Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative Analysis of Machine Learning Models to Predict Backfat Thickness in Hanwoo Cattle

Taeyong Yun
Yonsei University
Dawoon Jeong
Jeonbuk National University
Jinhyeon Yun
Chonnam National University
Woongsup Lee
woongsup.lee@yonsei.ac.kr

Yonsei University

Article

Keywords:

Posted Date: March 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9026112/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Comparative Analysis of Machine Learning Models to Predict Backfat Thickness in Hanwoo Cattle

Taeyong Yun[1], Dawoon Jeong[2], Jinhyeon Yun[3,4], and Woongsup Lee[1,*]

[1]Graduate School of Information, Yonsei University, Seoul, 03722, South Korea
[2]Department of Animal Resources Biotechnology, Jeonbuk National University, Jeonju, 54896, South Korea
[3]Department of Animal Science, Chonnam National University, Gwangju, 61186, South Korea
[4]Graduate School of Data Science, Chonnam National University, Gwangju, 61186, South Korea
[*]woongsup.lee@yonsei.ac.kr

ABSTRACT

Accurate prediction of carcass traits is essential for genetic improvement and value optimization in the beef industry. Backfat thickness is particularly important because it directly affects both market value and consumer preference. This study evaluated the predictive performance of machine learning (ML) models for estimating backfat thickness in Hanwoo cattle. A total of 386 Hanwoo carcass records were used, and 10 carcass features served as input variables, with backfat thickness being the prediction target. Model performance was assessed primarily by mean absolute error (MAE) and secondarily by the coefficient of determination (R^2). Support vector regression (SVR) showed the best predictive performance, achieving the lowest MAE (2.796) and the highest R^2 (0.375) after hyperparameter tuning. In contrast, other models showed higher MAE values, generally ranging from 3.0 to 4.0. Pearson correlation analysis identified intramuscular fat score ($r = 0.37$), quality grade ($r = -0.33$), and sex ($r = 0.34$) as the most influential predictors of backfat thickness. These findings support SVR as a robust approach for predicting carcass backfat thickness in Hanwoo cattle and provide a practical framework for ML-based precision phenotyping and breeding strategies.

Introduction

Backfat thickness in Hanwoo cattle is a key determinant of beef quality and market value because it directly influences carcass grade and consumer preference[1,24]. Although moderate backfat can improve palatability through its association with marbling, excessive fat deposition reduces yield grade and economic return. Accurate prediction of this trait is therefore important to maintain carcass uniformity and optimize profitability. To date, most predictive studies have relied on genomic approaches. For example, Lopes et al.[25] showed that multi-trait analyses improved prediction accuracy by leveraging genetic correlations among traits. However, genomic approaches mainly focus on heritability and molecular markers and may underrepresent the combined effects of environmental, phenotypic, and management-related factors that also influence fat deposition.

In this context, machine learning (ML) offers a promising alternative to build predictive models of complex traits due to its ability to capture nonlinear relationships and high-order interactions among predictive variables, thereby improving prediction accuracy[6]. Over the past decade, ML has gained increasing attention in animal science for predicting diverse phenotypic traits. Ruchay et al.[36], for example, reported that ML models outperformed traditional linear models for cattle body-weight prediction based on morphometric measurements, with random forest achieving the highest accuracy. Comparative analyses have also shown that support vector machines can provide strong predictive performance for quantitative traits in cattle and pigs[28,42].

More recent livestock-focused studies further indicate that model superiority is trait- and population-dependent: ML methods improved genomic prediction of growth traits in composite beef cattle[15], while comparative ML analyses in Holstein cattle showed clear model-specific differences in predictive performance for behavior-related traits[32]. Given the wide range of available ML algorithms and their distinct learning mechanisms, it is essential to perform a comparative evaluation, as the optimal model often depends on the specific trait and dataset characteristics. Identifying the best-performing model is therefore critical to enhancing the accuracy, robustness, and practical utility of prediction models in livestock breeding and management.

Despite growing interest in applying ML to livestock phenotyping, few studies have systematically compared a broad range of algorithms to identify effective models for predicting backfat thickness in Hanwoo cattle. Here, we benchmarked 14 classical and ensemble-based ML algorithms using 10 carcass-related variables as predictors. Rather than using deep learning methods, which typically require larger datasets, we focused on approaches better suited to the moderate sample size available for Hanwoo carcass traits. Because comparative results in moderate-sized datasets can be sensitive to data splitting and hyperparameter tuning[3,17,22,23], we applied a consistent evaluation procedure and a transparent tuning strategy to enable fair comparisons across models. By identifying top-performing algorithms under this setting, we aim to establish an optimized

predictive framework tailored to Hanwoo carcass characteristics and to provide practical guidance for improving prediction accuracy.

Results

Correlation Analysis

Pearson's correlation analysis revealed that backfat thickness was most strongly correlated with intramuscular fat score (fat number, $r = 0.37$, $p < 0.001$), sex ($r = 0.34$, $p < 0.001$), and meat quality grade (grade score, $r = -0.33$, $p < 0.001$). Pairwise correlation coefficients among all carcass variables are visualized in Figure 1.

