Poultry Science

Machine learning predictive modeling for condemnation risk assessment in antibiotic-free raised broilers

Pranee Pirompud,* Panneepa Sivapirunthep,† Veerasak Punyapornwithaya,‡ and Chanporn Chaosap†,[1]

*Doctoral Program in Innovative Tropical Agriculture, Department of Agricultural Education, Faculty of Industrial Education and Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand 10520; †Department of Agricultural Education, Faculty of Industrial Education and Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand 10520; and ‡Department of Veterinary Biosciences and Veterinary Public Health, Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai 50100, Thailand

ABSTRACT The condemnation of broiler carcasses in the poultry industry is a major challenge and leads to significant financial losses and food waste. This study addresses the critical issue of condemnation risk assessment in the discarding of antibiotic-free raised broilers using machine learning (**ML**) predictive modeling. In this study, ML approaches, specifically least absolute shrinkage and selection operator (**LASSO**), classification tree (**CT**), and random forests (**RF**), were used to evaluate and compare their effectiveness in predicting high condemnation rates. The dataset of 23,959 truckloads from 2021 to 2022 contained 14 independent variables covering the rearing, catching, transportation, and slaughtering phases. Condemnation rates between 0.26% and 25.99% were used as the dependent variable for the analysis, with the threshold for a high conviction rate set at 3.0%. As high condemnation rates were in the minority (8.05%), sampling methods such as random over sampling (**ROS**), random under sampling (**RUS**), both sampling (**BOTH**), and random over sampling example (**ROSE**) were used to account for imbalanced datasets. The results showed that RF with RUS performed better than the other models for balanced datasets. In this study, mean body weight, weight per crate, mortality and culling rates, and lairage time were identified as the 4 most important variables for predicting high condemnation rates. This study provides valuable insights into ML applications for predicting condemnation rates in antibiotic-free raised broilers and provides a framework to improve decision-making processes in establishing farm management practices to minimize economic losses in the poultry industry. The proposed methods are adaptable for different broiler producers, which increases their applicability in the industry.

Key words: condemnation rate, weight per crate, mortality and culling rate, lairage time, sampling technique

2024 Poultry Science 103:104270
https://doi.org/10.1016/j.psj.2024.104270

INTRODUCTION

The condemnation of broiler carcasses poses significant financial losses and food waste challenges, presenting a critical issue within the poultry industry. Broiler condemnation occurs during the postmortem inspection process at slaughterhouses, where either the entire carcass or specific portions are rejected based on the severity of defects. Diseases and injuries are common causes of condemnation, with various factors such as contamination, traumatic injuries, septicemia, dermatitis, and ascitic syndrome contributing to partial or complete condemnation (Hortêncio, et al., 2022).

The condemnation rate in broilers is influenced by various factors, including health status, weather conditions, weight, age, stocking density and management practices (Buzdugan et al., 2020; Junghans et al., 2022; Lupo et al., 2009). Health status, a critical predictor of outcomes such as mortality, morbidity and functional status, plays a significant role in these rates (Düpjan and Dawkins, 2022). Pirompud et al. (2023) used traditional statistical methods to investigate risk factors for condemnation in antibiotic-free broiler rearing and identified the most important influencing factors, including transport time, sex, slaughter age, mortality rate, weight per crate, mean body weight, feed withdrawal time and rearing stocking density.

Despite numerous studies on risk factors, reducing broiler condemnation rates remains challenging due to the complex interplay of these factors. Condemnation

Received February 14, 2024.

Accepted August 22, 2024.

[1]Corresponding author: chanporn.ch@kmitl.ac.th

rates are influenced by variables such as weight, age, weather, health status, diseases and preslaughter handling practices such as feed withdrawal, lairage, and transport times (Lupo et al., 2009; Buzdugan et al., 2020; Junghans et al., 2022). However, in the broiler industry, there is little evidence of predictive models that classify and predict condemnation rates, prioritizing influencing factors. Two primary methods for predicting outcomes based on multiple variables are regression techniques (e.g. logistic regression, LASSO) and machine learning (**ML**) algorithms. ML, known for its ability to handle nonlinear systems and large datasets, has gained prominence in various fields (Sampson et al., 2011; Shahinfar et al., 2014). Within the field of statistical modeling, the LASSO model is regarded as a regression strategy since it enhances the accuracy and interpretability of regression models by applying machine learning concepts, particularly regularization techniques (Tibshirani, 1996; Hastie et al., 2009). ML algorithms have been shown to be effective in predicting various outcomes in livestock research (Avizheh et al., 2023; Magalhães et al., 2023; Punyapornwithaya et al., 2022; Pirompud et al., 2024). However, to our knowledge, there are no ML predictive models that specifically classify and predict condemnation rates in broilers.

