Animal 15 (2021) 100391

Contents lists available at ScienceDirect

Animal

The international journal of animal biosciences

Machine learning approaches for the prediction of lameness in dairy cows

S. Shahinfar [a,1,*], M. Khansefid [a,1], M. Haile-Mariam [a], J. E. Pryce [a,b]

[a] Agriculture Victoria Research, AgriBio, Centre for AgriBioscience, Bundoora, Victoria 3083, Australia
[b] School of Applied Systems Biology, La Trobe University, Bundoora, Victoria 3083, Australia

ARTICLE INFO

Article history:
Received 23 February 2021
Revised 22 September 2021
Accepted 28 September 2021
Available online 17 November 2021

Keywords:
Animal welfare
Logistic regression
Naïve bayes
Neural networks
Precision dairy farming

ABSTRACT

Lameness is one of the costliest health problems, as well as a welfare concern in dairy cows. However, it is difficult to detect cows with possible lameness, or the ones that are at risk of becoming lame e.g. in the next week or so. In this study, we investigated the ability of three machine learning algorithms, Naïve Bayes (**NB**), Random Forest (**RF**) and Multilayer Perceptron (**MLP**), to predict cases of lameness using milk production and conformation traits. The performance of these algorithms was compared with logistic regression (**LR**) as the gold standard approach for binary classification. We had a total of 2 535 lameness scores (2 248 sound and 287 unsound) and 29 predictor features from nine dairy herds in Australia to predict lameness incidence. Training was done on 80% of the data within each herd with the remainder used as validation set. Our results indicated that in terms of area under curve of receiver operating characteristics, there were negligible differences between LR (0.67) and NB (0.66) while MLP (0.62) and RF (0.61) underperformed compared to the other two methods. However, the F1-score in NB (27%) outperformed LR (1%), suggesting that NB could potentially be a more reliable method for the prediction of lameness in practice, given enough relevant data are available for proper training, which was a limitation in this study. Considering the small size of our dataset, lack of information about environmental conditions prior to the incidence of lameness, management practices, short time gap between production records and lameness scoring, and farm information, this study proved the concept of using machine learning predictive models to predict the incidence of lameness *a priori* to its occurrence and thus may become a valuable decision support system for better lameness management in precision dairy farming.

Implications

An alert system to predict incidence of lameness is essential in precision dairy farming and to address increasing animal welfare awareness. Incidence of lameness is complex and very hard to predict. Data-driven predictive models are promising in predicting lameness *a priori*, if enough data are available to train the models. Our study proved the concept that naïve bayes classifier is a good candidate for predicting the incidence of lameness in modern dairy farms.

Introduction

Lameness, mastitis and fertility problems are the most prevalent health issues in dairy cattle which have detrimental effects on the welfare and economic performance of the cows (Bruijnis et al., 2010). The direct economic impact of lameness which includes the costs of treatment and early culling is evident. However, the effects of lameness on reduced milk yield and impaired fertility are less obvious but have a large contribution in total economic loss due to lameness incidence (Green et al., 2002; Huxley, 2013). The costs associated with lameness cases depend on many factors, specifically the stage of lactation and early detection of the disease (Cha et al., 2010). Consequently, the economic loss of subclinical cases could be as low as USD 18, whereas acute clinical cases may cost USD 95–225 depending on the underlying cause of the lameness (Ettema and Østergaard, 2006; Bruijnis et al., 2010; Cha et al., 2010).

Genetic improvement to reduce lameness is difficult because lameness is multifactorial, largely affected by farm management and environment (the reported heritability is 0.02–0.15) and is often under-recorded on dairy farms (Khansefid et al., 2021). Moreover, the identified cases are often recorded as a binary trait, without any information related to the severity or cause of the disease

* Corresponding author.
E-mail address: shahinfar@uwalumni.com (S. Shahinfar).
[1] Majid Khansefid and Saleh Shahinfar contributed equally to this paper.

https://doi.org/10.1016/j.animal.2021.100391
1751-7311/

(Abdelsayed et al., 2017; Heringstad et al., 2018). Some earlier studies have reported that conformation traits are correlated with lameness and could be used as predictive traits, or auxiliary traits, in genetic prediction models to improve the accuracy of lameness predictions. However, the achieved improvement in accuracy was marginal and not consistent across studies (Laursen et al., 2009; van der Linde et al., 2010; Heringstad et al., 2018). This could highlight the importance of some predominantly non-genetic factors affecting the prevalence of lameness, either at the cow level (milk production level, stage of lactation, calving month and breed) or at herd level (milking and hoof-trimming managements, footbathing, walking distance, quality of tracks and smoothness of concrete surfaces at farm, herd size and stocking density) (Espejo and Endres, 2007; Beggs et al., 2015; Ranjbar et al., 2016; O'Connor et al., 2020).

