Session D Theatre 11

Time series data analysis to predict the status of mastitis in dairy cows by applying machine learning models to automated milking systems data

M. Dharejo[1], L. Minoque[2], T. Kabelitz[2], T. Amon[2], O. Kashongwe[2], M. Doherr[1]

[1] Institute for Veterinary epidemiology and biostatistics, school of veterinary medicine, Free university of Berlin, Konigsweg 67, 14163 Berlin, Germany, [2] Leibniz Institute for Agricultural Engineering and Bioeconomy, Max-Ayeth-Allee 100, 14469 Potsdam, Germany

Introduction: Mastitis in dairy cows is one of the most important issues that not only pose risk to animal health and welfare but also cause huge direct and indirect economic losses to the dairy sector. In recent times, automated milking systems (AMS) have gained sharp rise in popularity and adaptation. Mastitis detection under AMS operations becomes more difficult due to lack of direct human inspection of milk and udder during milking. The AMS technology consistently produces large amounts of milking records, which create the opportunity of developing algorithms to identify mastitis. The aim of this study was to predict mastitis in individual dairy cows through application of machine learning (ML) models on AMS generated high resolution data. Materials and Methods: The multivariable time series data with seven daily observed predictor variables and mastitis records of 1790 individual cows, was collected from two dairy farms situated in Saxony and Brandenburg states of Germany for a period of four years. We applied six ML models to correctly predict the status of mastitis (i) one day prior and (ii) on the day of clinical observation. Results: Each ML model varied in its efficiency for mastitis predictions. The overall accuracy, sensitivity and specificity scores of ML models ranged between (i) 0.80-0.90, 0.64-0.78 and 0.80-0.90 and, (ii) 0.84-0.93, 0.76-0.91 and 0.84-0.93 respectively. Conclusion: Our findings indicated moderate to high accuracy of ML models and demonstrated the robustness of time series AMS data by correctly predicting the future events of mastitis. However, as each model had its own strengths and weaknesses, therefore these findings have certain limitations. We propose inclusion of additional variables from AMS records and integration of other sensorial data for future studies.

Session D Theatre 12

Can AI accurately predict forage energy and protein values using chemical and textual data?

G. Tran[2], R. Genin[1], A. Lauront[1], M. Petitet[1], R. Rubrice[1], V. Heuzé[2], V. Guigue[1], A. Cornuejols[1]

[1] AgroParisTech, 22 place de l'Agronomie CS 20040, 91123 Palaiseau, France, [2] Association française de zootechnie, 22 place de l'Agronomie CS 20040, 91123 Palaiseau, France

Rationing is crucial in animal nutrition to meet protein and energy requirements, maximizing performance and minimizing feed costs. In France, the INRAE system calculates the energy and protein values of forages for ruminants, based on equations dependent on chemical composition, plant species, maturity, and other information. However, these calculations can be challenging when input data are lacking. We explored AI methods for calculating forage values according to the INRAE system, using six routine chemical parameters (dry matter, ash, protein, crude fibre, neutral detergent fibre, acid detergent fibre, lipids) and a text description of the forage. Output data are the energy values UFL (0.4 to 1.3 per kg) and UFV (0.3 to 1.4 per kg) and the protein values PDI (45 to 120 g/kg), PDIA (5 to 61 g/kg) and BPR (-60 to +150 g/kg). We used Python with machine learning and deep learning techniques, with a fine-tuned LLM on agricultural vocabulary. The training data (913 forages) came from the INRA 2018 tables for ruminants. Root mean square error (RMSE) is the evaluation metric. Using classical machine learning with decision tree boosting gave good results. The best were obtained using gradient boosting trees algorithms combining chemical data with the first two levels of text labels. RMSE of PDI, PDIA, and BPR were 2.65, 1.89, and 3.86 respectively. RMSE of UFL and UFV were 0.054 and 0.047 when chemical data were combined with a multiple correspondence analysis of labels. The deep learning approach used the BERT and camemBERTav2 LLMs to retrieve fodder description embeddings, combined with chemical data in a multi-layer perceptron. Fine-tuned BERT gave the lowest RMSE for PDI, PDIA, BPR, UFL and UFV: 4.21, 3.61, 5.04, 0.046, and 0.062 respectively. They were all higher than for the machine learning model except for UFL. Both classical machine learning and deep learning thus gave satisfactory results for the prediction of nutritional values of forages, with the former being more efficient than the latter. Further exploration is needed for practical deployment.