Machine learning, a branch of data science, is about training computers to make predictions based on data (Kuhle, et al., 2018). Better predictions are achieved through machine learning techniques that adapt and learn from incoming data. While traditional statistical methods are valuable, their effectiveness may be limited when dealing with the complexity of relationships, especially in scenarios with high condemnation rates. Therefore, exploring the potential of machine learning is crucial given its proven superiority, especially when processing large and complex datasets.

In this study, ML approaches, least absolute shrinkage and selection operator (**LASSO**), classification tree (**CT**), and random forests (**RF**) are used to evaluate and compare their ability to predict high condemnation rates. LASSO, a statistical method for regression in datasets with numerous variables, uses a penalty term to shrink less influential variables to zero. This prevents overfitting, facilitates variable selection, and results in an accurate and interpretable model, making LASSO valuable for statistical modeling and machine learning (Ghosh, et al., 2021). CT and RF are machine learning methods that establish rules for categorizing and predicting outcomes. The CT approach creates a decision tree with binary answers ("yes" or "no") that shows the path from the root main node to the leaf node. The "what-if" results of CT are easy to understand (Blockeel, et al., 2023; Mlambo, et al., 2022). The RF algorithm, on the other hand, generates a multidimensional decision tree and the final model outcome is selected by majority voting based on the results of the individual trees (Ahmad, et al., 2018). Imbalanced datasets, where the minority class constitutes less than 40% of the total data (Google Developer, 2022), can lead to biased ML models that favor the majority class. To address this, various data balancing strategies are applied, including random over sampling (**ROS**), random under sampling (RUS), both sampling (**BOTH**) and synthetic sampling or random over sampling example (**ROSE**). These methods help to improve the predictive accuracy of models when dealing with imbalanced data (Demir and Şahin, 2022; Lunardon et al., 2014; Pirompud et al., 2024).

Despite the limited use of ML in predicting condemnations in poultry production, this study aims to bridge the gap by evaluating the predictive performance of LASSO, CT, and RF models for categorizing condemnation percentages. In addition, 4 sampling methods - ROS, RUS, BOTH, and ROSE - were used to optimize the imbalance of the original datasets and improve the prediction accuracy. The results can support the incorporation of appropriate ML algorithms into commercial broiler production and help decision-makers establish farm management practices and minimize economic losses. In addition, the methods can be adapted for different broiler producers, increasing the applicability of the results.

MATERIALS AND METHODS

Data Collection and Description of the Variables

The data for this study were sourced from a broiler producer primarily exporting poultry to Europe. In compliance with export regulations, the company ensures antibiotic-free rearing of broilers. The Ross 308 and Cobb 500 commercial broiler strains were raised in controlled commercial environments with regulated temperature, lighting, and ventilation to optimize growth. After the rearing phase, the birds were manually sorted by weight, placed into transport crates, and loaded onto trucks. Each truck, holding 495 crates, transported the birds to the slaughterhouse. Following reversible electrical stunning, the birds were automatically eviscerated, with carcasses and internal organs conveyed together. Inspectors then identified and separated condemned carcasses and organs by removing them from the conveyor. Various causes of condemnation were noted during this inspection, including purulent abscesses, cellulitis, arthritis, viscera unfit for human consumption, cachexia, emaciated carcasses, abnormal carcass color, and carcasses without offal.

