Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

Explainable Gradient Boosting for Egg Production Anomaly Detection in Laying Hen Farms

Intra Swadaya Hidayat, Febria Sri Handayani, Yarza Aprizal, Almaheri Adhi Pratama, Muhammad Jhonsen Syaftriandi
Institut Teknologi dan Bisnis Palcomtech, Indonesia
intra.swadaya@palcomtech.ac.id, febria_sri@palcomtech.ac.id, yarza_afrizal@palcomtech.ac.id, rendy_almaheri@palcomtech.ac.id, m_jhonsen@palcomtech.ac.id

Abstract- Small-scale laying-hen farms require analytical tools that can detect abnormal egg-production patterns while remaining interpretable for operational decision support. This study develops and evaluates an explainable gradient boosting framework for low-production anomaly detection using farm-level environmental, management, and production records. The dataset contained 481 observations with eight variables, including flock size, feed amount, ammonia, temperature, humidity, light intensity, noise, and total egg production. Exact-duplicate inspection found 407 duplicated rows; therefore, the 74 unique observations were used as the primary evaluation dataset. Egg production rate was derived by normalizing total egg output by flock size, and low-production anomalies were defined using the 10th percentile threshold, resulting in eight anomaly cases. Supervised tree-based models were evaluated using repeated stratified four-fold cross-validation with 20 repeats, and unsupervised methods were used as anomaly baselines. XGBoost achieved the best full-feature performance, with F1-score of 0.928, recall of 1.000, ROC-AUC of 0.997, and PR-AUC of 0.984. Sensitivity analysis without flock size showed that environmental and management variables retained useful signal. SHAP analysis indicated that flock size, light intensity, and temperature were the dominant predictors. The findings support interpretable operational monitoring of low egg-production patterns, although the results require validation on larger expert-labeled datasets.

Keywords: egg production, anomaly detection, gradient boosting, SHAP, machine learning

1. INTRODUCTION

Egg production anomaly detection is an operational monitoring task that aims to identify unusually low normalized egg output from farm-level environmental, management, and production records. In laying-hen farms, short-term reductions in laying performance may reflect unfavorable temperature and humidity conditions, gas accumulation, lighting variation, feed-related changes, flock-size effects, or broader management [1], [2], [3], [4], [5], [6]. For small-scale farms, these signals are often inspected manually and retrospectively. As a result, abnormal production patterns may be detected late, and the basis for corrective action may remain weak.

Precision livestock farming and smart poultry systems provide the technological basis for data-driven farm monitoring, but sensing infrastructure alone does not solve the anomaly-detection problem. IoT, low-cost sensing, sound-based monitoring, remote poultry-management systems, and computer-vision-based behavior monitoring have been proposed to improve data capture and farm visibility [7], [8], [9], [10]. Machine learning has also shown strong potential for poultry forecasting, laying-rate prediction, explainable egg-production modelling, and health-related monitoring-[1], [11], [12], [13], [14]. However, these studies mainly address prediction, monitoring infrastructure, or disease-risk estimation rather than interpretable low-production anomaly detection.

Existing anomaly-detection research in laying-hen production shows that abnormal egg-production patterns are farm-specific and difficult to define without contextual information. Van Veen et al. [15] developed an adaptive expert-in-the-loop approach using daily production curves, incremental one-class support vector machines, and expert feedback. This work is the closest reference point for the present study. Nevertheless, a complementary gap remains for small, non-temporal, duplicate-prone tabular datasets in which streaming production curves and expert-confirmed anomaly annotations are not available.

Recent precision livestock farming literature emphasizes continuous monitoring, real-time warnings, and the integration of sensors, artificial intelligence, and farm decision support [16], [17]. More recent poultry and smart-farming studies also highlight the need for affordable, explainable, and deployable monitoring tools [18]-[22]. These studies strengthen the motivation for an anomaly-detection framework that is not only accurate but also interpretable and reproducible under limited data conditions.

This study develops and evaluates an explainable gradient boosting framework for operational low-production anomaly detection in small-scale laying-hen farms. The completed experiment transforms total egg production into egg production rate, constructs a low-production anomaly label using the 10th percentile threshold, audits duplicate observations, compares supervised tree-based models with unsupervised anomaly-detection baselines, evaluates two feature settings with repeated stratified cross-validation, and applies SHAP to interpret the best-performing XGBoost model. Because expert-confirmed biological stress or disease annotations are unavailable, the target class is framed as an operational low-production anomaly rather than a confirmed welfare or pathology label.

