Calving Prediction Algorithm Development Using Activity and Rumination Time Data in Dairy Cattle

Gyeonglim Ryu*, Hyeonguk Baek*, Jihwan Lee*, Seulgi Lee*, Seongwoo Kim*, Enkhmanlai Ganbataar* and Tae-jeong Choi*
*Dairy Science Division, National Institute of Animal Science, Rural Development Administration, Cheonan 31000, Korea
e-mail: ryuglim@korea.kr

젖소의 개체 및 활동량, 반추시간 데이터를 활용한 분만 예측 모델 개발

유경림*, 백형욱*, 이지환*, 이슬기*, 김성우*, Enkhmanlai Ganbataar*, 최태정*
*농촌진흥청 국립축산과학원 낙농과

Abstract

The objective of this study was to develop a machine learning-based algorithm to predict calving in dairy cattle using activity and rumination data combined with individual information. A total of 349 calving records were collected from 231 cows between 2020 and 2024. Five machine learning models were compared, and their performance was evaluated using sensitivity, specificity, precision, and F2 score. Among the models, RandomForest achieved the highest sensitivity and F2 score, while RandomForest and XGBoost demonstrated the best balance between sensitivity and specificity. These results suggest that integrating biometric data(activity and rumination time) with machine learning can improve the accuracy of calving management and support farm labor efficiency.

1. Introduction

Calving management in dairy cows is a critical factor that directly affects farm productivity. Accurately detecting the expected calving date and providing precise, efficient care before and after calving have a significant impact on overall milk yield, reproductive performance, and calf survival rates. Most farms still rely heavily on the experience of farmers and visual observation by workers to identify signs of calving. However, these traditional methods have limitations, including low prediction accuracy and high labor demands.

With the advancement of smart dairy technologies, it has become possible to continuously collect and analyze behavioral data from cows in real time. This development has led to growing interest in utilizing such data for calving prediction. In particular, behavioral indicators such as activity levels and rumination time are closely linked to the physiological changes that occur before calving, making them valuable metrics for predicting calving events. Nevertheless, effectively using these data requires algorithms capable of handling large datasets and identifying complex relationships among multiple variables.

In this study, we aimed to develop a machine learning-based algorithm to predict calving by integrating individual cow information (e.g., age, parity, last insemination date) with behavioral data (activity and rumination time). Five machine learning models—Decision Tree, Random Forest, XGBoost, Gradient Boosting, and LightGBM—were applied and evaluated. The model with the best performance was identified and its potential application in practical dairy farm management is discussed.

2. Materials and Methods

2.1 Experimental Animals

A total of 231 cows (199 Holstein and 32 Jersey) and 349 calving records collected from 2020 to 2024. All experimental animals and designs were approved by the Nationalk Institute of Animal Science Animal Care and Ethics Committee (NIAS-2023006).

2.2. Data Collecting

2.2.1. Activity and Rumination Time

Behavioral data were continuously collected every 2 hours using a cow health and reproduction monitoring system (SCR Heatime® HR System, SCR by Allflex). Daily activity and rumination records were collected from -14 to 0 days relative to the calving date (Day 0 = calving date).

2.2.2. Individual Data

Parity, age, and last insemination data were collected and recorded for each cow.

2.3. Statistical Analysis

Statistical analyses and model development were conducted using Python 3.13.2. Five machine learning models were evaluated(Decision Tree, RandomForest, XGBoost, GradientBoosting, and LightGBM). Performance was assessed using Sensitivity, Specificity, Precision, and F2 score.

3. Results&Discussion

Five machine learning models were evaluated for their performance in predicting calving events, and the results are summarized in Table 1.

First, the RandomForest model showed the highest overall performance, with a sensitivity of 0.671 ± 0.090 and specificity of 0.723 ± 0.036, demonstrating a strong ability to correctly identify cows close to calving while maintaining a relatively low false positive rate. F1 score was 0.134 ± 0.013, and the AUC reached 0.783 ± 0.036 and it's the highest among the tested models.

The Decision Tree model showed the highest sensitivity at 0.734 ± 0.096. However, its specificity was relatively lower at 0.657 ± 0.085. The F1 score was 0.124 ± 0.016, and the AUC was 0.750 ± 0.034, indicating moderate overall performance. While its prediction power was slightly lower than RandomForest model, the Decision Tree provided better, which is useful for understanding key contributing factors to calving prediction.

