperform all other methods for all traits included in this study.

In conclusion, the findings of the present research demonstrated that the RF and SMO methods and their implementation of backward algorithm represent a sensible complement to survival analysis. However, the verification of the present findings in external cohorts as well as the translation of the present findings for prevention strategies and management recommendations should be a matter for future research.

CRediT authorship contribution statement

B.B. Odevci: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **E. Emsen:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review & editing. **M.N. Aydin:** Supervision, Validation, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgements and have given us their written permission to be named. If we have not included an Acknowledgements, then that indicates that we have not received substantial contributions from non-authors.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2021.105995.

References

Aktaş, A.H., Dursun, Ş., Dogan, Ş., Kiyma, Z., Demirci, U., Halıcı, I.i., 2015. Effects of ewe live weight and age on reproductive performance, lamb growth, and survival in central anatolian merino sheep. Arch Anim Breed. 58, 451–459.

Alvarenga, F., Borges, I., Palkovič, L., Rodina, J., Oddy, V., Dobos, R., 2016. Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. Appl. Anim. Beh. Sci. 181, 91–99.

Bangar, Y.C., Pachpute, S.T., Nimase, R.G., 2016. The survival analysis of the potential risk factors affecting lamb mortality in deccani sheep. Journal of Dairy, Veternary and Animal Research 4, 266–270.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Cloete, S.W.P., Scholtz, A.J., 1998. Lamb survival in relation to lambing and neonatal behaviour in medium wool Merino lines divergently selected for multiple rearing ability. Aust. J. Exp. Agric. 38 (8), 801–811.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273e297.

Dwyer, C.M., 2003. Behavioural development in the neonatal lamb: effect of maternal birth related factors. Theriogenology 59, 1027–1050.

Davis, G.H., Kelly, R.W., Hanrahan, J.P., Rohloff, R.M., 1983. Distribution of litter size within flocks at different levels of fecundity. Proc. N. Z. Soc. Anim. Prod. 43, 25–28.

Dwyer, C., 2013. Maternal behaviour and lamb survival: from neuroendocrinology to practical application. Animal: Int. J. Animal Biosci. 8, 1–11. https://doi.org/10.1017/S1751731113001614.

Emsen, E., Diaz, C.A.G., Yaprak, M., Koycegiz, F., Kutluca, M., Emsen, H., 2012. Effect of inter-breed embryo transfer on lamb growing performance and survival. Reprod. Domest. Anim. 47 (1), 8–11. https://doi.org/10.1111/j.1439-0531.2008.01200.x.

Dwyer, C.M., Lawrence, A.B., Bishop, S.C., Lewis, M., 2003. Ewe-lamb bonding behaviors at birth are affected by maternal undernutrition in pregnancy. Br. J. Nutr. 89, 123–136.

Dwyer, C.M., Morgan, C.A., 2006. Maintenance of body temperature in the neonatal lamb: effects of breed, birth weight and litter size. J. Anim. Sci. 84, 1093–1101.

Fogarty, E.S., Swain, D., Cronin, G.M., Moraes, L.E., Bailey, D.W., Trotter, M., 2020a. Potential for autonomous detection of lambing using Global Navigation SatelliteSystem technology. Animal Production Science 60 (9), 1217–1226. https://doi.org/10.1071/AN18654.

Fogarty, S., Swain, D.L., Cronin, G.M., Moraes, L.E., Trotter, M., 2020b. Can accelerometer ear tags identify behavioural changes in sheep associated with parturition? Animal Reprod. Sci. 216 (106345), 0378–4320. https://doi.org/10.1016/j.anireprosci.2020.106345. ISSN 0378-4320.

Goursaud, A.P., Nowak, R., 1999. Colostrum mediates the development of mother preference by Newborn Lambs. Physiol. Behav. 67, 49–56. https://doi.org/10.1016/S0031-9384(99)00037-2.

González-Recio, O., Rosa, G.J.M., Gianola, D., 2014. Machine learning methods and predictive ability metrics for genome-wide prediction of complex traits. Livest. Sci. 166, 217–231.

Gökçe, E., Atakişi, O., Kırmızıgül, A.H., Ünver, A., Erdoğan, H.M., 2014. Passive immunity in lambs: serum lactoferrin concentrations as a predictor of IgG concentration and its relation to health status from birth to 12 weeks of life. Small Ruminant Res 116 (219–228), 2014. https://doi.org/10.1016/j.smallrumres.2013.11.006.

Hutchinson, J.M.C., Gigerenzer, G., 2005. Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. Behav. Process. 69 (2), 97e124. https://doi.org/10.1016/j.beproc.2005.02.019.

Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., 2008. Random survival forests. Ann. Appl. Statist. 2 (3), 841–860.

Kaler, J., Mitsch, J., Vazquez Diosdado, J., Bollard, N., Dottorini, T., Ellis, Keith, 2020. Automated detection of lameness in sheep using machine learning approaches: novel insights into behavioural differences among lame and non-lame sheep. R. Soc. Open Sci. https://doi.org/10.1098/rsos.190824.

