Prediction Quality of Cattle Behavior Traits Evaluated Through Different Cross-Validation Strategies Using Wearable Sensor Data and Machine Learning Algorithms

Leonardo A. C. Ribeiro[a], **Tiago Bresolin**[b], Guilherme J. M. Rosa[b], Daniel R. Casagrande[a], Marina A. C. Danes[a], João R. R. Dórea[b]

[a] Department of Animal Science, Federal University of Lavras, Minas Gerais, 37200–900, Brazil
[b] Department of Dairy Science, University of Wisconsin, Madison, 53706, United States

Introduction

Wearable sensors have been adopted as an alternative for real-time monitoring of cattle feeding behavior in grazing systems. However, even using machine learning (ML) techniques confounding effects such as cross-validation strategy may inflate the prediction quality. The objective was to evaluate the effect of different cross-validation strategies on the prediction of grazing activities in cattle using wearable sensor data and ML algorithms.

Material and Methods

- Six Nellore bulls (345 ± 21 kg) had their behavior visually classified as grazing or not-grazing for a period of 15 days.

- GLM, RF, and ANN were employed to predict behavior (grazing or not-grazing) using 3-axis accelerometer data.

Material and Methods

- Three cross-validation strategies were evaluated: holdout, leave-one-animal-out (LOAO), and leave-one-day-out (LODO).

- Algorithms were trained using similar dataset sizes (holdout: n = 57,862; LOAO: n = 56,786; LODO: n = 56,672).

Results

The GLM achieved the worst prediction accuracy (53%) compared to the ML techniques (65% for both RF and ANN).

The ANN performed slightly better than RF for LOAO (73%) and LODO (64%) cross-validation strategies.

The holdout yielded the highest accuracy values for all three ML approaches (GLM: 59%, RF: 76%, and ANN: 74%), followed by LODO (58%) and LOAO (55%).

Conclusions

The **GLM** approach was **not adequate to predict grazing behavior**, regardless of the cross-validation strategy.

The greater prediction accuracy observed for **holdout cross-validation** may simply indicate the **lack of data independence** and the presence of **carry over effects** from animals and grazing management.

Our results suggest that generalizing **predictive models to unknown** (not used for training) animals or grazing management may incur in **poor prediction quality**.

The results highlight the **need of using biological knowledge** to define the **validation strategy** that is closer to the real-life situation.