The contributions of this study are fivefold. First, it formulates egg-production anomaly detection using normalized egg production rate rather than raw egg counts. Second, it constructs and evaluates a reproducible operational anomaly label for tabular poultry datasets without expert-confirmed labels. Third, it compares supervised tree-based learning and

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

unsupervised anomaly baselines after duplicate auditing. Fourth, it performs feature-exclusion sensitivity analysis to test the effect of flock-size information. Fifth, it integrates SHAP and ROC-based analysis to connect predictive performance with interpretable farm-level insight.

2. RESEARCH METHODOLOGY

Figure 1 illustrates the overall research methodology of this study. The methodology consists of three main stages: (1) dataset preparation and audit, (2) anomaly label construction and feature setting, and (3) model development, evaluation, and interpretability analysis. Detailed procedures for each stage are presented in Sections 2.1–2.3.

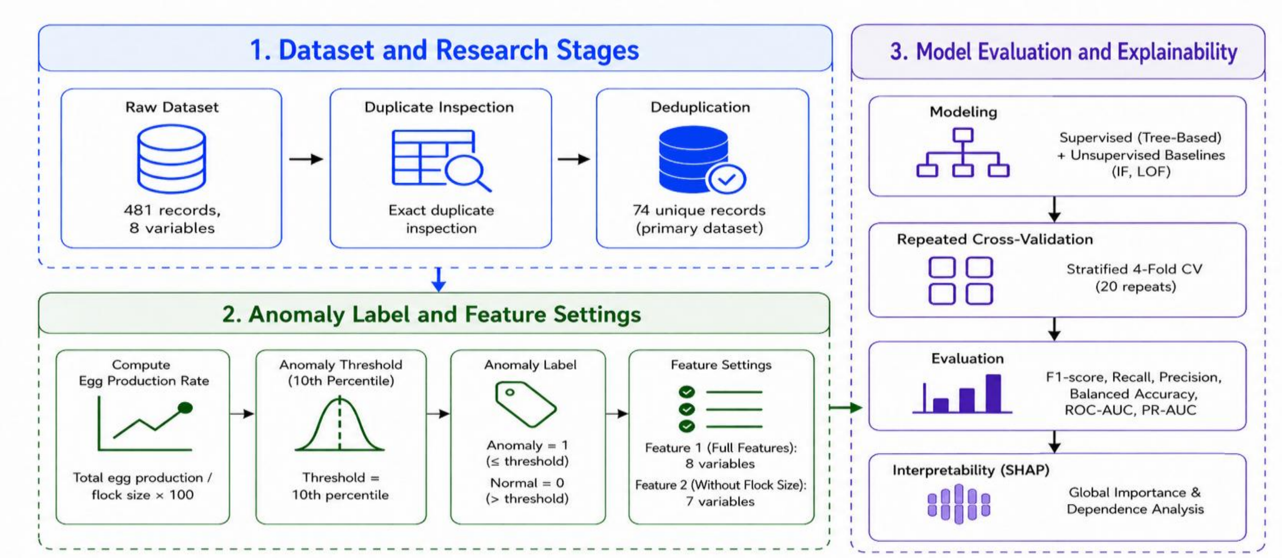

Figure 1. Research methodology framework..

2.1 Dataset and Research Stages

The dataset was obtained from Kaggle under the title Environmental Effect on Egg Production, authored by Md Anowar Hossen Faysal. The downloaded file used in this study was Egg_Production.csv. The dataset is publicly available on Kaggle under the CC0: Public Domain license, allowing reuse for research and analytical purposes. The dataset contained 481 observations from small-scale laying-hen farm records, and Table 1 lists the variables and analytical roles used in this study.

Initial inspection found no missing values. However, exact-duplicate inspection identified 407 duplicated records, leaving 74 unique observations. Because duplicated records can inflate model performance when identical observations appear in both training and test partitions, the deduplicated dataset was used as the primary evaluation dataset.

Table 1. Dataset variables and analytical roles

Variable	Description	Role
Amount_of_chicken	Number of chickens in the observed flock	Input feature
Amount_of_Feeding	Amount of feed provided	Input feature
Ammonia	Ammonia concentration indicator	Input feature
Temperature	Farm temperature	Input feature
Humidity	Farm humidity	Input feature
Light_Intensity	Light intensity indicator	Input feature
Noise	Noise level indicator	Input feature
Total_egg_production	Total egg output	Outcome used to derive production rate

2.2 Anomaly Label and Feature Settings

Egg production rate was calculated as total egg production divided by flock size and multiplied by 100. Operational anomalies were defined as observations whose egg production rate was at or below the 10th percentile. This thresholding approach was selected because the dataset did not include expert-confirmed disease, stress, or welfare labels.