The study used datasets of 27,111 truckloads from 2021 to 2022, which contained 14 independent variables (Table 1) with condemnation percentage as the dependent variable. To refine the dataset, records with missing variables, and condemnation rates of less than 0.26% and more than 25.99% were excluded, leaving 23,959 truckloads with condemnation rates between 0.26% and 25.99% for analysis. For the condemnation risk factors analyzed in this study, a condemnation rate of more than 3.0% was set as the threshold for a high condemnation rate to meet customer requirements for corrective action following McDonald's data collection guidance

Table 1. Description of features used to predict the condemnation rate.

Data type	Variable	Description
Category	Season[1]	1 = Winter, 2 = Summer, 3 = Rainy
	Time of transport[2]	1 = Night, 2 = Morning, 3 = Day
	Sex[3]	1 = Male, 2 = Female, 3 = Mixed sex
Numeric	Age (day)	Slaughter age
	Flock size (n)	Number of birds per house
	Mean body weight (g)	Average of body weight per birds per truckload
	Rearing stocking density (kg/m2)	Total amount of kilograms of bird per 1 m^2
	Mortality and culling (%)	(Total number of dead and culled birds x 100)/total number of chicks placed
	Weight per crate (kg)	Total amount of kilograms of bird in the crate
	Birds per crate (n)	Number of birds per crate
	Transport duration (min)	Time from the end of loading to arrival at slaughterhouse
	Distance (km)	Length from farm to slaughterhouse
	Lairage time (min)	Time from entering holding area to leaving holding area
	Feed withdrawal time (min)	Time from feed removal on farm to leaving holding area

[1]Season in Thailand: winter (November−February), summer (March−May), and rainy (June−October).
[2]Time of transport is time of day the vehicle left the farm after loading (Night = 18:00−04:00, Morning = 04:00−08:00, Day = 8:00−18:00).
[3]Sex included male, female, and mixed sex (housing both males and females together).

(McDonald, 2020). Records with condemnation rates higher than 25.99% were excluded due to the presence of infections in those batches, which resulted in abnormally high condemnation rates. The condemnation rate was treated as a binary characteristic and labeled "0″ for a low condemnation rate from 0.26% to 3.00% and "1″ for a high condemnation rate from more than 3.0% to 25.99%. In the refined dataset of 23,959 truckloads, 1,928 (8.05%) were classified as having a high condemnation rate, while 22,031 (91.95%) were classified as having a low condemnation rate.

Pearson correlation coefficients were used to examine the relationships between continuous predictors and the condemnation rate, using the "cor" function from R-Base. A correlation matrix was then created and visualized using the 'corrplot' package, which illustrates the potential linear relationships between the predictors and the condemnation rate and highlights the strength and direction of these relationships.

Machine Learning Algorithms

Prediction models for high condemnation rates (0 = normal/low condemnation, 1 = high condemnation) have been developed using ML algorithms, in particular LASSO, CT, and RF. These algorithms have different prediction strengths. LASSO regularizes the linear regression and introduces a penalty term based on the absolute values of the regression coefficients to simplify the model and automatically select features by shrinking coefficients, making it more interpretable and efficient. The tuning parameter controls the strength of this penalty and plays a crucial role in determining the trade-off between model simplicity and accuracy (Hastie, et al., 2009; Tibshirani, 1996). CT uses iterative partitioning and pruning to create a decision tree that effectively classifies the observations while avoiding overfitting by simplifying the tree structure based on certain criteria such as the classification error rate and the Gini index (Breiman et al., 2017; Shahinfar, et al. 2014). In contrast, RF uses bootstrap aggregation or bagging to create numerous decision trees. Each tree makes independent predictions, and their collective decisions contribute to the overall predictions (Shahinfar, et al. 2014).

The ML models were created with R version 4.3.1 and the algorithms LASSO, CT, and RF (R Core Team, 2023). Important packages used in the development were caret, dplyr, glmnet, tidyverse, partykit, e1071, ROCR, randomForest, and ggplot2.

Model Building and Sampling Techniques

ML algorithms in R were employed to forecast the likelihood of a high condemnation rate. The original dataset contained 22,031 truckloads (91.95%) with a low condemnation rate of 0.26% to 3.0% and 1,928 truckloads (8.05%) with a high condemnation rate of greater than 3.0% to 25.99%. In this study, the minority class accounted for 1% to 20%, which is a moderate imbalance (Google Developer, 2022). ML algorithms perform best when samples are evenly distributed across classes, so sampling methods were used to balance the dataset.