Considering the complexity of predicting the incidence of lameness, machine learning (**ML**) methods were shown to be promising for detecting the risk level of lameness at the herd level. For this purpose, routinely precollected records related to management, housing, production, reproduction, longevity and genetics merit at the farm level were assessed to be used as predictor features (Warner et al., 2020). Although the risk factors affecting lameness prevalence were reported in previous studies, the application of ML approaches at cow level to predict lameness has yet to be studied (Solano et al., 2015; O'Connor et al., 2020).

In precision dairy farming predicting lameness incidence at the cow level is vital for management efficiency and improving animal welfare. Detection of susceptible cows (high-risk category) is necessary for provision of better management, or earlier treatment to prevent the escalation to severe cases and therefore will culminate in reducing the prevalence and economic loss due to lameness. In recent years, there have been advancements in the use of accelerometers (O'Leary et al., 2020; Taneja et al., 2020) and video analysis via deep learning techniques to detect lameness cases on farm (Wu et al., 2020). Although these advanced systems can be highly accurate, they are costly to set up and operate. However, the prediction of lameness using routinely measured production data and type traits on farms can be more cost-effective and although comparatively less accurate, may serve as an early warning system for the farmer. Hence, the objective of this study was to evaluate the usefulness of ML approaches for prediction of lameness incidence at cow level and compare it with classic binary classification method.

Material and methods

Data

Lameness scores, milk production and conformation trait data used in this study were collected from all lactating cows in nine Australian dairy farms in spring 2018 (seven herds in Victoria and two herds in Tasmania). The lameness scoring was performed by trained classifiers after morning milking according to Dairy Australia's guidelines (Dairy Australia, 2015); where 0 = walking evenly, 1 = walking unevenly, 2 = moderate difficulty in walking and 3 = severe lameness. In this study, cows were classified as either sound (score 0) or unsound (score 1–3) because there were a limited number of cows with non-zero scores.

The milk production traits were test-day milk yield, fat, protein and lactose percentage as well as somatic cell count measured within the week of a lameness scoring visit with average (±SD) of −1.1(±3.6), meaning on average, milk record was taken a day before lameness scoring and only a single herd-test event was used in the analysis (i.e. we did not consider repeated records). Ideally, we would have the production records at least 1 week prior to the lameness scoring but limitation in the data collection did not allow it.

The conformation traits were scored only once in the first lactation by professional classifiers (as per standard dairy industry practice internationally). The traits evaluated were body condition score, Mammary System (udder depth, udder texture, median suspensory ligament, fore attachment, front teat placement, rear attachment height, rear attachment width, rear teat placement, teat length and front end height), Feet and Legs (foot angle, heel depth, rear legs – rear view and rear legs – side view), Rump (pin set, pin width, and loin strength) and Dairy Strength (stature, muzzle width, chest width, body depth, angularity, and bone quality) (DataGene, 2021). Furthermore, we also investigated the following potential predictors in our study; parity, age at calving (months), age at lameness scoring visit (months), days in milk at lameness scoring and test-day visit.

After quality control on the data, any column or row with more than 50% missing values was excluded (2 267 rows and 14 features were excluded). The remaining data comprised a total of 2 535 lameness scores (2 248 sound and 287 unsound) with records of lameness and 41 predictor features. Missing values for about 30% of lactose percentage were imputed using rfImpute procedure from randomForest packages in R (Liaw and Wiener, 2002). Feature selection was performed only on the training pool (described below) using a combination of mean reduction in Gini index and mean decrease in accuracy from randomForest package as follows; random forest was performed on the training set with 10 different ntree and four different mtry (40 times in total) and in each iteration, top 10 features for the mean decrease accuracy and top 10 for the mean decrease Gini were selected and saved in an array. Final selected features were unique value of this array. Feature selection method used in the current study was chosen because of ease of use, reliable results in previous studies and robustness to small training set which was the main limitation of our study. In total, 29 features were selected as predictors of lameness incidence. Three of these features were categorical; herd (nine levels); parity (1, 2, 3, 4, and 4+); and month of calving (12 levels). The summary statistics of the retained predictors used in this study is provided in Table 1. Breed, feet and legs, stature, udder texture, bone quality, muzzle width, pin set, rear set of leg, rear leg rear view, udder fore attachment, fore teat placement, and rear teat placement are the 12 features that were not selected for final models.

The distribution of lame cows among herds in this dataset is represented in Fig. 1.