Two feature settings were evaluated. The full feature setting used Amount_of_chicken, Amount_of_Feeding, Ammonia, Temperature, Humidity, Light_Intensity, and Noise. The second setting excluded Amount_of_chicken to test whether model performance depended strongly on flock-size information, since egg production rate already normalizes total egg output by flock size.

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

2.3 Model Evaluation and Explainability

Operational anomaly detection was treated as a binary classification task. Decision Tree and Random Forest were used as supervised baselines, while Gradient Boosting, XGBoost, and LightGBM were evaluated as boosting models. Isolation Forest and Local Outlier Factor were included as unsupervised anomaly-detection baselines.
Supervised models were evaluated using repeated stratified four-fold cross-validation with 20 repeats. Precision, recall, F1-score, balanced accuracy, ROC-AUC, and PR-AUC were used as evaluation metrics. Model performance was then summarized across full-feature and environment-management-only settings.
Table 2 summarizes the evaluation design, models, and metrics used for supervised learning, unsupervised baseline comparison, and explainability analysis.

Table 2. Experimental design

Task	Models	Main metrics
Operational anomaly classification, full features	Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM	Precision, recall, F1-score, balanced accuracy, ROC-AUC, PR-AUC
Operational anomaly classification, no flock-size feature	Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM	Precision, recall, F1-score, balanced accuracy, ROC-AUC, PR-AUC
Unsupervised anomaly baseline	Isolation Forest, Local Outlier Factor	Precision, recall, F1-score, balanced accuracy, ROC-AUC, PR-AUC
Explainability analysis	XGBoost with SHAP/TreeSHAP	Mean absolute SHAP value and ranked feature contribution

3. RESULT AND DISCUSSION

3.1 Dataset Audit and Operational Anomaly Label

The dataset audit confirmed that the original dataset contained 481 observations and eight variables. No missing values were observed. The duplicate audit was methodologically important because 407 exact duplicates were found, leaving 74 unique observations. The original dataset produced 56 anomaly observations using the 10th percentile production-rate threshold, whereas the deduplicated dataset produced eight anomaly observations.
In the deduplicated dataset, egg production rate ranged from 55.33% to 82.76%, with a mean of 72.05%. The 10th percentile threshold was 65.06%. Therefore, the anomaly label represents low egg-production-rate observations and should not be interpreted as confirmed disease, biological stress, or welfare status.
Table 3 reports the effect of duplicate removal on the effective sample size and operational anomaly count.

Table 3. Dataset audit and operational anomaly label summary

Dataset version	Rows	Anomaly cases	Q10 threshold
Original data	481	56 (11.64%)	64.81%
Deduplicated data	74	8 (10.81%)	65.06%
Primary evaluation data	74	8 (10.81%)	65.06%

Figure 2 illustrates the egg production rate distribution and the 10th percentile threshold used to define operational anomalies.

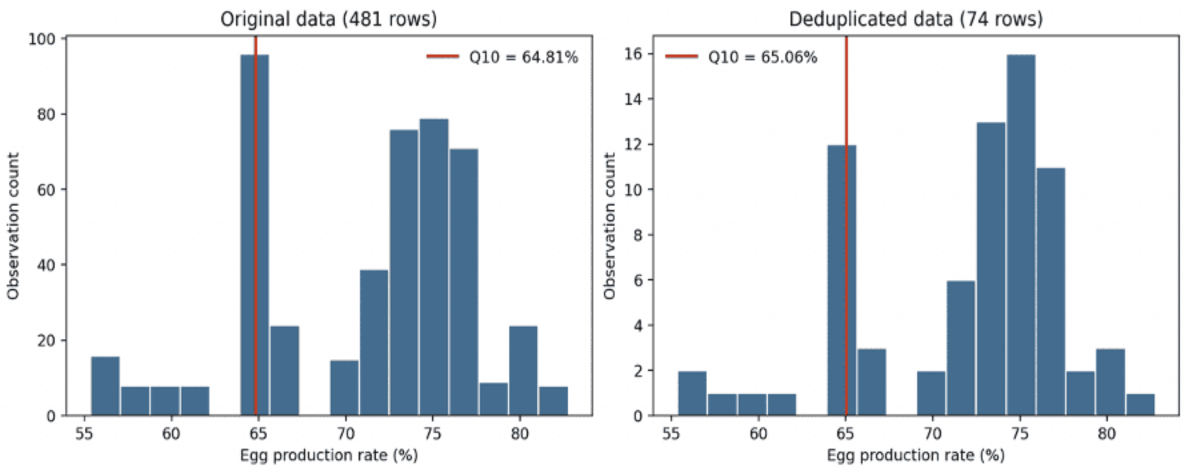

Figure 2. Distribution of egg production rate and the 10th percentile operational anomaly threshold.

