A Model for the Prediction of Lifetime Profit Estimate of Dairy Cattle (Student Abstract)

Vahid Naghashi,[1,2] **Abdoulaye Banire Diallo,** [1,2]

[1] Departement d'Informatique, UQAM, Montreal, QC, Canada
[2] Laboratoire d'Algebre, de Combinatoire, et d'Informatique Mathematique, UQAM, Montreal, QC, Canada
PO Box 8888 Downtown station, Montréal, QC H3C 3P8
Montreal, Canada
naghashi.vahid@ens.uqam.ca

Abstract

In livestock management, the decision of animal replacement requires an estimation of the lifetime profit of the animal based on multiple factors and operational conditions. In Dairy farms, this can be associated with the profit corresponding to milk production, health condition and herd management costs, which in turn may be a function of other factors including genetics and weather conditions. Estimating the profit of a cow can be expressed as a spatio-temporal problem where knowing the first batch of production (early-profit) can allow to predict the future batch of productions (late-profit). This problem can be addressed either by a univariate or multivariate time series forecasting. Several approaches have been designed for time series forecasting including Auto-Regressive approaches, Recurrent Neural Network and a very deep stack of fully-connected layers. In this paper, we proposed a LSTM based approach coupled with attention and linear layers to better capture the dairy features. We compare the model, with three other architectures including NBEATs, ARIMA, MUMU-RNN using dairy production of 292181 dairy cow samples. The results highlight the performance of the proposed model and those of the compared architectures. We also highlight that such architecture could allow to predict late-profit with an error less than 3$ per month, opening the way of better resource management in the dairy industry.

Introduction

Profit estimation and prediction is often defined as forecasting the gained profits from livestock in the future taking into account several factors such as heath, productivity in the past months or years, environmental conditions and management policies. Recently, machine learning and deep learning models have been applied in agriculture and animal science fields (Sun et al. 2020). Deep learning techniques dealt with sequential and temporal agricultural data in different applications (Xu et al. 2020). Promising results were obtained in automatic cropping of cow's body region and cow's pattern identification for individual animals (Zin et al. 2018). Detection of the key parts of dairy cow's body (cow's head, back and legs) is performed using deep convolutional neural networks and Filter Layer YOLOv3 algorithm (Jiang et al. 2019). The problem of profit prediction can be addressed as a univariate or multivariate time-series forecasting problem. Recently, machine learning and deep learning models have been used in time-series prediction. One of the well-known classical time-series prediction models is the Auto-ARIMA (Contreras et al. 2003). This model corresponds to univariate time-series prediction methods. UniMu-RNN and MuMu-RNN (Frasco et al. 2020) are another univariate and multivariate time-series forecasting models (Hochreiter and Schmidhuber 1997) which consists of two stacks of LSTM layers followed by a linear layer for prediction of the milk production profits in upcoming months. Meanwhile, for the univariate time series, N-BEATS has been shown to provide the most prominent outcomes. N-BEATS uses a stack of fully connected neural network layers to forecast the future values of a target series in a single pass (Oreshkin et al. 2019).

Problem Definition

The lifetime profit estimate problem can be defined as follows. For each input example (ith sample) of size T, i.e. $x_i = (x_i^1, ..., x_i^T) \, \epsilon \, \Re^{p \times T}$ with p as the number of input features and T the length of the series, a prediction model predicts the future profits of M steps ahead, $\hat{p}_i \, \epsilon \, \Re^{1 \times M}$. Therefore, our aim is to learn a set of nonlinear functions f to map the input data $X \, \epsilon \, \Re^{b \times T \times p}$ to the estimated profit values in the future months $\hat{Y} \, \epsilon \, \Re^{b \times M}$ (b is the number of the cow samples): $\hat{Y} = f(X)$.

Materials and Methods

To assess the lifetime profit estimate, we collected a dataset of 292181 milking cow samples of Holstein breed collected during 39 months. For each month and for each cow the profit associated with the milk yield during the test dates is collected as well as 14 other features (see Table 1 in supplementary material part). Only the data for cows from age 18 in month to 46 of the cows were used in this study. In the dataset, we defined early test-date as the one for cows with age from month 18 to 46 and future test-date for cows with age from month 47 to 56. Thus the input data is represented as a tensor of shape $\Re^{b \times T \times p}$, where $b = 292181$, $T = 29$ and $p = 15$. In this study we proposed a LSTM model to capture the sequence patterns of the features. We coupled it with an attention mechanism allowing to better handle the

RMSE results				
Univariate		Multivariate		
ARIMA	N-BEATS	ConvRNN	MuMu	PM
12.69	7.36	8.15	8.19	**2.36**