Figure 1 outlines the process of creating the original data model, starting with a 70:30 segmentation of the dataset. The original dataset of 23,959 truckloads was randomly divided into a train dataset and a test dataset, with 16,772 truckloads for the train dataset and 7,187 truckloads for the test dataset. The train dataset comprised 15,422 truckloads with low condemnation rates (majority class) and 1,350 truckloads with high condemnation rates (minority class). This segmentation was used for model building, parameter tuning and performance evaluation.

The model development included a 10-fold cross-validation with the train dataset. In addition, ROS, RUS, BOTH, and ROSE parameter tunning methods from the ROSE package (R Core Team, 2023) were used to remove class imbalances and improve model predictions. In ROS, the instances of the minority class are randomly duplicated to equalize the class distribution (Demir and Şahin, 2022). RUS, on the other hand, randomly removes

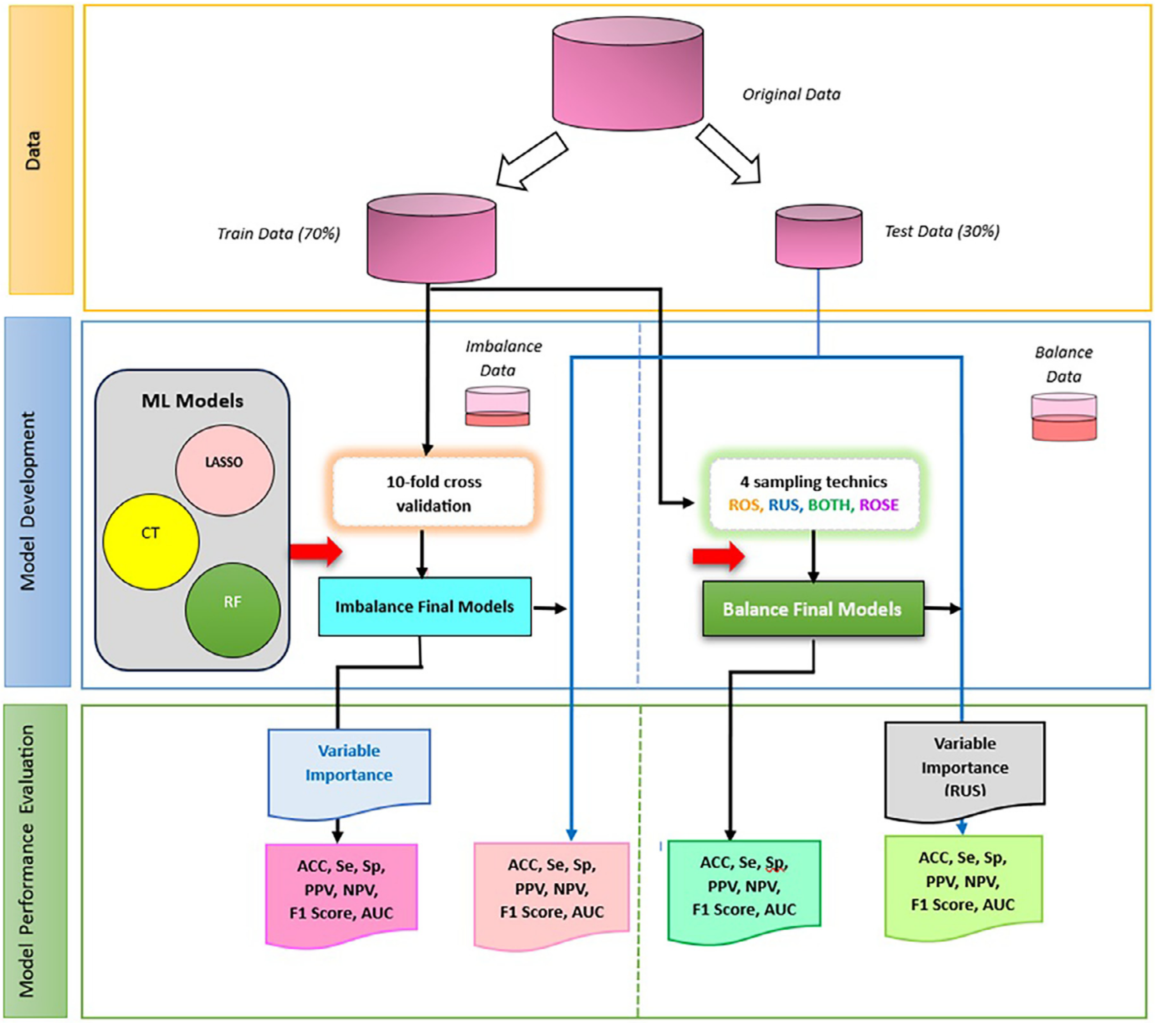

Figure 1. The process of machine learning methods for condemnation rate prediction in original imbalance and balance datasets.

instances from the majority class to balance the class distribution (Demir and Şahin, 2022). BOTH is a combination of oversampling and undersampling (Lunardon, et al., 2014). In synthetic sampling, which is implemented by ROSE in particular, synthetic instances of the minority class are generated (Lunardon, et al., 2014).

Given that the imbalanced data were split into train and test datasets, the test dataset ($n = 7{,}187$ truckloads) was reserved for evaluating all machine learning models generated in this study. The imbalanced train data were initially used to train these models. However, the application of oversampling and undersampling techniques inevitably altered the size of the train dataset compared to the original imbalanced data. The details of the sample size for the train dataset for each sampling technique and ML method are provide in Supplementary Table 3. For example, ROS, an oversampling technique, increase minority class size from 1,350 to 15,422 truckloads.

Model Evaluation

The evaluation of model performance was based on the agreement between the actual results and the predictions of the model, using metrics such as true-positive (**TP**), true-negative (**TN**), false-positive (**FP**), and false-negative (**FN**). The total number of observations (truckloads) was represented by N, where $N = TP + TN + FP + FN$. Standard performance measures defined by (Shahinfar, et al., 2014) were used to evaluate model performance:

- Accuracy (ACC): Measures the proportion of correctly identified observations for both classes (positive = high condemnation, negative = low condemnation).

$ACC = (TP + TN)/(TP + TN + FP + FN)$

- Sensitivity (Se): TP rate or recall, measures the proportion of positives correctly identified.

$Se = TP/(TP + FN)$

- Specificity (SP): TN rate, measures the proportion of negatives correctly identified.

$SP = TN/(TN + FP)$

- Positive Predicted Value (PPV): Precision, measures the proportion of predicted positives that are correct.

$PPV = TP/(TP + FP)$

- Negative Predicted Value (NPV): Measures the proportion of predicted negatives that are correct.

$$\mathrm{NPV} = \mathrm{TN}/(\mathrm{TN} + \mathrm{FN})$$

the F1 score and the area under the receiver operating characteristic (ROC) curve were also used to evaluate model performance (Vihinen, 2012). The F1 score is calculated using the formula:

$$\mathrm{F1} = \frac{2 \text{ x Precision x Recall}}{\text{Presion} + \text{Recall}}$$

the ROC curve for each ML model was generated by plotting the true-positive rates (sensitivity) against the false-positive rates (1-specificity) using the functions provided by the ROCR package (Sing, et al., 2005). The discriminatory or predictive ability of the predictors was assessed by determining the area under the ROC curve (AUC) for each ML model. As stated by Hosmer, et al. (2013), the interpretation of the AUC values was as follows: An area under the curve (**AUC**): 0.5 indicates the absence of discrimination; 0.6 to 0.7 indicates fair discrimination; 0.7 to 0.8 indicates acceptable discrimination; 0.8 to 0.9 indicates great discrimination; and 0.9 to 1.0 denotes outstanding discrimination.

Variable Importance

It is critical in the field of machine learning to assess the significance of variables to obtain information regarding the effect of removing predictors from the model on prediction accuracy. This evaluation was performed by ranking the predictors using the varImp function from the caret package (Punyapornwithaya, et al., 2022). In this analysis, ML models developed with the train dataset were considered. In the context of predictive modeling, the variable importance scores, which were normalized to a maximum of 100, provided a clear indication of the predictors' relative importance, with higher scores signifying greater significance.

