

Milk Prediction Using Conformation Traits as Privileged Information

In the domain of dairy farming, prediction of milk production represents a crucial aspect of dairy cattle breeding, facilitating advancements in animal monitoring, herd management, and decision-making for farmers [1]. Milk yield is influenced by a wide range of factors, including health, environment, management practices, and genetics [1]. Traditional approaches include statistical models, such as autoregressive moving average (ARMA) and seasonal autoregressive integrated moving average (SARIMA), as well as deep learning (DL) methods (recurrent neural networks & multilayer perceptrons) [1]. However, these models often address regression or forecasting separately, limiting their adaptability. DL models lack flexibility due to fixed input sizes, while statistical models face scalability issues as the population size increases.

Our study aims to develop a model able to predict future lactation curves using historical lactation data and phenotypic information, such as estimated breeding values and conformation traits, while minimizing computational expenses. More specifically, we predict two sets of target variables depending on the type of task: either a single-output or a multi-output. We also address two problems: regression and forecasting.

To overcome these limitations, we propose **LSTMDropout**, a novel model that leverages training-exclusive features through the learning using privileged information paradigm [2]. Combining long-short-term memory (LSTM) networks with heteroscedastic dropout improves variance estimation and uncertainty modeling, enhancing prediction robustness. Our architecture consists of three (3) main steps. The **first one** addresses the dimensional mismatch between the baseline data $X_{base} \in \mathbb{R}^{n_{samples}, n_{seqlength}, n_{features_{base}}}$ and privileged data $X_{priv} \in \mathbb{R}^{n_{samples}, n_{seqlength}, n_{features_{priv}}}$. Separate LSTM blocks are employed to learn representations from these two inputs. Once transformed, these representations are passed through a linear projection layer, a normalization layer, and a ReLU activation layer, ensuring that both inputs are aligned for subsequent processing. Our approach handles heterogeneous inputs, time series data (X_{base}), and phenotypic traits (conformation traits & EBVs, X_{priv}). This adaptation adds complexity but ensures that each type of data is processed optimally. This **second step** involves estimating the variance of the heteroscedastic dropout from the two previously learned representations. To calculate the variance of the heteroscedastic dropout during training, we apply two distinct linear layers to the two representations learned in the first step. These learned representations will be the inputs for the third step. Note that during testing, we will only have access to the X_{base} representations as inputs in the second step. We will duplicate these representations before applying the linear layers and the block dropout. The **third and final step** is common to both training and testing. However, the architecture varies depending on the nature of the task:

- *Regression.* A feed-forward layer is used to output a single value.
- *Forecasting.* We first apply either a linear layer (feed-forward simple) or an LSTM layer (feed-forward RNN), followed by a linear layer to obtain the final output dimension.

In the benchmark dataset, which contains dairy production records of 11,000 individuals, our LSTMDropout model surpasses ARMA, SARIMA, and N-BEATS in regression (errors RMSE: 10.55 and MAE: 8.42) and outperforms the temporal fusion transformer in forecasting, achieving lower errors in both single-output (10.55, 8.42) and multi-output tasks (5.49, 4.48). These results highlight the effectiveness of incorporating privileged information with heteroscedastic dropout, offering accuracy, scalability, and computational efficiency for the dairy industry.

References

- [1] C. G. Frasco., M. Radmacher., R. Lacroix., R. Cue., P. Valtchev., C. Robert., M. Boukadoum., M. Sirard., and A. B. Diallo. Towards an effective decision-making system based on cow profitability using deep learning. In *Proceedings of the 12th ICAART - Volume 2*, pages 949–958. SciTePress, 2020.
- [2] V. Vapnik and A. Vashist. A new learning paradigm: Learning using privileged information. *Neural Networks*, 22(5):544–557, 2009. ISSN 0893-6080. Advances in Neural Networks Research: IJCNN2009.