Proc. Assoc. Advmt. Anim. Breed. Genet: 24: 369-372

EXPLORING MACHINE LEARNING APPROACHES TO PREDICT THE INCIDENCE OF LAMENESS IN DAIRY COWS

S. Shahinfar[1*], M. Khansefid[1], M. Haile-Mariam[1] and J.E. Pryce[1,2]

[1]Agriculture Victoria Research, AgriBio, Centre Bundoora, Victoria 3083, Australia
[2]School of Applied Systems Biology, La Trobe University, Bundoora, Victoria 3083, Australia

ABSRTACT

In this study, we investigated the ability of three machine learning algorithms, Naïve Bayes (NB), Random Forest (RF) and Multilayer Perceptron (MLP), in the prediction of cases of lameness. Performance of these algorithms were compared with logistic regression (LR) as the gold standard approach for binary classification. There were negligible differences between LR, NB and RF, while MLP underperformed the other three methods. However, the F1-score in NB (22%) outperformed LR (11%), suggesting NB potentially could be a more reliable method for prediction of lameness in practice if there is enough relevant data available for proper training.

INTRODUCTION

Lameness along with 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 are evident. However, the effects of lameness on reduced milk yield and impaired fertility are less obvious but have large contribution in total economic loss due to lameness incidence (Green *et al.* 2002; Huxley 2013).

Genetic improvement to reduce lameness is difficult because the accuracy of lameness predictions is often low. Considering the complexity of prediction of lameness incidence, machine learning (ML) was shown to have promise to detect the risk level of lameness at the herd level according to 20 routinely pre-collected farm based records related to management, housing, production, reproduction, longevity and genetics merits (Warner *et al.* 2020).

Predicting lameness incidence at the cow-level can help farmers detect susceptible cows (high risk category). Hence, the objective of this study was to evaluate the usefulness of ML approaches in the prediction of lameness incidence and compare it with classic binary classification method.

MATERIAL AND METHODS

Data. Lameness scores, milk production and conformation traits data were collected from 11 Australian dairy farms in spring 2018. The lameness scoring was performed by trained classifiers after morning milking according to Dairy Australia guidelines[1]; where 0=walking evenly, 1=walking unevenly, 2=moderate difficulty in walking and 3=severe lameness. In this study, cows were classified to either sound (score 0) or unsound (score 1-3) group 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 (SCC) measured within a week before the lameness scoring visit. Further, we also investigated the following potential predictors in our study; breed, parity, age at calving (in months), age at lameness scoring visit (in months), days in milk (DIM) at lameness scoring and test-day visit.

Any column or row with more than 50% missing values was excluded. The remaining data comprised 2,640 cows in 11 herds with records of lameness and 42 predictor features. Missing

---

[1] https://www.dairyaustralia.com.au/dairytas/animal-management-and-milk-quality/animal-health/lameness

values for about 30% of lactose percentage and parity number were imputed using rfImpute procedure from randomForest packages in R (Liaw and Wiener 2002). Feature selection was performed using combination of mean reduction in Gini index and mean decrease in accuracy from randomForest Package combined with the potential predictor traits reported in previous literature (Solano *et al.* 2015; Ranjbar *et al.* 2016; O'Connor *et al.* 2020). In total 31 features were selected as predictors of lameness incidence. Four of these features were categorical; breed (Holstein, Jersey, Holstein × Jersey and Holstein × non-Jersey crossbreds); herd (11 levels); parity (1, 2, 3, 4, and 4+); and month of calving (MOC; 12 levels). The summary statistics of the rest of the predictors used in this study is provided in Table 1.

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). *Multilayer Perceptron (MLP)* is a feedforward artificial neural network that calculates a sequential linear combination of inputs into a set of appropriate outputs via its hidden layers and activation functions (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). 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 bootstrap samples drawn from the training data (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 50% of randomly selected data. Training and testing of models were performed using 10-fold cross validation and repeated 10 times. Performance metrics were aggregated. The entire training and validation process was conducted in R v4.0.2 programming language (R-Core-Team 2020).

RESULTS AND DISCUSSION

Table 2 shows model performance metrics for algorithms used in this study to predict incidence of lameness. There was not a consistent best performer among algorithms used to predict lameness. In terms of accuracy (ACC) and precision (PRE) LR outperformed the ML algorithms at 0.86 and 0.28 respectively. Among the ML algorithms, MLP had the lowest false positive rate (FPR) at 0.04, however, it had a high standard deviation in performance. Considering true positive rate (TPR), it was NB that outperformed the other methods with a relatively low standard deviation (0.26). As the current study encountered an unbalanced classification problem (unbalanced numbers of lame to sound cows), using F1 score (harmonic average of precision and recall) was a more suitable metrics for comparing different classification algorithm. The naïve Bays classifier had the highest TPR and F1 score (0.22) and moderate precision relative to other tested algorithms. In real life different types of misclassification error varies in cost, without considering those costs, identifying the optimum classifier is not possible (Shahinfar *et al.* 2015). In the absence of misclassification cost, we base our classifier selection on F1-score.

The Area under ROC curve (AUC) indicates the overall performance of classifier asymptotically. In the current study LR had the highest AUC at 0.65 followed by NB (AUC= 0.63). Warner *et al.* (2020), reported AUC = 0.73-0.75 for risk prediction of lameness at the herd level.

Considering all the performance criteria, NB had significantly higher F1 Score compare to LR, therefore NB would be the recommended algorithm to predict incidence of lameness. Nevertheless, NB still misclassified a large proportion of animals (i.e. high FPR and low PRE). This sub-optimal performance can be firstly due to the fact that the training data set was limited in size and highly imbalanced; and secondly, 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 (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).