Training and testing of models were performed as follows; first 20% of data within each herd was selected at random and set aside as testing set and the remaining 80% was considered as training pool. Second, 90% of the training pool was selected as a training set, where hyper-parameter tuning was done on 90% randomly selected from this training set (90% × 90% = 81%; i.e. each time the grid search was done on the randomised 81% of the training set). Any data split was done separately on the lame and sound animals within herds to keep the total proportion of lame to sound animals constant in all the steps. These steps were repeated 10 times and performance metrics – Accuracy (**ACC** = (TP + TN)/(TP + FP + TN + FN)), Precision (**PRE** = TP/(TP + FP)), True positive rate (**TPR** = TP/(TP + FN)), False positive rate (**FPR** = FP/(FP + TN)) and F1-Score (**F1**) were aggregated – where TP is true positive, TN is true negatives, FP is false positive, and FN is false negative. As the current study encountered an unbalanced classification problem, using F1-score (harmonic average of precision and recall) is a more suitable metrics for comparing different classification

Table 1
Summary statistics and feature importance* of the predictors selected for the prediction of lameness incidence in dairy cows.

Trait (Unit)	Mean	SD	Min	Max	Mean Decrease Accuracy %*	Mean Decrease Gini*
General Features						
Age at calving (month)	48.0	22.6	22	161	6.8	16.9
Age at lameness scoring (month)	52.7	23.0	23	162	7.3	15.4
DIM at lameness scoring (day)	142.2	146.5	1	485	6.9	18.1
DIM at milk recording (day)	118.6	106.4	2	314	6.1	21.8
Herd**	–	–	–	–	6.2	20.7
Month of calving**	–	–	1	12	4.5	16.7
Parity**	–	–	1	5	2.5	7.8
Production Traits						
Fat percentage (%)	3.86	0.97	1.13	9.84	1.7	25.3
Milk yield (kg)	27.14	8.95	3.20	6.06	6.8	22.0
Lactose percentage (%)	5.05	0.27	3.61	5.84	7.6	24.3
Protein percentage (%)	3.44	0.38	2.00	5.86	2.8	22.6
Somatic cell count (cells/ml)	129.7	460.8	1.00	9 590	2.5	16.3
Composite Type Traits						
Dairy Strength	11	1.7	3	16	3.4	11.6
Mammary System	10.3	1.4	5	14	3.7	9.5
Overall Type	9.9	1.3	1	13	1.4	10.1
Rump	10.9	2.1	1	16	1.0	14.2
Linear Type Traits						
Body condition score	3.57	0.75	1	8	3.6	14.7
Angularity	5.6	1.0	2	8	2.3	7.3
Body depth	6.0	1.1	2	9	1.7	8.9
Median suspensory ligament	6.4	1.1	2	9	2.6	7.7
Chest width	4.8	1.2	1	9	2.2	10.3
Foot angle	5.4	0.9	2	9	0.4	8.1
Heel depth	5.6	0.8	2	9	0.2	9.2
Loin strength	6.3	0.9	2	9	2.3	8.5
Pin width	6.3	1.3	2	9	2.2	12.9
Rear attachment height	6.3	1.3	1	9	1.5	10.8
Rear attachment width	5.6	1.3	1	9	4.0	9.7
Teat length	4.2	1.3	1	8	0.5	9.0
Udder depth	5.3	1.4	1	9	0.9	10.3

* Feature importance was only measured on the training set.
** These features were considered as factor.