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

3.2 Model Performance

Table 4 reports the cross-validated supervised classification metrics for all five supervised models under the full-feature and environment-management-only settings. XGBoost achieved the best full-feature performance with an F1-score of 0.928 and balanced accuracy of 0.988, while Figure 3 visualizes the F1-score comparison across supervised models and feature settings.

When Amount_of_chicken was removed from the predictors, the Decision Tree model achieved the highest environment-management-only F1-score of 0.860, while XGBoost retained strong discrimination with ROC-AUC of 0.971 and PR-AUC of 0.846. Because the anomaly label was derived from egg production rate, which already normalizes egg output by flock size, the full-feature result should be interpreted as an upper-bound model, whereas the environment-management-only setting provides a more operationally relevant validation of the environmental and management signals.

Table 4. Cross-validated supervised classification metrics on the deduplicated dataset

Feature setting / model	Precision	Recall	F1-score	Bal. Acc.
XGBoost (full features)	0.883 +/- 0.175	1.000 +/- 0.000	0.928 +/- 0.111	0.988 +/- 0.019
Random Forest (full features)	0.994 +/- 0.056	0.869 +/- 0.221	0.910 +/- 0.152	0.934 +/- 0.111
Decision Tree (full features)	0.914 +/- 0.182	0.881 +/- 0.214	0.875 +/- 0.176	0.932 +/- 0.109
Gradient Boosting (full features)	0.961 +/- 0.180	0.819 +/- 0.267	0.861 +/- 0.214	0.907 +/- 0.133
LightGBM (full features)	0.609 +/- 0.233	0.831 +/- 0.263	0.682 +/- 0.214	0.878 +/- 0.135
Decision Tree (env-management only)	0.803 +/- 0.252	0.962 +/- 0.155	0.860 +/- 0.204	0.960 +/- 0.090
XGBoost (env-management only)	0.737 +/- 0.266	0.906 +/- 0.226	0.783 +/- 0.220	0.926 +/- 0.114
Random Forest (env-management only)	0.719 +/- 0.346	0.688 +/- 0.350	0.676 +/- 0.314	0.829 +/- 0.177
LightGBM (env-management only)	0.488 +/- 0.218	0.819 +/- 0.278	0.584 +/- 0.200	0.846 +/- 0.142
Gradient Boosting (env-management only)	0.522 +/- 0.413	0.488 +/- 0.398	0.469 +/- 0.349	0.728 +/- 0.187

Table 5 reports the threshold-independent supervised discrimination metrics. The PR-AUC column is included because the anomaly class is small after deduplication.

Table 5. Cross-validated supervised ROC-AUC and PR-AUC on the deduplicated dataset

Feature setting / model	ROC-AUC	PR-AUC
XGBoost (full features)	0.997 +/- 0.010	0.984 +/- 0.061
Random Forest (full features)	0.988 +/- 0.029	0.955 +/- 0.096
Decision Tree (full features)	0.932 +/- 0.109	0.823 +/- 0.243
Gradient Boosting (full features)	0.935 +/- 0.116	0.883 +/- 0.202
LightGBM (full features)	0.961 +/- 0.052	0.819 +/- 0.208
Decision Tree (env-management only)	0.960 +/- 0.090	0.799 +/- 0.258
XGBoost (env-management only)	0.971 +/- 0.054	0.846 +/- 0.219
Random Forest (env-management only)	0.977 +/- 0.039	0.866 +/- 0.205
LightGBM (env-management only)	0.924 +/- 0.069	0.662 +/- 0.229
Gradient Boosting (env-management only)	0.962 +/- 0.047	0.802 +/- 0.220

Table 6 shows that the unsupervised baselines were weaker than supervised models. Isolation Forest and Local Outlier Factor produced low F1-scores and weak PR-AUC values, especially in the environment-management-only setting. This indicates that the low-production anomaly label is better captured as a supervised classification problem than as a generic density-based or isolation-based outlier problem.