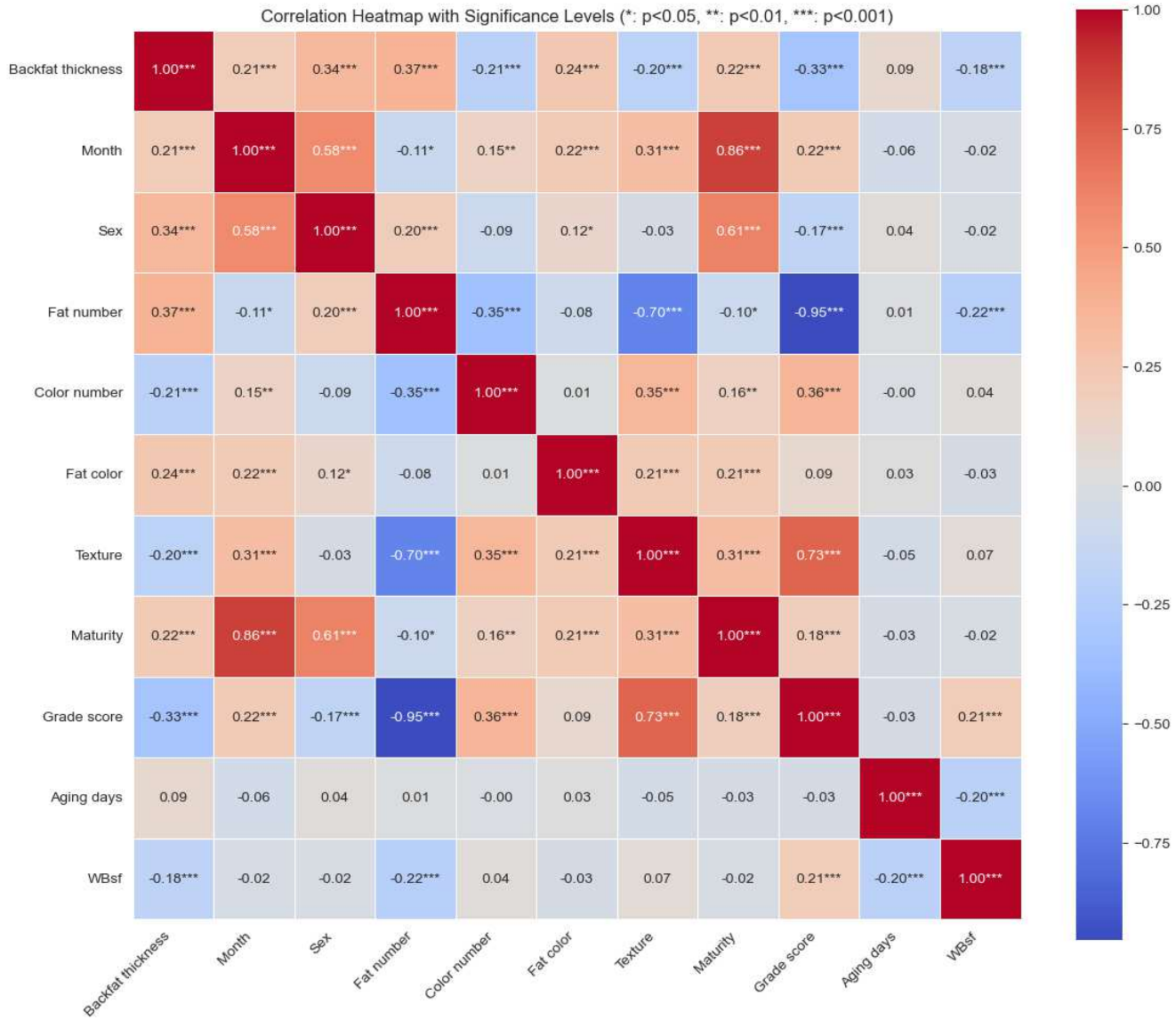


Figure 1. Heatmap illustrating Pearson's correlation coefficients among the carcass variables. Statistical significance is indicated as ns ($p \geq 0.05$), * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

Model Performance Comparison

Among the 14 ML models evaluated, SVR achieved the lowest MAE before and after hyperparameter optimization (2.94 mm and 2.80 mm, respectively) and the highest R^2 (0.357 and 0.375, respectively). The MAE and R^2 values for all models are presented in Figures 2 and 3, respectively.