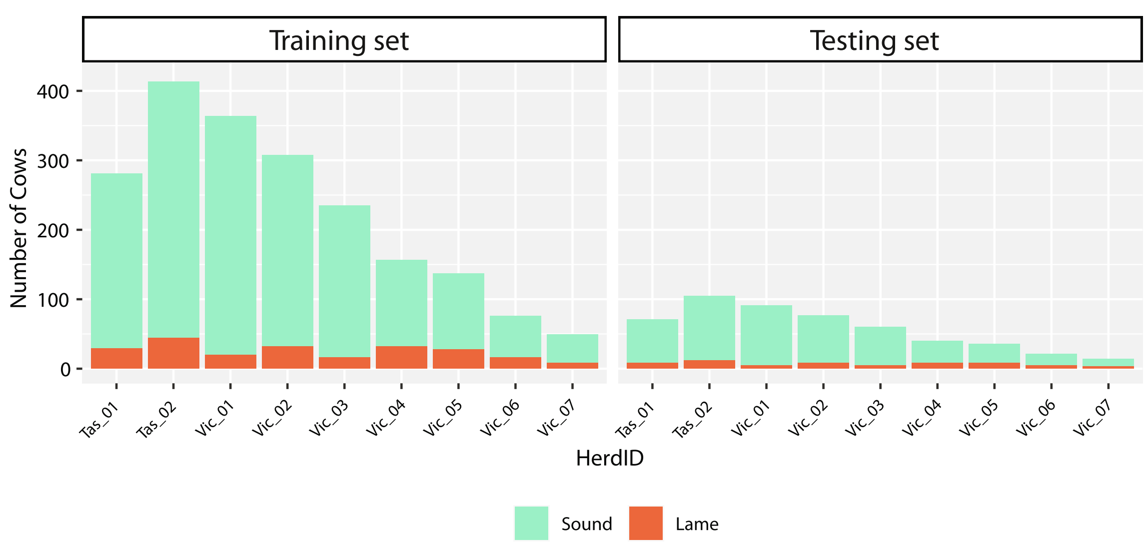

Fig. 1. Distribution of lameness cases of dairy cattle across herds in the training and testing sets.

algorithms. The F1-score ranges between 0 (total disagreement) and 1 (perfect classification) and it can be calculated as F1-score = (2 * PRE * TPR)/(PRE + TPR). The entire training and validation process was conducted in R v4.0.2 programming language (R-Core-Team, 2020).

Lameness prediction

Three machine learning methods were used in this study and their performance was compared with the classic binary prediction method, logistic regression (**LR**) (Cramer, 2002).

*Multilayer Perceptron (**MLP**)* is essentially a feedforward artificial neural network that takes a vector of real value inputs and calculates a sequential linear combination of these inputs into a set of appropriate outputs via its hidden layers and activation functions. It is well-suited for cases in which the instance space is complex, noisy and inter-correlated (Mitchell, 1997). Package 'h2o' in R was used for this purpose (LeDell et al., 2020).

*Naïve Bayes (**NB**)* is one of the most efficient and effective inductive learning algorithms for machine learning and data mining. It is a statistical classifier based on Bayes rule (Domingos and Pazzani, 1997), and it is the simplest form of Bayesian network, in which all features are independent, given the value of the outcome. Simplicity, computational feasibility, and robustness make this method suitable for practical use. Package 'e1071' in R was used for this purpose (Meyer et al., 2019).

*Random Forest (**RF**)* is one of the ensemble prediction methods in which predictor trees are trained on m bootstrap samples drawn from the training data. A random selection of a subset of features to generate each of those predictors is imbedded in the algorithm to break down the correlation between features (Ho, 1995; Breiman, 2001). Package 'randomForest' in R was used for this purpose (Liaw and Wiener, 2002).

Hyper-parameter tuning was conducted via a grid search on 90% of randomly selected training set in each iteration. Tuned hyper-parameters were number of trees to grow (20:200) and number of variables randomly sampled as candidates at each split (5:15) for randomForest. Laplace smoothing coefficient (1:10) was tuned for naïveBayes. For MLP, number of epochs (25, 50, 100, 500), number of hidden layers (2:3), activation functions (Rectifier, Tanh, Maxout), input dropout ratio (0, 0.1, 0.25, 0.5), and l1-norm coefficient (1e-5, 1e-7) were tuned.

Results and discussion

Lameness is a complex trait and it is difficult to predict its incidence. Table 2 shows model performance metrics for algorithms used in this study to predict incidence of lameness. There was no consistent winner among algorithms used to predict lameness. In terms of ACC and FPR, LR outperformed other ML algorithms at 0.88 and 0.00, respectively, indicating that the algorithm just classified almost all of cows as sound. In ML algorithms, MLP had the highest PRE at 0.41, however, it had a high SD in performance.

Considering TPR, it was NB that outperformed other methods, with a relatively low SD of 0.03. The naïve Bayes classifier showed the highest TPR and F1-score (0.27) and moderate precision relative to other algorithms tested in this study. Although in real life, different types of misclassification incur different cost and without considering those costs, identifying the optimum classifier is invalid (Shahinfar et al., 2015). In the absence of misclassification cost, we base our classifier selection on F1-score. In a study that also considered Australian dairy cattle, Bonfatti et al. (2020) attempted to predict lameness using mid-infrared spectra using partial least square discriminant analysis. They reported model performance metrics lower than reported here. Therefore, it seems that production and conformation traits are better choices as predictors when compared to mid-infrared spectra for the incidence of lameness in dairy farms, at least in Australia.

Fig. 2 presents receiver operating characteristic curve for the predictor algorithms used in this study. The Area under curve (**AUC**) indicates the overall performance of classifier asymptotically. In the current study, LR had the highest AUC at 0.67 followed by NB (AUC = 0.66). Warner et al. (2020) reported AUC = 0.73–0.75 for risk prediction of lameness at the herd level.