Table 1: Prediction results in terms of RMSE (whole dataset). PM refers to our proposed model and MuMu is an abbreviation for MuMu-RNN model (Frasco et al. 2020)

dairy features. The two layers of LSTMs capture the temporal relations among different time steps and encode this information through hidden vectors. The sequence of hidden states are fed to the attention layer to assign weights (scores) to each hidden state. Then, a weighted hidden state vector is computed using the attention scores and all of the hidden states. The final prediction is obtained by passing the weighted hidden vector through a linear layer (for more details see supplemental Figure 1). In the model we provided two learning units allowing to capture dry cows and milking cows patterns, separately. This option allows an optimisation of the model. We compared the proposed model (PMS) with two univariate time-series methods: 1) Auto-ARIMA (Contreras et al. 2003); 2) N-BEATS (Oreshkin et al. 2019) and with two multivariate models used in the literature: 1) ConvRNN(Maggiolo and Spanakis 2019); 2) MuMu-RNN (Frasco et al. 2020).

Experimental Settings

The train-test split was applied to the main dataset with 70:30 ratio. In other words, the models were trained with 200k samples and tested using 92181 remaining cow samples. All the models were trained for 50 epochs using GPUs (4 x NVIDIA T4 Turing and 16GB GDDR6 memory). MSE (Mean Square Error) loss function has been used in the training process and being optimized with Adam optimizer. The evaluation criteria used to assess the models' performance is the Rooted Mean Square Error (RMSE) which is calculated based on the actual and the predicted profits.

Results

Table 1 shows the RMSE of all the trained models using the dairy dataset. The NBEATS performs better than ARIMA and it is even slightly better than the two multi-variate models (ConvRNN, MuMu-RNN). However, the proposed model (PM) adapted for dairy provides better RMSE at 2.36$ per month. An example of profit prediction for a cow's sample using the proposed method is illustrated in Figure 1 together with the actual profits over the forecasting months. In this figure, the reason of a sudden spike from month 47 to 48 is that the true profit value for this cow sample in month 47 were missing and the missing profit was imputed with the value of -5 (the cow treated as dry). But the same cow was profitable and without missing profit value in month 48 (with a profit close to 15). Therefore, the trained model was trying to predict the profits accurately, leading to that sudden changes both in estimated profits and the ground truth ones.
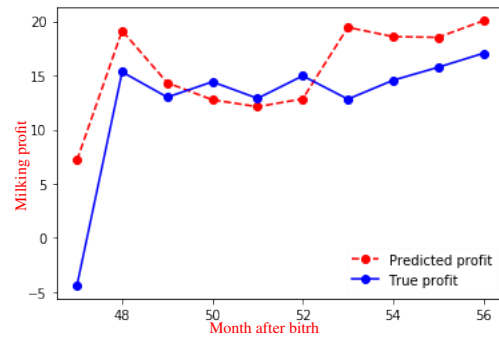

Figure 1: An example of a well-predicted cow sample using our proposed model.

References

Contreras, J.; Espinola, R.; Nogales, F. J.; and Conejo, A. J. 2003. ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems*, 18(3): 1014–1020.

Frasco, C. G.; Radmacher, M.; Lacroix, R.; Cue, R.; Valtchev, P.; Robert, C.; Boukadoum, M.; Sirard, M.-A.; and Diallo, A. B. 2020. Towards an Effective Decision-making System based on Cow Profitability using Deep Learning. In *ICAART (2)*, 949–958.

Hochreiter, S.; and Schmidhuber, J. 1997. LSTM can solve hard long time lag problems. *Advances in neural information processing systems*, 473–479.

Jiang, B.; Wu, Q.; Yin, X.; Wu, D.; Song, H.; and He, D. 2019. FLYOLOv3 deep learning for key parts of dairy cow body detection. *Computers and Electronics in Agriculture*, 166: 104982.

Maggiolo, M.; and Spanakis, G. 2019. Autoregressive convolutional recurrent neural network for univariate and multivariate time series prediction. *arXiv preprint arXiv:1903.02540*.

Oreshkin, B. N.; Carpov, D.; Chapados, N.; and Bengio, Y. 2019. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv:1905.10437*.

Sun, Z.; Di, L.; Fang, H.; and Burgess, A. 2020. Deep Learning Classification for Crop Types in North Dakota. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 2200–2213.

Xu, J.; Zhu, Y.; Zhong, R.; Lin, Z.; Xu, J.; Jiang, H.; Huang, J.; Li, H.; and Lin, T. 2020. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. *Remote Sensing of Environment*, 247: 111946.

Zin, T. T.; Phyo, C. N.; Tin, P.; Hama, H.; and Kobayashi, I. 2018. Image technology based cow identification system using deep learning. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, 236–247.