Ensemble methods ranked second overall: AdaBoost yielded an MAE of 3.10 mm ($R^2 = 0.30$) and random forest 3.14 mm ($R^2 = 0.29$). Other ensemble models, including XGBoost (3.22 mm, $R^2 = 0.27$), CatBoost (3.25 mm), gradient boosting (3.28 mm, $R^2 = 0.26$), and LightGBM, showed intermediate performance. Linear regression models, including Bayesian ridge, ridge, lasso, and ordinary linear regression, produced MAE values of approximately 3.25–3.28 mm and R^2 values of approximately 0.29–0.30. Instance-based and neural models showed the highest prediction errors: KNN yielded an MAE of 3.47 mm ($R^2 = 0.14$), decision tree 3.51 mm, and multi-layer perceptron (MLP) 3.57 mm ($R^2 = 0.13$).

To improve interpretability of model ranking, we additionally summarized the relative gain of the optimized SVR model against representative alternatives (Table 1). Compared with the strongest competing ensemble models, SVR reduced MAE by approximately 9.7–10.8%. Relative to the linear-model group, MAE reduction was approximately 13.8–14.6%, and larger reductions were observed versus KNN (19.3%) and MLP (21.6%). In parallel, SVR provided absolute R^2 gains of 0.075–0.245 over these comparison models. These effect-size summaries indicate that SVR delivered not only the best rank but also practically meaningful error reduction under the current experimental setting.

Table 1. Relative performance gain of the optimized SVR model over representative alternatives.

Comparison model	MAE (mm)	MAE reduction vs SVR (%)	ΔR^2 vs SVR
AdaBoost	3.10	9.7	0.075
Random forest	3.14	10.8	0.085
Linear-model group[†]	3.25–3.28	13.8–14.6	0.075–0.085
KNN	3.47	19.3	0.235
MLP	3.57	21.6	0.245

Note: Optimized SVR performance: MAE = 2.80 mm, $R^2 = 0.375$.

[†] Bayesian ridge, ridge, lasso, and ordinary linear regression models.

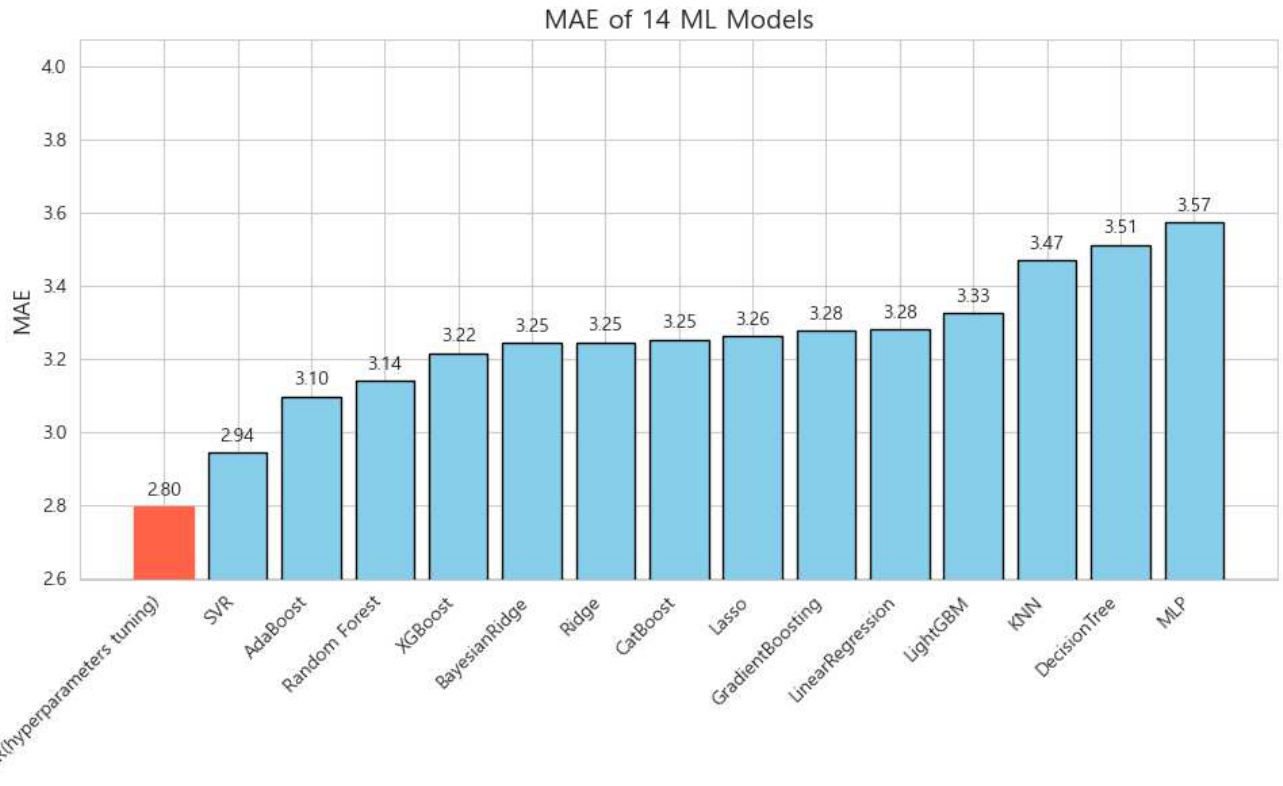


Figure 2. MAE comparison across 14 machine learning models to predict backfat thickness in Hanwoo cattle.