Lift analysis is defined as an increase in expected response in a selected subset of population (Sheng et al., 2014; Shahinfar et al., 2015). In the context of lameness prediction, we selected the top 25% of our data based on their predicted lameness probability by each algorithm. We calculated model performance metrics for this subset of data and compared it with the whole dataset (Table 2). Improvement in performance on top 25% of population was not obvious. F1-score and TPR just improved marginally, while ACC and FPR worsened and PRE did not change much.

According to previous studies, NB can accommodate dependencies between features very well and can often outperform more elaborate methods, such as rule learners and decision tree learners (Clark and Niblett, 1989; Cestnik, 1990). Also NB is quite intuitive and easy to understand, which is an advantage for this algorithm (Kononenko, 1990). Considering all the performance criteria, NB had significantly higher F1-score compared to LR, therefore, NB would be our recommended algorithm to predict the incidence of lameness. Nevertheless, NB still misclassified a large proportion of animals (i.e. high FPR and low PRE). This suboptimal performance can be due to the following: 1) our training dataset was limited in size and highly imbalanced; 2), lameness is indeed a very complex trait affected by genetics, environment, and management factors such as nutrition, production level, bedding, weather, walking track, laneway quality and pasture condition, while their causal-effect pathways still needed to be discovered and implemented in the prediction model (Ranjbar et al., 2016; O'Connor et al., 2020). Thus, for an accurate prediction of lameness incidence, a very comprehensive dataset of management factors affecting lameness (both at farm and animal level) is needed, which is often not accurately and consistently collected in dairy farms (O'Connor et al., 2020); 3), subjective definition of lameness by evaluators and

Table 2
Model performance metrics (± SD) for algorithms used in prediction of incidence of lameness in dairy cows for the whole testing set (ALL) compared with top 25% (Top 25%) of the test set.

Algorithm		ACC	Pre	TPR	FPR	F1
ALL						
	LR	0.88(±0.00)^{a}	0.09(±0.19)^{b}	0.00(±0.01)^{c}	0.00(±0.00)^{a}	0.01(±0.02)^{d}
	MLP	0.87(±0.01)^{a}	0.41(±0.30)^{a}	0.03(±0.03)^{c}	0.01(±0.01)^{ab}	0.05(±0.04)^{c}
	NB	0.83(±0.01)^{b}	0.28(±0.03)^{ab}	0.27(±0.03)^{a}	0.10(±0.02)^{c}	0.27(±0.02)^{a}
	RF	0.87(±0.01)^{a}	0.33(±0.09)^{a}	0.08(±0.03)^{b}	0.02(±0.01)^{b}	0.12(±0.05)^{b}
Top 25%						
	LR	0.80(±0.02)^{a}	0.39(±0.46)^{a}	0.02(±0.03)^{c}	0.01(±0.01)^{a}	0.04(±0.05)^{c}
	MLP	0.80(±0.05)^{a}	0.31(±0.24)^{a}	0.08(±0.08)^{c}	0.05(±0.06)^{ab}	0.11(±0.09)^{c}
	NB	0.60(±0.05)^{b}	0.28(±0.03)^{a}	0.68(±0.10)^{b}	0.42(±0.07)^{c}	0.40(±0.03)^{a}
	RF	0.77(±0.04)^{a}	0.33(±0.09)^{a}	0.20(±0.08)^{a}	0.10(±0.05)^{b}	0.24(±0.09)^{b}

Abbreviations: ACC: accuracy; PRE: precision; TPR: true positive rate; FPR: false positive rate; F1: F1-scores: LR: logistic regression; MLP: multilayer perceptron; NB: naïve bayes; RF: random forest.
The values with different superscript letters in each column (in either all or top 25% sections) are significantly different (P < 0.05) according to Tukey-HSD multiple comparison test.