Lopez-Villalobos, N., Garrick, D., 1999. Genetic parameter estimates for lamb survival in Romney sheep. Proc. New Zealand Soc. Animal Prod. 59, 121–124.

Mcnamara, J.M., Green, R.F., Olsson, O., 2006. Bayes' theorem and its applications in animal behaviour. Oikos. https://doi.org/10.1111/j.0030-1299.2006.14228.x.

Morel, P.C.H., Morris, S.T., Kenyon, P.R., 2008. Effects of birth weight on mortality in triplets born lambs. Austr. J. Exper. Agricult. 48, 984–987.

Matheson, S., Rooke, J., Mcilvaney, K., Jack, M., Ison, S., Bünger, L., Dwyer, C., 2011. Development and validation of on-farm behavioural scoring systems to assess birth assistance and lamb vigour. Animal: Int. J. Animal Biosci. 5, 776–783.

Mora-Medina, P., Orihuela-Trujillo, A., Arch-Tirado, E., Roldan-Santiago, P., Terrazas, A., Mota-Rojas, D., 2016. Sensory factors involved in mother-young bonding in sheep: a review. Veterinární medicína 61 (11), 595–611.

Moraes, A.B.D., Poli Cesar, H.E.C., Fischer, V., Fajardo, N.M., Aita, M.F., Porciuncula, G.C.D., 2016. Ewe maternal behavior score to estimate lamb survival and performance during lactation. Acta Scient Anim Sci. 38 (3), 327–332.

Nadimi, E.S., Søgaard, H.T., Bak, T., 2008. ZigBee-based wireless sensor networks for classifying the behaviour of a herd of animals using classification trees. Biosyst. Eng. 100 (2), 167e176. https://doi.org/10.1016/j.biosystemseng.2008.03.003.

Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., Getz, W.M., 2011. Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. J. Exp. Biol. 215, 986–996. https://doi.org/10.1242/jeb.058602.

Nowak, R.F., Lindsay, D.R., 1992. Discrimination of Merino ewes by their newborn lambs: important for survival? Appl. Anim. Behav. Sci. 34, 61–74.

Nowak, R., Poindron, P., 2006. From birth to colostrums: early steps leading to lamb survival. Repro. Nutr. Dev. 46, 431–446.

NRC, 2007. Nutrient requirements of small ruminants: sheep, goats, cervids and new world camelids. Washington, DC. The National Academies Press.

Oldham, M.C., Thompson, A.N., Ferguson, M.B., Gordon, D.J., Kearney, G.A., Paganoni, B.L., 2011. The birthweight and survival of Merino lambs can be predicted from the profile of liveweight change of their mothers during pregnancy. Animal Prod. Sci. 51 https://doi.org/10.1071/AN10155.

Pérez-Enciso, M., 2017. Animal breeding learning from machine learning. J. Animal Breed. Genet. 134, 85–86.

Safari, E., Fogarty, N.M., Gilmour, A.R., 2005. A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. Livest Prod Sci. 92, 271–289.

Sanz, J.A., Fernandes, A.M., Barrenechea, E., Silva, S., Santos, V., Gonçalves, N., Paternain, D., Jurio, A., Melo-Pinto, P., 2016. Lamb muscle discrimination using hyperspectral imaging: comparison of various machine learning algorithms. J. Food Eng. 174, 92–100.

Sebastiani, P., Abad, M.M., Ramoni, M.F., 2005. Bayesian Networks. In: Maimon, O., Rokach, L. (Eds.), Data Mining and Knowledge Discovery Handbook. Springer, New York.

Southey, B.R., Rodriguez-Zas, S.L., Leymaster, K.A., 2001. Survival analysis of lamb mortality in a terminal sire composite population. J. Anim. Sci. 79 (9), 2298–2306.

Tomaszyk, K., Dobek, A., Moliński, K., Gut, A., Szwaczkowski, T., 2014. Synergy factors in the analysis of lamb survival. Ital. J. Anim. Sci. 13, 735–740.

Wang, H., Ge Li, X., Yu Liu, T., Wu, Q., 2014. Data mining and its application in animal husbandry management system. Adv. Mater. Res. 926–930, 2525–2528. https://doi.org/10.4028/www.scientific.net/AMR.926-930.2525.

Yang, Hongwei, 2013. The Case for Being Automatic: Introducing the Automatic Linear Modeling (LINEAR) Procedure in SPSS Statistics. Multiple Linear Regression Viewpoints. 39.

Yapi, C.V., Boylan, W.J., Robinson, R.A., 1990. Factors associated with causes of preweaning lamb mortality. Prev. Vet. Med. 10, 145–152.

Zhang, G.P., 2005. Neural networks for data mining. In: Maimon, O., Rokach, L. (Eds.), Data Mining and Knowledge Discovery Handbook. Springer, New York.

Zhang, S., 2012. Decision tree classifiers sensitive to heterogeneous costs. J Syst Software 85, 771–779.

Zupan, B., Demsar, J., Kattan, M.W., Beck, R.J., Bratko, I., 2000. Machine learning for survival analysis: a case study on recurrence of prostate cancer. Artif. Intell. Med. 20 (1), 59–75.