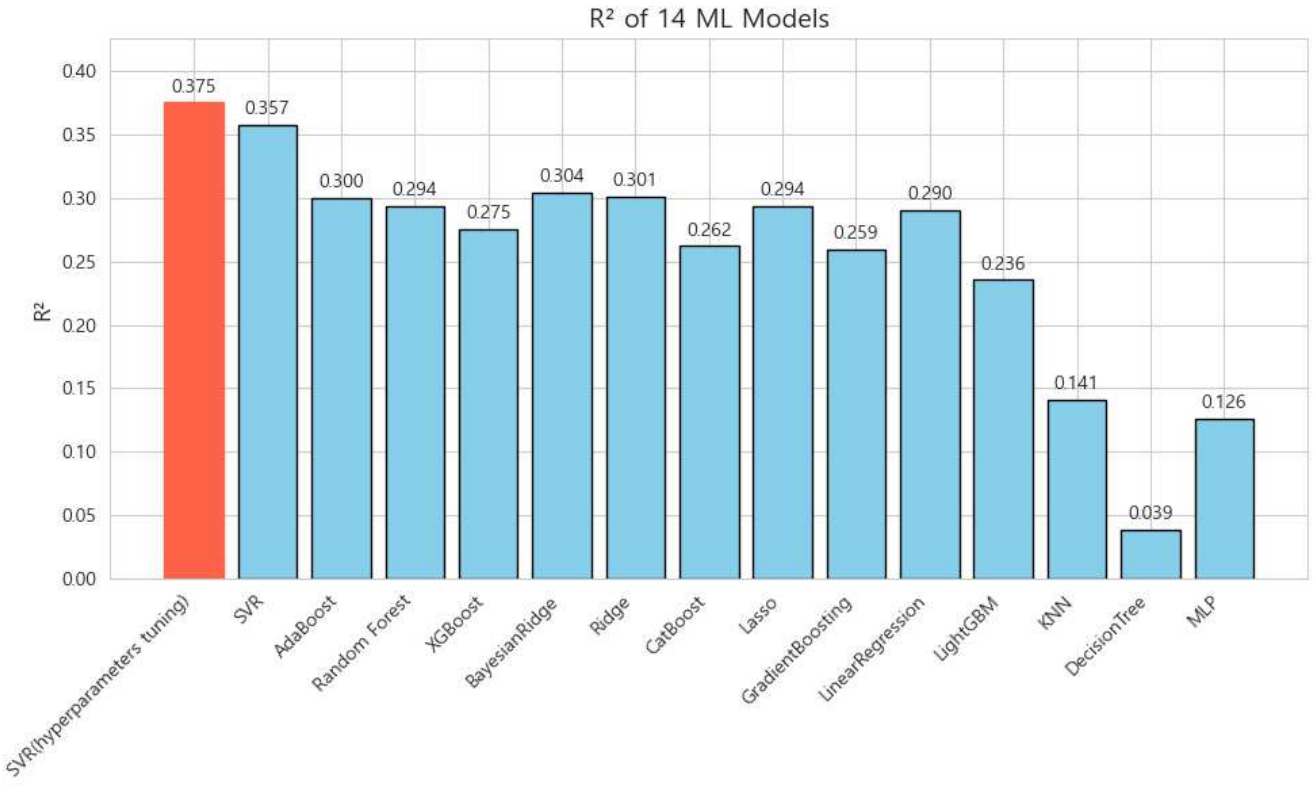


Figure 3. R^2 score comparison across 14 machine learning models to predict backfat thickness in Hanwoo cattle.

SVR Partial Dependence by Sex

Figure 4 shows the relationship between intramuscular fat score (fat number) and backfat thickness, stratified by sex, with SVR-predicted values obtained by averaging all other input features at each fat-number level. A positive trend between fat number and backfat thickness was observed in steers and cows/heifers (including hysterectomized heifers). In contrast, bulls did not show a consistent directional trend across fat-number levels.

Discussion

To our knowledge, this is the first study to systematically compare 14 ML regression algorithms for predicting backfat thickness in Hanwoo cattle using 10 carcass traits. The optimized SVR model achieved the best predictive performance (MAE $\approx$ 2.80 mm; $R^2 \approx 0.38$), indicating that it effectively captured complex nonlinear relationships among carcass traits. This finding is broadly consistent with recent livestock genomic prediction studies showing that nonparametric or kernel-based models can be highly competitive for trait prediction in cattle populations[15,30,42]. In contrast, our results differ from some reports on body weight prediction in cattle, where ensemble methods, particularly random forest, typically achieved the highest accuracy[36]. This discrepancy highlights the importance of evaluating multiple algorithms within each specific application, as optimal modeling strategies are highly context-dependent, influenced by target traits, data characteristics, and dataset size.