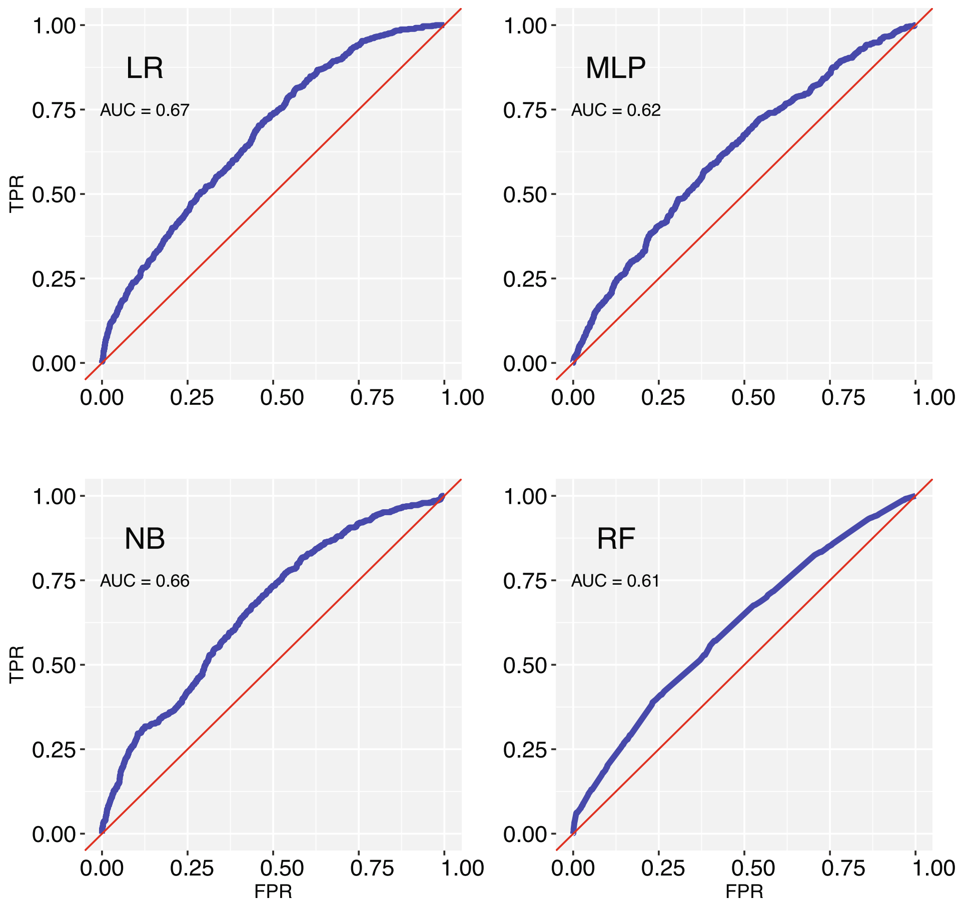

Fig. 2. Receiver operating characteristic plots for algorithms used in prediction of lameness in dairy cattle. Abbreviations: TPR: true positive rate; FPR: false positive rate; LR: logistic regression; MLP: multilayer perceptron; NB: naïve bayes; RF: random forest; AUC: area under curve.

farmers can cuntribute to ineficiency in modeling such a trait. Dairy Australia has provided guidelines for lameness scoring (Dairy Australia, 2015) but in practice, the scores can vary from one evaluator, or region, to another on what can be considered as lameness or not. In the current study because there was only one evaluator per herd, we assumed that it would be captured as part of the herd effect and therefore we did not correct for evaluator nor did we include it in our predictor set.

To best of our knowledge, this study is the first of its kind to predict the incidence of lameness at the cow level using predictive models and production and type phenotypic data.

Conclusion

Prediction of the incidence of lameness in dairy cattle is a difficult task. Multiple environmental effects influence lameness and their interactions and causal-effect pathways are often not considered in lameness prediction. Prediction of the incidence of lameness at the cow level is possible with Naive Bayes classifier and logistic regression. Our study was limited by a lack of comprehensive data. Although the classification performance was suboptimal in our study, we expect additional information at the herd level, such as grazed vs housed, nutrition, the distance cows walk and weather may improve the prediction accuracy. Nevertheless, this study proved the concept of predicting lameness on an individual basis for precise management in dairy farms. Further research and development are required to develop a robust predictive model for lameness that can perform on the acceptable scale for the industry. A more comprehensive large-scale data collection on traits related to the lameness are recommended.

Ethics statement

Not applicable.

Data and model availability statement

None of the data were deposited in an official repository.

Author ORCIDs

Saleh Shahinfar: https://orcid.org/0000-0003-0730-7577
Majid Khansefid: https://orcid.org/0000-0002-5091-7293
Mekonnen Haile-Mariam: https://orcid.org/0000-0001-5476-7475
Jennie Pryce: https://orcid.org/0000-0002-1397-1282

Author contributions

Majid Khansefid prepared the data. **Saleh Shahinfar** carried out the machine learning analysis. **Majid Khansefid** and **Saleh Shahinfar** contributed equally to the writing of the manuscript. **Mekonnen Haile-Mariam** and **Jennie Pryce** managed the project and contributed to writing of the manuscript.

Declaration of interest

Authors declare no conflict of interest.

Acknowledgements

The staff of DataGene (Bundoora, Victoria, Australia) are gratefully acknowledged for providing the data used in this study.