Table 6. Unsupervised anomaly-detection baseline performance on the deduplicated dataset

Feature setting / model	Precision	Recall	F1-score	Bal. Acc.	ROC-AUC	PR-AUC
Isolation Forest (full features)	0.125	0.125	0.125	0.509	0.513	0.136
Local Outlier Factor (full features)	0.000	0.000	0.000	0.439	0.511	0.114
Isolation Forest (environment-management only)	0.000	0.000	0.000	0.439	0.324	0.087
Local Outlier Factor (environment-management only)	0.000	0.000	0.000	0.439	0.305	0.083

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

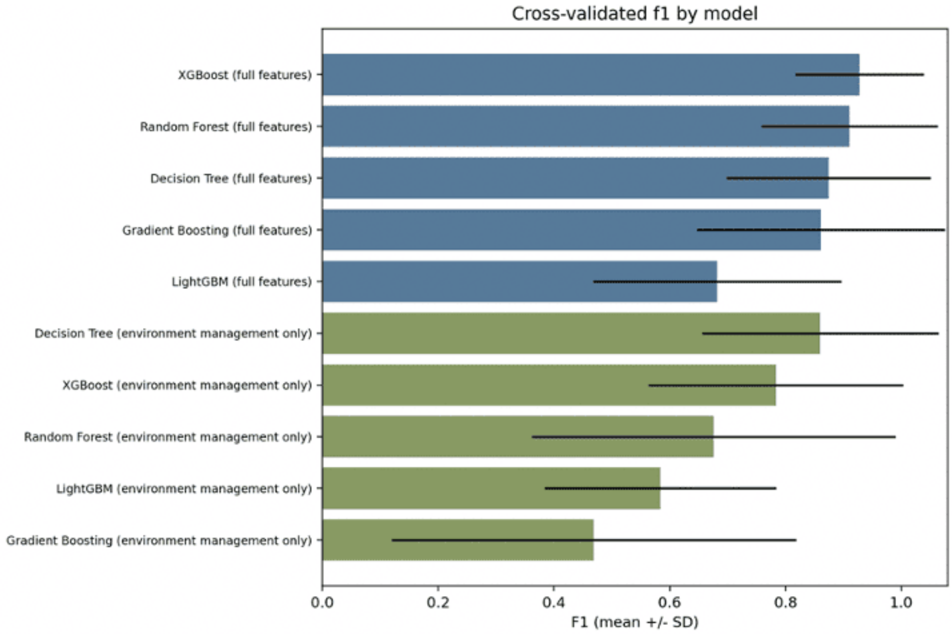

Figure 3. Cross-validated F1-score comparison across supervised models and feature settings.

Figure 4 compares ROC-AUC values across all supervised models, showing that high discrimination was observed in both full-feature and environment-management-only settings.

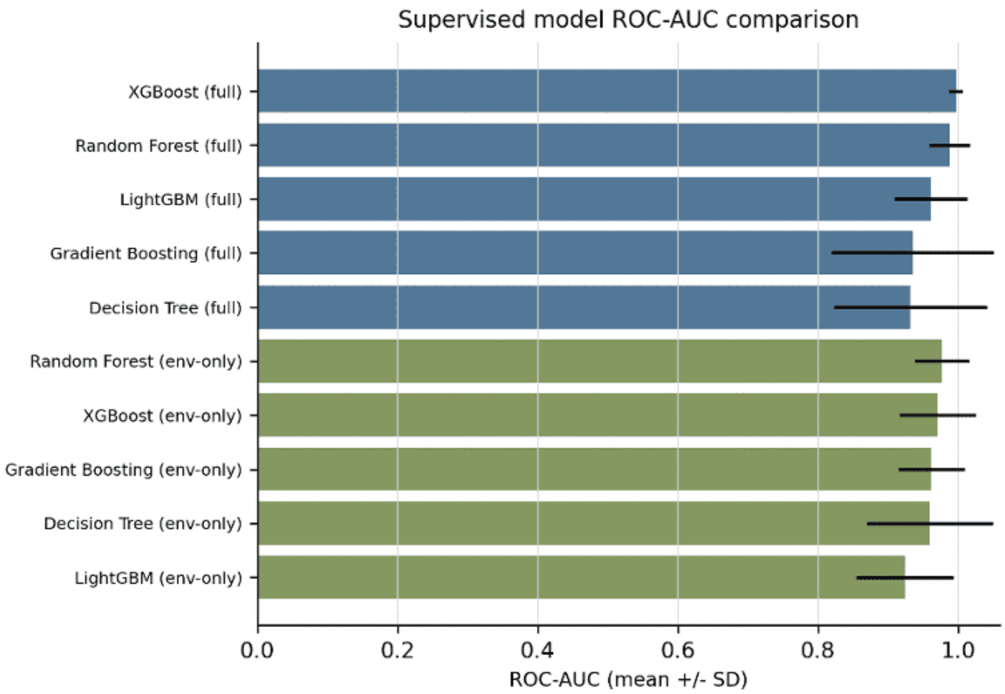

Figure 4. Cross-validated ROC-AUC comparison across supervised models and feature settings.