For XGBoost, the sensitivity and specificity were 0.570 ± 0.083 and 0.780 ± 0.029, respectively. This model showed a good balance between sensitivity and specificity. The F1 score was 0.140 ± 0.026, and the AUC was 0.759 ± 0.037, which was slightly lower than RandomForest but still competitive, making it suitable for practical use where a balance between missed detections and false alarms is critical.

The Gradient Boosting model demonstrated the highest specificity at 0.800 ± 0.030, meaning it was most effective in minimizing false positives. However, its sensitivity was lower at 0.531 ± 0.096, indicating a tendency to miss actual calving events. Its F1 score and AUC were 0.141 ± 0.021 and 0.761 ± 0.042, respectively, showing comparable performance to XGBoost in terms of classification power.

Finally, the LightGBM model achieved a sensitivity of 0.584 ± 0.076 and specificity of 0.765 ± 0.032, with an F1 score of 0.136 ± 0.021 and an AUC of 0.764 ± 0.037. While its sensitivity was slightly lower than RandomForest and Decision Tree, its balanced performance made it a reliable alternative model.

Overall, RandomForest provided the most consistent performance across metrics, while XGBoost and Gradient Boosting also showed potential depending on whether higher sensitivity or specificity is prioritized. Decision Tree, despite lower AUC, remains valuable for interpretability and as a decision support tool for understanding behavioral and individual factors influencing calving prediction.

[Table 1] Performance comparison of five machine learning models for calving prediction. (Values represent mean ± standard deviation.)

Model	Sensitivity	Specificity	F1_Score	AUC
LightGBM	0.584 (±0.076)	0.765 (±0.032)	0.136 (±0.021)	0.764 (±0.037)
Gradient Boosting	0.531 (±0.096)	0.800 (±0.030)	0.141 (±0.021)	0.761 (±0.042)
XGBoost	0.570 (±0.083)	0.780 (±0.029)	0.140 (±0.026)	0.759 (±0.037)
Decision Tree	0.734 (±0.096)	0.657 (±0.085)	0.124 (±0.016)	0.750 (±0.034)
Random Forest	0.671 (±0.090)	0.723 (±0.036)	0.134 (±0.013)	0.783 (±0.036)

4. Conclusion

This study demonstrated the potential of machine learning algorithms for predicting calving in dairy cattle using continuous behavioral data and individual cow information. The RandomForest model showed the most promising performance, while XGBoost also provided reliable results. Implementing such algorithms on dairy farms could enhance labor efficiency by allowing farmers to prepare for calving events in advance, reducing unnecessary waiting time and improving animal welfare. Future research will focus on integrating real-time sensor data and validating the algorithm under diverse farm conditions for broader application.

References

[1] S. R. Ahuja, "Multimedia Conferencing System", IEEE Conference on Computer Workstation, pp. 52-58, March, 1998.

[2] C. E. F. Clark, N. A. Lyons, L. Millapan, S. Talukder, G. M. Cronin, K. L. Kerrisk, and S. C. Garcia, "Rumination and activity levels as predictors of calving for dairy cows," Animal, vol. 9, no. 4, pp. 691-695, 2015.

[3] M. R. Borchers, Y. M. Chang, K. L. Proudfoot, B. A. Wadsworth, A. E. Stone, and J. M. Bewley, "Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle," J. Dairy Sci., vol. 100, pp. 5664-5674, 2017.

[4] A. Raza, K. Abbas, T. Swangchan-Uthai, H. Hogeveen, and C. Inchaisri, "Behavioral adaptations in tropical dairy cows: Insights into calving day predictions," Animals, vol. 13, no. 18, pp. 1-13, 2024.

[5] Y. Wang, A. Perea, H. Cao, M. Bakir, and S. Utsumi, "A two-stage machine learning approach for calving detection in rangeland cattle," Agriculture, vol. 15, no. 13, pp. 1434, 2025.

[6] R. Antanaitis, L. Anskienė, G. Palubinskas, K. Džermeikaitė, D. Bačėninaite, L. Viora, and A. Rutkauskas, "Ruminating, eating, and locomotion behavior registered by innovative technologies around calving in dairy cows," Animals, vol. 13, no. 7, pp. 1257, 2023.

[7] S. Kishi, T. Kojima, C. Y. Huang, K. Yayou, and K. Fujioka, "A feasibility study on predicting cow calving time over 40 h in advance using heart rate and financial technical indicators," Scientific Reports, vol. 14, no. 1, pp. 1-10, 2024.