The comparable R^2 values observed between ensemble methods (AdaBoost, $R^2 = 0.30$; random forest, $R^2 = 0.29$) and regularized linear models (ridge, lasso, Bayesian ridge, $R^2 \approx 0.29$–0.30) suggest that most predictive relationships among carcass traits and backfat thickness were largely linear, leaving limited potential for improvement through nonlinear ensemble approaches under the current dataset size and feature structure. The higher prediction error of KNN (MAE = 3.47 mm) is consistent with the known sensitivity of instance-based methods to high-dimensional feature spaces[33]. The low performance of MLP (MAE = 3.57 mm, $R^2 = 0.13$) and decision tree (MAE = 3.51 mm) likely reflects overfitting and instability under a moderate sample size ($N = 386$), consistent with previous livestock prediction studies[40].

The correlation analysis identified intramuscular fat score, sex, and quality grade as the most influential predictors of backfat thickness. The positive association between marbling and subcutaneous fat accumulation is consistent with previous studies in Hanwoo cattle, reflecting shared genetic and nutritional factors influencing lipid deposition[1,24]. The sex-stratified partial dependence analysis further showed that the positive fat number–backfat relationship was evident in steers and cows/heifers,

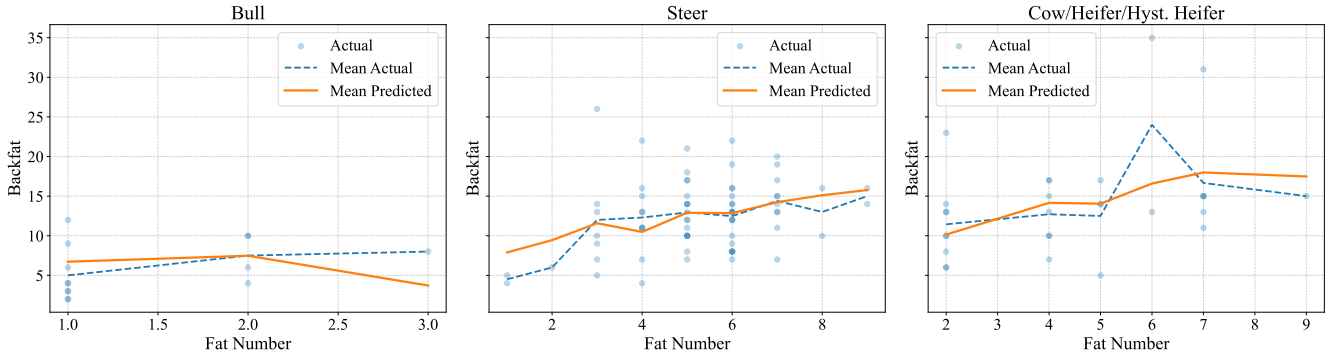


Figure 4. Relationship between intramuscular fat score (fat number) and backfat thickness by sex, with SVR predictions. Each blue point represents an individual observation of backfat thickness at a given fat number. The dashed line indicates the empirical mean backfat thickness for each fat number level, providing a baseline reference.

while bulls showed inconsistent patterns, likely attributable to the small number of bull samples and the common practice of early castration in Hanwoo production[19]. The negative correlation between quality grade and backfat thickness ($r = -0.33$) indicates that cattle with superior quality grades tend to accumulate less subcutaneous fat, a pattern consistent with the grading criteria applied in the Korean beef industry[24].

Accurate prediction of backfat thickness offers significant benefits to livestock producers and breeders. Reliable predictive frameworks, such as SVR, can support data-driven decisions on feeding management, optimal slaughter timing, and genetic selection, thereby improving carcass yield, beef quality, and overall economic outcomes. Incorporating informative predictors such as fat number into ML frameworks may further improve prediction of economically important carcass and beef-quality traits[38,40].

The major strength of this study is the systematic comparison of diverse ML paradigms under consistent experimental conditions. However, several limitations should be acknowledged. First, the moderate dataset size ($N = 386$) limited the applicability of more complex models, such as deep neural networks (DNNs), and may reduce generalizability. Second, although SVR was extensively tuned, other models underwent less intensive optimization; broader hyperparameter searches across all models could alter relative performance rankings. Third, the exclusion of genomic and environmental covariates may constrain external validity. Recent Hanwoo and Nellore studies suggest that expanding reference populations, targeted variable selection, and multi-omics integration can improve prediction for growth and carcass-related traits[13,21,31]. Future studies should therefore use larger multi-breed and multi-environment datasets and extend this framework to multi-trait prediction for precision livestock management.

From a translational perspective, interpretability is important for practical adoption. Although the SVR model achieved the best predictive performance in our study, practitioners often require explanations for why predicted backfat thickness increases or decreases for specific animals before using model outputs to guide feeding strategies or selection decisions. Post hoc explanation tools such as SHAP and LIME can provide complementary feature-level insights, particularly for assessing whether influential predictors (for example, marbling-related scores, sex, and grading variables) act consistently across subgroups. Integrating interpretability analyses with standard performance reporting may improve transparency and support deployment in breeding and carcass-management workflows[26,35]. Given the moderate sample size and the single-cohort nature of our dataset, external validation on independent populations and periodic model recalibration will be important before routine deployment[17,22,23].

