Using Test Day Data and Neural Network Regression to rank cow based on their future Lifetime Milk Yield Revenue

Main Author: D. Lefebvre*

Lactanet, Sainte-Anne-de-Bellevue, QC Canada

Presenting Author: D. Lefebvre*

Lactanet, Sainte-Anne-de-Bellevue, QC Canada

Other Authors: L. Fadul, R. Lacroix, D. Warner and M. Ayat

Lactanet, 555 Boulevard des Anciens-Combattants, Sainte-Anne-de-Bellevue, QC H9X 3R4 Canada

*Corresponding Author: dlefebvre@lactanet.ca

Increasing the longevity of dairy cows can lead to a more sustainable dairy industry and enhance farm profitability. Farmers often face important herd management decisions, such as determining which animals to keep. Integrating data with advanced analytics can help streamline this process. This study aimed to develop a model to predict future cow lifetime revenue based on yields using Dairy Herd Improvement (DHI) data. Data was sourced from Lactanet (Canadian Network for Dairy Excellence), covering 2,296 herds over eight years (2017-2024), with a total of 539,072 cows and 12.5 million test records. For model training, 80% of the data (1,836 herds) was used, while 20% (460 herds) was reserved for testing. A multilayer perceptron regressor from the scikit-learn package in Python was used to predict lifetime cumulative milk, fat, and protein yields at any given age in the future. These predicted yields were then used to estimate projected income. The model utilized 14 input features, including age on test day, parity, days in milk, fat and protein yields, somatic cell count linear score, current lactation and lifetime yields, and age at test. Up to eight consecutive test day records were randomly selected for each cow for prediction, with a minimum of one test. The inclusion of multiple breeds ensured the model's breed-agnostic nature. The model demonstrated high accuracy in predicting milk, fat, and protein yields, with R^2 values of 0.97 for each yield (MAPE: 5.68%, 5.92%, and 5.51%, respectively). Additionally, the projected income from the predicted yields showed strong correlation with observed income (R^2=0.97, MAPE=5.53%). Consistent results were found across breeds (e.g., Holstein R^2=0.97, Jersey R^2=0.94, Ayrshire R^2=0.97). The model effectively predicted cumulative yields at any future age, with an average prediction range of 5.4 years ahead. These findings suggest that the neural network can help farmers make informed decisions by predicting future cow lifetime yields and ranking animals within a herd for any age class.

Keywords: Test-day data, neural network regression, life time milk, dairy herd improvement.