RESULTS AND DISCUSSION

Summary of Descriptive Statistics of All Studied Variables

The study used data from 23,959 truckloads from the years 2021 and 2022. Table 2 shows the mean condemnation percentage for each categorical variable, while Table 3 provides descriptive statistics of continuous variables related to condemnation percentages. Additional details on the distinction between high and low condemnation rates for each variable can be found in Supplementary Tables 2 and 3.

Table 2. Mean percentage of condemnation.

Category variable	Truckload n	Truckload %	% Condemnation
Season			
Winter	9,558	39.89	1.41
Summer	4,799	20.03	1.41
Rainy	9,602	40.08	1.12
Time of transport			
Night	10,826	45.18	1.21
Morning	1,950	8.14	1.28
Day	11,183	46.68	1.37
Sex			
Male	1,793	7.48	1.02
Female	1,558	6.50	0.91
Mixed sex	20,608	86.02	1.34

Condemnation percentages for all truckloads ranged from 0.26% to 25.99%, with a mean of 1.29% and a median of 0.75%. The original data set included 22,031 truckloads (91.95%) with a low condemnation rate between 0.26% and 3.0% and 1,928 truckloads (8.05%) with a high condemnation rate between 3.0% and 25.99%. The proportion of minority classes in the dataset was between 1 % and 20 %, indicating a moderate imbalance (Google Developer, 2022). Table 4 shows the percentages of the causes of condemnation. It was found that the highest cause of condemnation was viscera not fit for human consumption (0.96% of 1.29%), representing 74.42% of the causes of condemnation.

In the correlation analysis, almost all pairs of variables showed significant correlations, with p-values less than 0.05, as shown in Figure 2. The results show a strong positive correlation between distance and transportation duration ($r = 0.87$, $P < 0.001$). Similarly, weight per crate and mean body weight are strongly correlated ($r = 0.76$, $P < 0.001$), while mean body weight shows a moderate correlation with rearing stocking density ($r = 0.68$, $P < 0.001$). In addition, the percentage of condemnations shows a slight positive correlation with mortality and culling rate ($r = 0.17$, $P < 0.001$) and negative correlations with mean body weight ($r = -0.20$, $P < 0.001$), weight per crate ($r = -0.17$, $P < 0.001$) and rearing stocking density ($r = -0.14$, $P < 0.001$).

Model Performances With Original Datasets

The model performances of the original datasets are shown in Table 5. The high accuracy in the range of 92-93 % indicates a good overall performance of all 3 models. The high specificity (99%) of all 3 models show how well the models correctly identify negative cases or low condemnation rates. The sensitivity, which indicates the ability to correctly identify all positive cases or a high condemnation rate (Trevethan, 2017), is between 0.06 and 0.17 for the train dataset and 0.04 and 0.13 for the test dataset. In addition, the model performance was also evaluated by the F1 score is a valuable metric in classification models, offering a balanced measure of accuracy by considering both precision and recall. This metric is especially useful in handling imbalanced datasets, providing a comprehensive evaluation of a model's performance (Wood, 2016). In the current study, F1 scores are low: 0.11 to 0.30 for the train dataset and 0.09 to 0.22 for the test dataset. The AUC scores for the LASSO and CT models of train and test datasets were classified as fair (71−72%) and poor (56−57%) classification models, respectively. In contrast, the AUC for the RF model was 99% on the train dataset but dropped to 80% on the test dataset. This finding suggests potential overfitting, as the model performed significantly better on the train data than on the test data. However, since

Table 3. Descriptive statistics of continuous variables and condemnation rate.