Our comparative results further suggest that, when sample size is constrained (e.g., $N = 386$), modestly complex but well-regularized models may be preferable to highly flexible architectures. The relatively close performance of ensemble methods and regularized linear models, together with the instability of MLP, is consistent with a bias-variance tradeoff in which increasing model capacity does not necessarily improve generalization under limited data[17]. This consideration is common in livestock datasets where phenotypes are costly to collect and covariates may be heterogeneous across institutions and production environments.

From an implementation standpoint, SVR-based backfat predictions could be used as a screening tool to flag animals likely to exceed economically desirable fat ranges prior to slaughter planning. These predictions may be further combined with farm-level management factors, such as feeding regime and age-at-slaughter targets, to support decision-making in practice[9]. However, model performance may shift when applied to different regions, management systems, or genetic subpopulations. Future work should prioritize validation on independent cohorts and the development of larger multi-institutional datasets that

integrate phenotypic, genomic, and management variables to strengthen generalizability and accelerate translation[18].

Methods

Dataset and Pre-processing

We analyzed a total of 386 Hanwoo carcass records collected between January 1, 2007, and December 31, 2020, from Jeonbuk National University and the National Institute of Animal Science, South Korea. The dataset comprised 37 bulls, 104 cows, and 245 steers, representing the general composition of the Hanwoo beef industry. Only records with complete carcass trait measurements were retained. Specifically, the inclusion criteria required the availability of backfat thickness (mm) and all carcass-related features, while records containing duplicates, missing values, or inconsistent grading data were excluded to ensure data integrity. Backfat thickness (mm) was used as the target variable, while ten carcass-related features were used as the predictor variables: month (age in months at slaughter), fat number (intramuscular fat score; 1 = none/slight to 9 = abundant), color number (meat color; 1 = pale red to 7 = dark red), fat color (subcutaneous fat color; 1 = white to 7 = yellowish), texture (meat texture; 1 = fine to 5 = coarse), maturity (physiological maturity; 1 = young to 9 = mature), grade score (carcass quality grade; 0–4), aging days (days), Warner–Bratzler shear force (WBsf) (kg), and sex (categorical variable: bull, steer, cow/heifer/hysterectomized heifer). To facilitate model computation and enhance interpretability, categorical variables were numerically encoded before model training. For instance, the "sex" variable was encoded as: bull = 0, steer = 1, and cow/heifer/hysterectomized heifer = 2. The descriptive statistics summarizing the distribution of each variable are presented in Table 2.

Table 2. Descriptive statistics of Hanwoo carcass variables ($N = 386$).

Variables	Count	Mean	Std	Min	Max
Backfat thickness (mm)	386	12.60	5.64	2.00	35.00
Month	386	36.45	15.78	18.00	120.00
Sex	386	1.17	0.58	0.00	2.00
Fat number	386	4.51	1.95	1.00	9.00
Color number	386	5.00	0.53	4.00	7.00
Fat color	386	3.04	0.35	2.00	6.00
Texture	386	1.39	0.55	1.00	4.00
Maturity	386	3.15	1.87	2.00	9.00
Grade score	386	1.22	0.82	0.00	3.00
Aging days	386	6.19	12.12	0.00	60.00
WBsf (kg)	386	3.23	0.86	1.24	6.86

Note: Month: age in months at slaughter; sex: bull = 0, steer = 1, cow/heifer = 2; fat number: 1 = none/slight to 9 = abundant; color number: 1 = pale red to 7 = dark red; fat color: 1 = white to 7 = yellowish; texture: 1 = fine to 5 = coarse; maturity: 1 = young to 9 = mature; grade score: 0–4; WBsf: Warner–Bratzler shear force, kg.

Backfat thickness showed a mean of 12.60 ± 5.64 mm, ranging from 2.00 mm to 35.00 mm. Additionally, the intramuscular fat score averaged 4.51 ± 1.95, meat color averaged 5.00 ± 0.53, fat color grade averaged 3.04 ± 0.35, texture score averaged 1.39 ± 0.55, maturity averaged 3.15 ± 1.87, quality grade averaged 1.22 ± 0.82, aging days averaged 6.19 ± 12.12 days, and the WBsf averaged 3.23 ± 0.86 kg. To enable model development and performance assessment, the dataset was randomly partitioned into a training set (80%, 309 instances) and a testing set (20%, 77 instances), ensuring a representative distribution of all categorical and continuous features across both sets.