3.3 Explainability and Interpretation

Figure 5 shows that SHAP analysis of the full-feature XGBoost model ranked Amount_of_chicken, Light_Intensity, and Temperature as the three most influential predictors. The dominance of Amount_of_chicken indicates that the full-feature model may partly exploit a mathematical relationship between flock size and the derived egg-production-rate label. Therefore, the full-feature result should not be treated as deployment-ready evidence on its own.

Figure 6 shows that after Amount_of_chicken was excluded, Temperature, Humidity, and Light_Intensity became the most influential variables. This result supports the environment-management-only analysis as a more conservative validation that non-flock-size variables still contain useful signal for low-production anomaly detection. However,

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

because the dataset is small and observational, SHAP values should be interpreted as model-based associations rather than causal evidence.

Compared with prior poultry forecasting and smart-farming studies, this work contributes a narrower but practically useful framework for operational low-production anomaly detection in a small tabular dataset. The findings are preliminary and should not be interpreted as evidence that the framework is ready for farm deployment. The small number of unique observations and anomaly cases limits generalizability; future work should validate the framework using larger timestamped datasets, expert-confirmed anomaly labels, and independent farm-level test data.

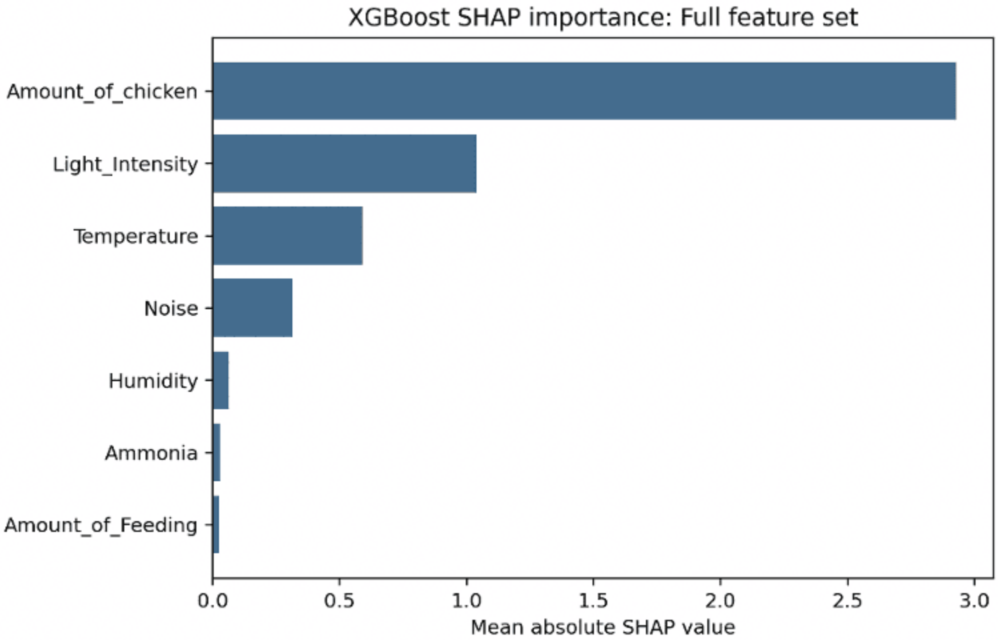

Figure 5. XGBoost SHAP feature importance using the full feature set.