Financial support statement

The authors thank DairyBio, jointly funded by Dairy Australia (Melbourne, Australia), The Gardiner Foundation (Melbourne, Australia), and Agriculture Victoria (Melbourne, Australia), for funding this project.

References

Abdelsayed, M., Haile-Mariam, M., Pryce, J.E., 2017. Genetic parameters for health traits using data collected from genomic information nucleus herds. Journal of Dairy Science 100, 9643–9655. https://doi.org/10.3168/jds.2017-12960.

Beggs, D.S., Fisher, A.D., Jongman, E.C., Hemsworth, P.H., 2015. A survey of Australian dairy farmers to investigate animal welfare risks associated with increasing scale of production. Journal of Dairy Science 98, 5330–5338.

Bonfatti, V., Ho, P.N., Pryce, J.E., 2020. Usefulness of milk mid-infrared spectroscopy for predicting lameness score in dairy cows. Journal of Dairy Science 103, 2534–2544. https://doi.org/10.3168/jds.2019-17551.

Breiman, L., 2001. Random forest. Machine Learning 45, 5–32.

Bruijnis, M.R.N., Hogeveen, H., Stassen, E.N., 2010. Assessing economic consequences of foot disorders in dairy cattle using a dynamic stochastic simulation model. Journal of Dairy Science 93, 2419–2432. https://doi.org/10.3168/jds.2009-2721.

Cestnik, B., 1990. Estimating probabilities: A crucial task in machine learning. In: Book of abstracts of the 9th European Conference on Artificial Intelligence, 6-10 August 1990, Stockholm, Sweden, p. 147–149.

Cha, E., Hertl, J.A., Bar, D., Gröhn, Y.T., 2010. The cost of different types of lameness in dairy cows calculated by dynamic programming. Preventive Veterinary Medicine 97, 1–8. https://doi.org/10.1016/j.prevetmed.2010.07.011.

Clark, P., Niblett, T., 1989. The CN2 induction algorithm. Machine Learning 3, 261–283.

Cramer, J.S., 2002. The Origins of Logistic Regression: Tinbergen Institute Discussion Papers. No 02-119/4, Tinbergen Institute. http://dx.doi.org/10.2139/ssrn.360300. Retrieved on 16 January 2021 from https://ssrn.com/abstract=360300.

Dairy Australia, 2015. Healthy Hooves Investigation Pack Healthy Hooves: Investigation Pack. Dairy Australia Ltd., Southbank, Victoria, Australia.

DataGene, 2021. Type ABVs explained – Genetics Backgrounder # 2. Retrieved on 15 March 2021 from https://datagene.com.au.

Domingos, P., Pazzani, M.J., 1997. On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning 29, 103–130.

Espejo, L.A., Endres, M.I., 2007. Herd-level risk factors for lameness in high-producing holstein cows housed in freestall barns. Journal of Dairy Science 90, 306–314. https://doi.org/10.3168/jds.S0022-0302(07)72631-0.

Ettema, J.F., Østergaard, S., 2006. Economic decision making on prevention and control of clinical lameness in Danish dairy herds. Livestock Science 102, 92–106. https://doi.org/10.1016/j.livprodsci.2005.11.021.

Green, L.E., Hedges, V.J., Schukken, Y.H., Blowey, R.W., Packington, A.J., 2002. The impact of clinical lameness on the milk yield of dairy cows. Journal of Dairy Science 85, 2250–2256. https://doi.org/10.3168/jds.S0022-0302(02)74304-X.

Heringstad, B., Egger-Danner, C., Charfeddine, N., Pryce, J.E., Stock, K.F., Kofler, J., Sogstad, A.M., Holzhauer, M., Fiedler, A., Müller, K., Nielsen, P., Thomas, G., Gengler, N., de Jong, G., Ødegård, C., Malchiodi, F., Miglior, F., Alsaaod, M., Cole, J.B., 2018. Invited review: Genetics and claw health: Opportunities to enhance claw health by genetic selection. Journal of Dairy Science 101, 4801–4821. https://doi.org/10.3168/jds.2017-13531.

Ho, T.K., 1995. Random decision forest. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, 14-16 August 1995, Montreal, QC, Canada, pp. 278–282.

Huxley, J.N., 2013. Impact of lameness and claw lesions in cows on health and production. Livestock Science 156, 64–70. https://doi.org/10.1016/j.livsci.2013.06.012.

Kononenko, I., 1990. Comparison of inductive and naive Bayesian learning approaches to automatic knowledge acquisition. In: Current Trends in Knowledge Acquisition (ed. Wielinga, B., Boose, J., Gaines, B., Schreiber, G. and van Someren, M.), IOS Press, Amsterdam, the Netherlands, pp. 190–197.