Model Selection

We evaluated 14 ML regression algorithms to predict backfat thickness: linear regression[29], ridge regression[16], lasso regression[39], Bayesian ridge regression[27], k-nearest neighbors (KNN)[10], support vector regression (SVR)[11], decision tree regression[4], random forest[5], gradient boosting machine[14], AdaBoost[12], extreme gradient boosting (XGBoost)[8], LightGBM[20], CatBoost[34], and multi-layer perceptron (MLP)[37]. These models were selected because they represent widely used and well-established ML algorithms covering diverse learning paradigms, including linear and regularized regression, kernel-based approaches, ensemble methods, and neural networks. This enables a balanced and comprehensive comparison of model complexity, interpretability, and predictive performance using moderate-sized livestock datasets.

SVR Hyperparameter Tuning

The performance of ML algorithms is highly dependent on the appropriate selection of hyperparameters. Among the 14 ML models, SVR exhibited the highest predictive accuracy. Accordingly, a detailed hyperparameter optimization procedure was performed using an extensive grid search across several critical SVR parameters: kernel type (radial basis function (RBF), polynomial, and sigmoid), regularization parameter $C \in \{0.1, 1, 10, 100, 1000\}$, kernel coefficient $\gamma \in \{\texttt{scale}, \texttt{auto}, 0.001, 0.01, 0.1, 1.0\}$, and polynomial degree $\in \{2, 3, 4\}$ (polynomial kernel only). The best-performing SVR model employed the Gaussian RBF kernel with optimized hyperparameters, achieving the lowest MAE of 2.7962.

Model Evaluation Metrics

Model performance was evaluated primarily using MAE, expressed in the same unit as backfat thickness (mm). MAE was chosen because it directly measures the average magnitude of prediction errors without directional bias, providing an intuitive and interpretable indication of model accuracy in carcass assessments[7,41]. Moreover, MAE is less sensitive to extreme deviations than mean squared error (MSE), making it a more robust indicator for datasets that may include biological or measurement variability[41].

In addition to MAE, the coefficient of determination (R^2) was used as a complementary evaluation metric to assess each model's explanatory power. R^2 quantifies the proportion of variance in backfat thickness explained by the model. While MAE indicates absolute prediction accuracy, R^2 provides insights into the overall fit of the model and its ability to capture relationships among variables. A higher R^2 value implies stronger predictive fidelity and a greater capacity to explain variability in carcass traits, offering a balanced interpretation of model performance beyond mere error magnitude.

Statistical Analyses

Preliminary statistical analyses were performed to examine the relationships among carcass traits and to identify potential sources of multicollinearity before model development. Specifically, Pearson's correlation coefficients (r) were calculated to quantify linear relationships between backfat thickness and each of the ten predictor variables. The statistical significance of each correlation was assessed using p-values, with thresholds of $p < 0.05$, $p < 0.01$, and $p < 0.001$ indicating increasing levels of significance. A correlation heatmap was generated to visualize the strength and direction of pairwise correlations, with significance levels marked by asterisks (*, **, ***).

Model evaluation protocol and bias control

To ensure fair model comparison, all algorithms were trained and evaluated using the same predictor set, the same target definition, and an identical train–test partition. Hyperparameter tuning was conducted using only the training data, and the held-out test set was reserved for final performance reporting to reduce optimistic bias[17,22]. Model performance was assessed using both MAE and R^2 to jointly capture absolute error magnitude and explained variance[7,41]. In addition, subgroup-oriented interpretation was conducted using sex-stratified partial dependence analysis to examine whether dominant predictor effects were consistent across biologically relevant strata.

Acknowledgements

This work was carried out with the support of "Cooperative Research Program for Agriculture Science & Technology Development (Project title: Advancement of Hanwoo beef tracing and management technique based on data, Project No. PJ0170202022)" Rural Development Administration, Republic of Korea.

Author contributions statement

T.Y. conceived and designed the study. T.Y. and D.J. collected and pre-processed the data. T.Y. developed the machine learning models and conducted the experiments. J.Y. provided domain expertise in animal science and contributed to data interpretation. W.L. supervised the research. All authors reviewed and approved the manuscript.

Funding

This work was carried out with the support of "Cooperative Research Program for Agriculture Science & Technology Development (Project title: Advancement of Hanwoo beef tracing and management technique based on data, Project No. PJ0170202022)" Rural Development Administration, Republic of Korea.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Data availability: The data used in this research are not publicly available but can be provided upon reasonable request.