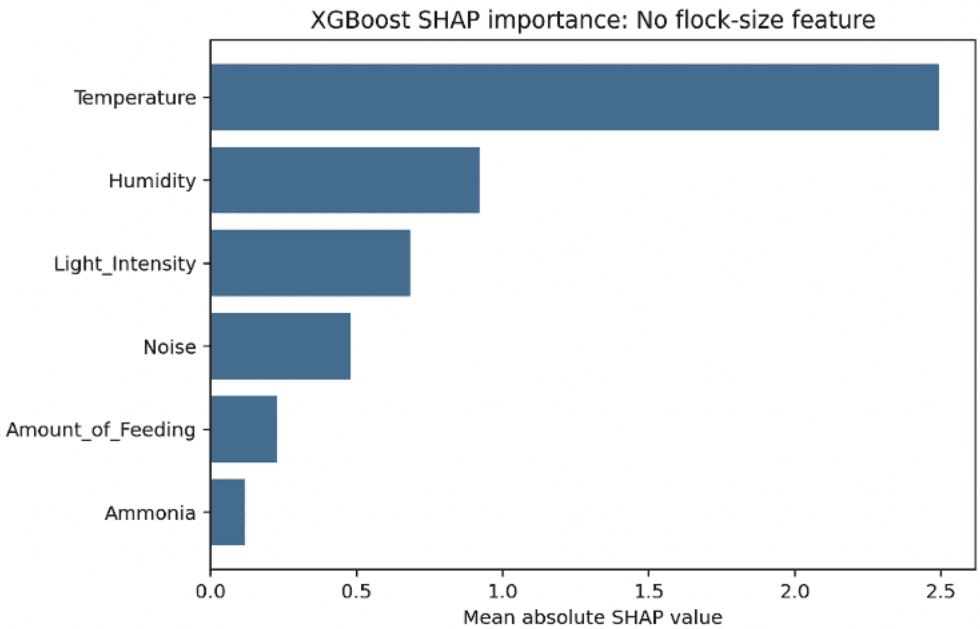

Figure 6. XGBoost SHAP feature importance after excluding flock-size information.

4. CONCLUSION

This study developed and evaluated an explainable gradient boosting framework for operational low-production anomaly detection in small-scale laying-hen farms. After duplicate removal, the dataset contained 74 unique observations and eight anomaly cases based on the 10th percentile of normalized egg production rate. XGBoost achieved the strongest

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

full-feature performance, with an F1-score of 0.928 and ROC-AUC of 0.997, while the environment-management-only analysis showed that non-flock-size variables still contained useful anomaly-detection signal. SHAP analysis showed that flock size, light intensity, and temperature were the most influential predictors in the full model, while temperature, humidity, and light intensity dominated in the environment-management-only model. These results should be regarded as preliminary because the dataset is small, non-temporal, and lacks expert-confirmed anomaly labels. Future research should evaluate the framework on larger, timestamped, multi-farm datasets with validated biological, welfare, or management annotations before operational deployment.

REFERENCES

[1] H. Yin *et al.*, 'A Novel Method to Predict Laying Rate Based on Multiple Environment Variables', *IEEE Access*, vol. 9, pp. 115488–115496, 2021, doi: 10.1109/ACCESS.2021.3105189.

[2] Y. A. Adejola, T. Z. Sibanda, I. Ruhnke, J. Boshoff, S. Pokhrel, and M. Welch, 'Analyzing the Risk of Short-Term Losses in Free-Range Egg Production Using Commercial Data', *Agriculture*, vol. 15, no. 7, p. 743, 2025, doi: 10.3390/agriculture15070743.

[3] L. Paura, I. Arhipova, L. Jankovska, N. Bumanis, G. Vitols, and M. Adjutovs, 'Evaluation and association of laying hen performance, environmental conditions and gas concentrations in barn housing system', *Ital. J. Anim. Sci.*, vol. 21, no. 1, pp. 694–701, 2022, doi: 10.1080/1828051X.2022.2056528.

[4] A. K. Carranza-Díaz, D. Larios-Martínez, A. Ruiz-González, S. Godbout, and S. Fournel, 'Modelling greenhouse gas and ammonia emissions from housing and manure storage in three laying hen production systems', *Biosyst. Eng.*, vol. 265, p. 104423, 2026, doi: 10.1016/j.biosystemseng.2026.104423.

[5] M. Li *et al.*, 'Multivariate analysis for data mining to characterize poultry house environment in winter', *Poult. Sci.*, vol. 103, no. 5, p. 103633, 2024, doi: 10.1016/j.psj.2024.103633.

[6] A. M. Pranta, S. M. A. Islam, and R. I. Khan, 'Development of a sensor-integrated AI automation model for decision-based heat stress management in layer chickens under subtropical climate conditions', *Smart Agricultural Technology*, vol. 12, p. 101306, 2025, doi: 10.1016/j.atech.2025.101306.

[7] R. B. Bist, S. Subedi, X. Yang, and L. Chai, 'A Novel YOLOv6 Object Detector for Monitoring Piling Behavior of Cage-Free Laying Hens', *AgriEngineering*, vol. 5, no. 2, pp. 905–923, 2023, doi: 10.3390/agriengineering5020056.

[8] R. Finistrosa, C. Mañoso, Á. P. de Madrid, and M. Romero, 'Low-Cost IoT and LoRaWAN-Based System for Laying Hen Identification in Family Poultry Farms', *Applied Sciences*, vol. 15, no. 9, p. 4856, 2025, doi: 10.3390/app15094856.