Khansefid, M., Haile-Mariam, M., Pryce, J.E., 2021. Including milk production, conformation, and functional traits in multivariate models for genetic evaluation of lameness. Journal of Dairy Science 104, 10905–10920. https://doi.org/10.3168/jds.2020-20074.

Laursen, M.V., Boelling, D., Mark, T., 2009. Genetic parameters for claw and leg health, foot and leg conformation, and locomotion in Danish Holsteins. Journal of Dairy Science 92, 1770–1777. https://doi.org/10.3168/jds.2008-1388.

LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T., Nykodym, T., Aboyoun, P., Kurka, M., Malohlava, M., 2020. h2o: R Interface for the 'H2O' Scalable Machine Learning Platform. R package version 3.30.0.1. Retrieved on 16 February 2021 from https://CRAN.R-project.org/package=h2o.

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2019. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-3. Retrieved on 16 February 2021 from https://CRAN.R-project.org/package=e1071.

Mitchell, T.M., 1997. Artificial neural networks. Machine Learning. McGraw-Hill, New York.

O'Connor, A.H., Bokkers, E.A.M., de Boer, I.J.M., Hogeveen, H., Sayers, R., Byrne, N., Ruelle, E., Shalloo, L., 2020. Associating mobility scores with production and reproductive performance in pasture-based dairy cows. Journal of Dairy Science 103, 9238–9249. https://doi.org/10.3168/jds.2019-17103.

O'Leary, N.W., Byrne, D.T., O'Connor, A.H., Shalloo, L., 2020. Invited review: Cattle lameness detection with accelerometers. Journal of Dairy Science 103, 3895–3911. https://doi.org/10.3168/jds.2019-17123.

O'Connor, A.H., Bokkers, E.A.M., de Boer, I.J.M., Hogeveen, H., Sayers, R., Byrne, N., Ruelle, E., Engel, B., Shalloo, L., 2020. Cow and herd-level risk factors associated with mobility scores in pasture-based dairy cows. Preventive Veterinary Medicine 181,. https://doi.org/10.1016/j.prevetmed.2020.105077 105077.

R-Core-Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ranjbar, S., Rabiee, A.R., Gunn, A., House, J.K., 2016. Identifying risk factors associated with lameness in pasture-based dairy herds. Journal of Dairy Science 99, 7495–7505. https://doi.org/10.3168/jds.2016-11142.

Shahinfar, S., Guenther, J.N., Page, C.D., Kalantari, A.S., Cabrera, V.E., Fricke, P.M., Weigel, K.A., 2015. Optimization of reproductive management programs using lift chart analysis and cost-sensitive evaluation of classification errors. Journal of Dairy Science 98, 3717–3728. https://doi.org/10.3168/jds.2014-8255.

Sheng, V.S., Gu, B., Fang, W., Wu, J., 2014. Cost-sensitive learning for defect escalation. Knowledge Base Systems 66, 146–155.

Solano, L., Barkema, H.W., Pajor, E.A., Mason, S., LeBlanc, S.J., Zaffino Heyerhoff, J.C., Nash, C.G.R., Haley, D.B., Vasseur, E., Pellerin, D., Rushen, J., de Passillé, A.M., Orsel, K., 2015. Prevalence of lameness and associated risk factors in Canadian Holstein-Friesian cows housed in freestall barns. Journal of Dairy Science 98, 6978–6991. https://doi.org/10.3168/jds.2015-9652.

Taneja, M., Byabazaire, J., Jalodia, N., Davy, A., Olariu, C., Malone, P., 2020. Machine learning based fog computing assisted data-driven approach for early lameness detection in dairy cattle. Computers and Electronics in Agriculture 171,. https://doi.org/10.1016/j.compag.2020.105286 105286.

van der Linde, C., de Jong, G., Koenen, E.P.C., Eding, H., 2010. Claw health index for Dutch dairy cattle based on claw trimming and conformation data. Journal of Dairy Science 93, 4883–4891. https://doi.org/10.3168/jds.2010-3183.

Warner, D., Vasseur, E., Lefebvre, D.M., Lacroix, R., 2020. A machine learning based decision aid for lameness in dairy herds using farm-based records. Computers and Electronics in Agriculture 169, 105193.

Wu, D., Wu, Q., Yin, X., Jiang, B., Wang, H., He, D., Song, H., 2020. Lameness detection of dairy cows based on the YOLOv3 deep learning algorithm and a relative step size characteristic vector. Biosystems Engineering 189, 150–163. https://doi.org/10.1016/j.biosystemseng.2019.11.017.