References

1. Baldassini, W. et al. Meat quality traits of Nellore bulls according to different degrees of backfat thickness: a multivariate approach. *Animal Production Science* **57**, 363–370 (2016).
2. Alam, M. Z. et al. Genome-wide association study to identify QTL for carcass traits in Korean Hanwoo cattle. *Animals* **13**, 2737 (2023).
3. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13**, 281–305 (2012).
4. Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. *Classification and Regression Trees* (Wadsworth International Group, Belmont, CA, 1984).
5. Breiman, L. Random forests. *Machine Learning* **45**, 5–32 (2001).
6. Chafai, N., Hayah, I., Houaga, I. & Badaoui, B. A review of machine learning models applied to genomic prediction in animal breeding. *Frontiers in Genetics* **14**, 1150596 (2023).
7. Chai, T. & Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development* **7**, 1247–1250 (2014).
8. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 785–794 (2016).
9. Cho, H. et al. Repeatability of feed efficiency and its relationship with carcass traits in Hanwoo steers during their entire growing and fattening period. *Animal Bioscience* **37**, 1568–1580 (2024).
10. Cover, T. M. & Hart, P. E. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* **13**, 21–27 (1967).
11. Drucker, H. et al. Support vector regression machines. *Advances in Neural Information Processing Systems* **9**, 155–161 (1997).
12. Freund, Y. & Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* **55**, 119–139 (1997).
13. Frezarim, G. B. et al. Multi-omics integration identifies molecular markers and biological pathways for carcass and meat quality traits in Nellore cattle. *Scientific Reports* **15**, 10467 (2025).
14. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of Statistics* **29**, 1189–1232 (2001).
15. Hay, E. H. Machine learning for the genomic prediction of growth traits in a composite beef cattle population. *Animals* **14**, 3014 (2024).
16. Hoerl, A. E. & Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).
17. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R* (Springer, New York, 2013).
18. Jang, J.-H., Lee, H.-D., Kim, J.-J. & Haque, M. A. Evaluation of genomic breeding values and accuracy for carcass traits in Korean Hanwoo cows using whole-genome SNP chip panels. *Mammalian Genome* **36**, 827–841 (2025).
19. Kang, S. M., Lee, I. S. & Lee, S. K. Carcass characteristics of Hanwoo (Korean cattle) from different sex conditions, raising altitudes and slaughter seasons. *Livestock Science* **123**, 283–287 (2009).
20. Ke, G. et al. LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems* **30**, 3149–3157 (2017).
21. Koo, Y., Alkhoder, H., Choi, T.-J., Liu, Z. & Reents, R. Genomic evaluation of carcass traits of Korean beef cattle Hanwoo using a single-step marker effect model. *Journal of Animal Science* **101**, skad104 (2023).
22. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proc. 14th Int. Joint Conf. Artificial Intelligence*, 1137–1145 (1995).
23. Kuhn, M. & Johnson, K. *Applied Predictive Modeling* (Springer, New York, 2013).

24. Lee, J.-M. et al. Effect of backfat thickness on the carcass grade factors and carcass price in Hanwoo cows and steers. *Food Science of Animal Resources* **31**, 280–289 (2011).

25. Lopes, F. et al. Genome-enabled prediction of meat and carcass traits using Bayesian regression, single-step genomic best linear unbiased prediction and blending methods in Nelore cattle. *Animal* **15**, 100006 (2021).

26. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* **30**, 4765–4774 (2017).

27. MacKay, D. J. C. Bayesian interpolation. *Neural Computation* **4**, 415–447 (1992).

28. Mammadova, N. & Keskin, I. Application of the support vector machine to predict subclinical mastitis in dairy cattle. *The Scientific World Journal* **2013**, 603897 (2013).

29. Montgomery, D. C., Peck, E. A. & Vining, G. G. *Introduction to Linear Regression Analysis* (John Wiley & Sons, 2021).

30. Mota, L. F. M. et al. Benchmarking machine learning and parametric methods for genomic prediction of feed efficiency-related traits in Nellore cattle. *Scientific Reports* **14**, 6404 (2024).

31. Mota, L. F. M. et al. Variable selection strategies for genomic prediction of growth and carcass related traits in experimental Nellore cattle herds under different selection criteria. *Scientific Reports* **15**, 22266 (2025).

32. Pedrosa, V. B. et al. Machine learning methods for genomic prediction of cow behavioral traits measured by automatic milking systems in North American Holstein cattle. *Journal of Dairy Science* **107**, 4758–4771 (2024).

33. Pestov, V. Is the k-NN classifier in high dimensions affected by the curse of dimensionality? *Computers & Mathematics with Applications* **65**, 1427–1437 (2013).

34. Prokhorenkova, L. et al. CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems* **31**, 6639–6649 (2018).

35. Ribeiro, M. T., Singh, S. & Guestrin, C. "Why should I trust you?": Explaining the predictions of any classifier. *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 1135–1144 (2016).

36. Ruchay, A. et al. Predicting the body weight of Hereford cows using machine learning. *IOP Conference Series: Earth and Environmental Science*, 012056 (2021).

37. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).

38. Shataee, S., Kalbi, S., Fallah, A. & Pelz, D. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. *International Journal of Remote Sensing* **33**, 6254–6280 (2012).

39. Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B* **58**, 267–288 (1996).

40. Tirink, C. Comparison of Bayesian regularized neural network, random forest regression, support vector regression and multivariate adaptive regression splines algorithms to predict body weight from biometrical measurements in Thalli sheep. *Kafkas University Veterinary Faculty Journal* **28** (2022).

41. Willmott, C. J. & Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research* **30**, 79–82 (2005).

42. Zhao, W. et al. Applications of support vector machine in genomic prediction in pig and maize populations. *Frontiers in Genetics* **11**, 598318 (2020).