VETERINARY HYGIENE, SANITATION AND EXPERTISE

DOI https://doi.org/10.30525/978-9934-26-350-7-20

DYSTOCIA DETECTION IN SIMMENTAL CATTLE USING THE K-NEAREST NEIGHBOR METHOD

ВИЯВЛЕННЯ ДИСТОЦІЇ У ВРХ СИММЕНТАЛЬСЬКОЇ ДОПОМОГИ МЕТОД К-БЛИЖЧОГО СУСІДА

Zaborski D.
Doctor of Agriculture,
Associate Professor at the Department of Ruminants Science,
West Pomeranian University of Technology
Szczecin, Poland

Заборський Д.
доктор сільського господарства,
доцент кафедри жуйних тварин
Західнопоморський технологічний університет
м. Щецин, Польща

Stadnytska O. I.
Doctor of Agriculture, Senior Research Associate,
Institute of Agriculture of the Carpathian Region of the National Academy of Agrarian Sciences of Ukraine
Obroshyne, Lviv region, Ukraine

Стадницька О. І.
доктор сільського господарства,
старший науковий співробітник
Інститут сільського господарства Карпатського регіону Національної академії аграрних наук України
с. Оброшине, Львівська область, Україна

Sobek Z.
Professor of Agriculture,
Professor at the Department of Genetics and Animal Breeding
Poznań University of Life Sciences
Poznań, Poland

Собек З.
професор сільського господарства,
професор кафедри генетики та розведення тварин
Познанський університет природничих наук
м. Познань, Польща

Introduction. Dystocia in dairy cattle is caused by numerous direct and indirect factors [1, p. 1] and results in many adverse consequences, which lead to substantial economic losses [2, p. 1]. It also negatively affects dam and calf welfare and influences acceptability of a production system by consumers.

Therefore, the aim of the present study was to detect dystocia in dairy cows using one of the statistical approaches, i.e. the k-nearest neighbor (KNN) method.

Materials and methods. A total of 12,712 calving records of Simmental cows from the herds located in Poland were analyzed. An initial dataset contained 32,817 records from the years 2005–2009 but was subsequently reduced after editing for missing, incorrect values and outliers. Two continuous predictor variables were used: AGE – cow's calving age (in months) and GEST – gestation length (in days). In addition, five categorical predictor variables were included in the model: SEASON – calving season (autumn – from October to December, winter – from January to March, spring – from April to June, summer – from July to September), LACT – lactation number (from the second to the seventeenth inclusive), SBREED – breed of the calf's sire (Simmental or other), SEX – calf sex (male or female), and PREV – the difficulty of the previous calving (two categories). The output (dependent) variable was calving difficulty class (easy vs. difficult).

Calving difficulty was originally recorded on the five-point (before 2006) or six-point (since 2006) scale according to the requirements set up by the Polish Federation of Cattle Breeders and Dairy Farmers. The five-point scale was as follows: 1 – an easy calving, 2 – a calving with help from man, 3 – a complicated calving, 4 – a very complicated calving, 5 – an abortion. The six-point scale was as follows: 1 – a spontaneous calving, 2 – an easy calving, 3 – a difficult calving, 4 – a very difficult calving, 5 – an abortion, 6 – cesarean section. Subsequently, the above ordinal scale was converted to a nominal one with only two categories: easy (scores 1 and 2 on both scales) and difficult (scores 3 and 4 on both scales and score 6 on the six-point scale). Abortions (score 5 on both scales) were excluded from the analysis. The whole dataset of calving records was randomly divided into a training set (L; 75% records) used for model development and a test set (T; 25% records) used for the objective verification of the model predictive performance. In the construction of the KNN model, the Euclidean distance between objects, the normalization of predictor variables and the 10-fold cross-validation (to find the optimum number of neighbors) were applied.

To evaluate model predictive performance, the following measures were used: sensitivity (*Se*), specificity (*Sp*), the *a posteriori* probability of true positive response (*PSTP*), the *a posteriori* probability of true negative response (*PSTN*) and accuracy (*Acc*) given by the following formulae [3, p. 3]:

$$Se = \frac{TP}{TP+FN},\ Sp = \frac{TN}{TN+FP},\ PSTP = \frac{TP}{TP+FP},\ PSTN = \frac{TN}{TN+FN},$$
$$Acc = \frac{TP+TN}{TP+TN+FP+FN},$$

where: *TP* is the number of correctly detected difficult calvings, *TN* is the number of correctly detected easy calvings, *FP* is the number of incorrectly detected easy calvings, *FN* is the number of incorrectly detected difficult calvings.

Results and discussion. The optimum number of the nearest neighbors in the applied model was five. Sensitivity (the proportion of correctly detected difficult calvings), specificity (the proportion of correctly classified easy calvings) and accuracy (the proportion of correctly classified calvings from both categories) on the independent T set were 0.00%, 100.00% and 97.77%, respectively. *PSTP* (showing the reliability of dystocia prediction by the model) was not possible to calculate due to the zero number of predicted difficult calvings. *PSTN* (showing the reliability of easy calving prediction by the model) was 97.77%. Based on the obtained results, it can be concluded that the data mining algorithm applied in the present study was ineffective in detecting cows with dystocia, which could have been caused by its low rate in the analyzed dataset (approximately 2.13%). This made it really hard for the classifier to correctly indicate such records. In addition, the use of a larger number of predictor variables would be required in the future research in order to improve the predictive performance of the KNN model.

Bibliography:

1. Ghavi Hossein-Zadeh N. Effect of dystocia on subsequent reproductive performance and functional longevity in Holstein cows. *Journal of Animal Physiology and Animal Nutrition*. 2016. Vol. 100, № 5. P. 860–867. DOI: https://doi.org/10.1111/jpn.12460.

2. Hiew M.W., Megahed A.A., Townsend J.R., Singleton W.L., Constable P.D. Clinical utility of calf front hoof circumference and maternal intrapelvic area in predicting dystocia in 103 late gestation Holstein-Friesian heifers and cows. *Theriogenology*. 2016. Vol. 85, № 3. P. 384–395. DOI: https://doi.org/10.1016/j.theriogenology.2015.08.017.

3. Zaborski D., Proskura W.S., Grzesiak W., Różańska-Zawieja J., Sobek Z. The comparison between random forest and boosted trees for dystocia detection in dairy cows. *Computers and Electronics in Agriculture*. 2019. Vol. 163. Art. № 104856. DOI: https://doi.org/10.1016/j.compag.2019.104856.