[9] A. Amirivojdan, A. Nasiri, S. Zhou, Y. Zhao, and H. Gan, 'ChickenSense: A Low-Cost Deep Learning-Based Solution for Poultry Feed Consumption Monitoring Using Sound Technology', *AgriEngineering*, vol. 6, no. 3, pp. 2115–2129, 2024, doi: 10.3390/agriengineering6030124.

[10] J. Chigwada, F. Mazunga, C. Nyamhere, V. Mazheke, and N. Taruvinga, 'Remote poultry management system for small to medium scale producers using IoT', *Sci. Afr.*, vol. 18, p. e01398, 2022, doi: 10.1016/j.sciaf.2022.e01398.

[11] N. Bumanis, A. Kviesis, L. Paura, I. Arhipova, and M. Adjutovs, 'Hen Egg Production Forecasting: Capabilities of Machine Learning Models in Scenarios with Limited Data Sets', *Applied Sciences*, vol. 13, no. 13, p. 7607, 2023, doi: 10.3390/app13137607.

[12] H. Ji, Y. Xu, and G. Teng, 'Predicting egg production rate and egg weight of broiler breeders based on machine learning and Shapley additive explanations', *Poult. Sci.*, vol. 104, no. 1, p. 104458, 2025, doi: 10.1016/j.psj.2024.104458.

[13] Y. Liu *et al.*, 'A Machine Learning Framework Based on Extreme Gradient Boosting to Predict the Occurrence and Development of Infectious Diseases in Laying Hen Farms, Taking H9N2 as an Example', *Animals*, vol. 13, no. 9, p. 1494, 2023, doi: 10.3390/ani13091494.

[14] M. Welch, T. Z. Sibanda, J. D. S. Vilela, M. Kolakshyapati, D. Schneider, and I. Ruhnke, 'An Initial Study on the Use of Machine Learning and Radio Frequency Identification Data for Predicting Health Outcomes in Free-Range Laying Hens', *Animals*, vol. 13, no. 7, p. 1202, 2023, doi: 10.3390/ani13071202.

[15] L. A. van Veen, H. van den Brand, A. C. M. van den Oever, B. Kemp, and A. Youssef, 'An adaptive expert-in-the-loop algorithm for flock-specific anomaly detection in laying hen production', *Comput. Electron. Agric.*, vol. 229, p. 109755, 2025, doi: 10.1016/j.compag.2024.109755.

[16] D. Berckmans, 'General introduction to precision livestock farming', *Animal Frontiers*, vol. 7, no. 1, pp. 6–11, Jan. 2017, doi: 10.2527/af.2017.0102.

[17] S. Neethirajan, 'The role of sensors, big data and machine learning in modern animal farming', Aug. 01, 2020, *Elsevier B.V.* doi: 10.1016/j.sbsr.2020.100367.

[18] A. Rajbongshi *et al.*, 'Leveraging explainable AI for sustainable agriculture: a comprehensive review of recent advances', *Artif. Intell. Rev.*, Jan. 2026, doi: 10.1007/s10462-025-11459-5.

Bulletin of Information Technology (BIT)
Vol 7, No 3, September 2026, Hal. 323 – 330
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.3071
https://journal.fkpt.org/index.php/BIT

[19] R. B. Bist, D. Wang, L. Chai, and Y. Xiong, 'Precision Farming Technologies for Monitoring Livestock and Poultry', Feb. 01, 2026, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/agriengineering8020064.

[20] S. Neethirajan, S. Scott, C. Mancini, X. Boivin, and E. Strand, 'Human-computer interactions with farm animals—enhancing welfare through precision livestock farming and artificial intelligence', *Front. Vet. Sci.*, vol. 11, 2024, doi: 10.3389/fvets.2024.1490851.

[21] A. Cartolano, A. Cuzzocrea, and G. Pilato, 'Analyzing and assessing explainable AI models for smart agriculture environments', *Multimed. Tools Appl.*, vol. 83, no. 12, pp. 37225–37246, Apr. 2024, doi: 10.1007/s11042-023-17978-z.

[22] Y. A. Adejola, T. Z. Sibanda, I. Ruhnke, J. Boshoff, S. Pokhrel, and M. Welch, 'Forecasting egg production performance and fluctuations in commercial free-range poultry systems using a random forest model', *Smart Agricultural Technology*, vol. 12, Dec. 2025, doi: 10.1016/j.atech.2025.101380.