Table 1. Summary of data used in this study

trait	mean	sd	min	max	mean decrease accuracy	mean decrease Gini
Age at calving	47.71	22.33	22	161	6.44	22.4
Age at lameness scoring	52.33	22.72	23	162	7.36	21.9
BCS	3.59	0.76	1.0	8.0	3.28	15.44
Dairy strength	11.02	1.67	3	16	3.68	14.3
Feet & legs	10.52	1.55	3	15	2.35	12.83
Mammary system	10.28	1.35	5	14	1.77	11.02
Overall type	9.88	1.32	1	13	2.98	10.81
Rump	10.88	2.12	1	16	2.1	17.57
DIM at lameness scoring	138.68	145.25	1	485	8.37	22.45
DIM at milk test-day	115.75	105.56	2	314	8.67	27.06
Fat %	3.85	0.97	1.13	9.84	4.3	27.72
Lactose%	5.05	0.26	3.61	5.84	7.26	29.04
Angularity	5.55	0.98	2	8	4.62	10.43
Body depth	6.03	1.09	2	9	2.36	9.98
Bone quality	6.8	1.11	1	9	0.22	12.24
Median suspensory	6.4	1.07	2	9	3.19	9.86
Foot angle	5.36	0.92	2	9	1.41	10.29
Heel depth	5.64	0.83	2	9	1.5	11
Loin strength	6.34	0.91	2	9	2.52	10.77
Pin width	6.24	1.32	1	9	2.11	14.5
Rear attachment width	5.63	1.31	1	9	5.1	12.38
Rear legs - rear view	5.92	1.03	1	9	1.3	13.99
Stature	6.29	1.51	1	9	1.46	12.19
Udder depth	5.31	1.39	1	9	2.13	13.36
Milk yield	27.25	8.93	32	606	8.23	27.08
Protein %	3.44	0.38	2.00	5.86	3.34	27.41
SCC	129.84	477.43	1	9590	2.88	26.81
Breed	*	*	*	*	1.21	2.1
Herd	*	*	*	*	12.18	27.86
Parity	*	*	*	*	3.59	9.15
MOC	*	*	*	*	5.46	19.39

'*'these features were considered as factor

Table 2. Model performance metrics for algorithms used in prediction of incidence of lameness in dairy cows. ACC=Accuracy; PRE=Precision; TPR=True Positive Rate; FPR=False Positive Rate; F1 = F1 scores

algorithm	ACC	PRE	TPR	FPR	F1-score
LR	**0.86(0.032)[ab]**	**0.28(0.072)[a]**	0.09(0.086)[b]	**0.04(0.046)[ab]**	0.11(0.072)[b]
MLP	**0.88(0.022)[a]**	0.17(0.173)[b]	0.03(0.038)[c]	**0.02(0.029)[a]**	0.04(0.045)[c]
NB	0.80(0.036)[c]	0.20(0.015)[b]	**0.26(0.070)[a]**	0.14(0.048)[c]	**0.22(0.020)[a]**
RF	0.84(0.057)[b]	**0.22(0.097)[ab]**	0.13(0.130)[b]	0.07(0.080)[b]	0.12(0.084)[b]

* The values with different superscript letters in each column are significantly different ($p<0.05$) according to Tukey-HSD multiple comparison test.

CONCLUSION

Prediction of 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 incidence of lameness on the cow level is possible with Naive Bayes classifier and logistic regression. Lack of a comprehensive dataset was the main limitation of this study. Although the classification performance was suboptimal in our study, we expect additional information on the herd level such as bedding, nutrition, and weather will improve prediction accuracy. Nevertheless, this study provided proof of concept for prediction of lameness at the cow level.

REFERENCES

Breiman L. (2001) *Mach. Learn.* **45:** 5.

Bruijnis M.R.N., Hogeveen H., Stassen E.N. (2010) *J. Dairy Sci.* **93:** 2419.

Domingos P., Pazzani M.J. (1997) *Machine Learning* **29:** 103.

Green L.E., Hedges V.J., Schukken Y.H., Blowey R.W., Packington A.J. (2002) *J. Dairy Sci.* **85:** 2250.

Ho T.K. (1995) *in Proc. 3rd ICDAR, Montreal, QC, Canada* 278.

Huxley J.N. (2013) *Livestock Science* **156:** 64.

LeDell E., Gill N., Aiello S., Fu A., Candel A., Click C., Kraljevic T., Nykodym T., Aboyoun P., Kurka M., Malonhava, M. (2020)

Liaw A., Wiener M. (2002) *R News* **2:** 18.

Meyer D., Dimitriadou E., Hornik K., Weingessel A., Leisch F. (2019)

Mitchell T.M., McGraw-Hill: New York. (1997)

O'Connor A.H., Bokkers E.A.M., de Boer I.J.M., Hogeveen H., Sayers R., Byrne N., Ruelle E., Shalloo L. (2020) *J. Dairy Sci.* **103:** 9238.

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) *Preventive Veterinary Medicine* **181:** 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) *J. Dairy Sci.* **99:** 7495.

Shahinfar S., Guenther J.N., Page C.D., Kalantari A.S., Cabrera V.E., Fricke P.M., Weigel K.A. (2015) *J. Dairy Sci.* **98:** 3717.

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) *J. Dairy Sci.* **98:** 6978.

Warner D., Vasseur E., Lefebvre D.M., Lacroix R. (2020) *Comput. Electron. Agric.* **169:** 105193.