Variables	Mean	SD	Min	Max	Median	% CV
Flock size (n)	27411.63	8974.59	5184	62000	30200	32.74
Age (day)	44.74	5.82	32	70	43	13.01
Mean body weight (g)	2935.34	308.60	1330	5040	2933	10.51
Rearing stocking density (kg/m^2)	29.09	3.28	13.27	41.23	29.11	11.24
Mortality and culling (%)	5.09	2.42	1.02	11.96	4.56	47.54
Weight per crate (kg)	18.04	2.15	6.35	35.21	17.96	11.92
Bird per crate (n)	6.15	0.49	4	10	6	7.96
Lairage time (min)	118.45	64.41	10	886	114	54.37
Transport duration (min)	151.23	87.64	30	488	126	57.95
Distance (km)	112.96	78.02	25	354	80	69.07
Feed withdrawal time (min)	535.74	80.41	300	852	550	15.02
Condemnation (%)	1.29	1.82	0.26	25.99	0.75	141.08

Table 4. Causes of condemnation.

Causes of condemnation	% of total slaughter	% of causes of condemnation
Purulent abscess	0.16	12.20
Cellulitis	0.01	1.03
Arthritis	0.04	3.24
Viscera not fit for human consumption	0.96	74.42
Cachexia and emaciated carcasses	0.06	4.55
Abnormal color carcasses	0.04	3.10
Carcass without offal	0.02	1.54
Condemnation percentage	1.29	100

the AUC on the test dataset is still above 80%, the RF model remains a good classifier (Ilyrek, 2023). Previous studies have also documented overfitting in RF models (Li et al., 2024; Song et al., 2021).

These results indicate that the predictive performance of the RF model was outstanding for the train dataset and great for the test dataset, based on the classification criteria by Hosmer, et al., (2013). Analysis of the ROC curve showed that RF had better predictive performance than CT and LASSO, as it had a larger area under the curve (Figure 3). In summary, the results show that none of the 3 models can effectively predict a high condemnation rate in the original imbalanced dataset due to the low sensitivity and F1 score. Although models such as RF showed high predictive power with an AUC of 0.8 for the original test dataset, the overall performance was affected by the imbalance. Therefore, the use of sampling methods to handle the imbalance parameters was deemed necessary to improve the predictive performance of the model.

Model Performances With Balanced Datasets

Handling imbalanced data is essential for enhancing the performance of ML models and ensuring the reliability of the results. If 1 class outperforms the others, biased models may perform well in the dominant class while underperforming in the minority class. By using sampling techniques to treat imbalanced data, ML projects can improve generalization and make predictions across classes, making the model more robust and effective (Nagidi, 2020). In this study, given the moderate imbalance in the original dataset, the 3 ML models were not appropriate for predicting a high condemnation rate due to low Se and F1 scores. Thus, 4 sampling techniques—ROS, RUS, BOTH, and ROSE—were employed to address the imbalanced dataset and improve the model performance. The model performance of the balance train and test datasets with 4 different sampling techniques is presented in Table 6, revealing that each sampling method has a remarkable impact on the sensitivity, with the RUS method having the highest sensitivity compared to the others. The F1 score was also improved by the sampling techniques, as it increased from 0.09-0.22 for the original test datasets to 0.29-0.40 for the balanced test datasets. The AUC values for the original and balanced train datasets differ significantly. All ML models trained on the balanced dataset demonstrated higher AUC values compared to those trained on the original dataset, suggesting a potential risk of overfitting. In contrast, the AUC values for the original and balanced test datasets did not show significant differences. Among the models, RF achieved the highest AUC, followed by LASSO and CT. Avizheh et al. (2023) reported similar AUC values between the original and balanced datasets following parameter tuning. The areas under the ROC curve for the balanced dataset were very similar across the 4 sampling techniques for each model, with the RF model exhibiting the largest area under the ROC curve, as illustrated in Figure 4.

In this study, the most appropriate model for predicting a high condemnation rate should be RF as it had higher AUC compared to CT and LASSO. The AUC of RF was 0.8 which was in the great performance criteria according to Hosmer, et al. (2013). The RUS method was required to treat the imbalance of original data before RF was performed, as it provided higher sensitivity with a comparable F1 score compared to other sampling techniques. However, the RUS sampling technique reduced the ACC value from about 90 % of the original dataset to about 70 % of the balanced dataset. In general, a higher ACC value is preferable. Occasionally, however, a model with a lower ACC value but higher precision or recall may be better (Google Developer, 2022).

Variable Importance

In the original imbalanced datasets, mean body weight (**BW**), weight per crate (**WC**), rearing stocking

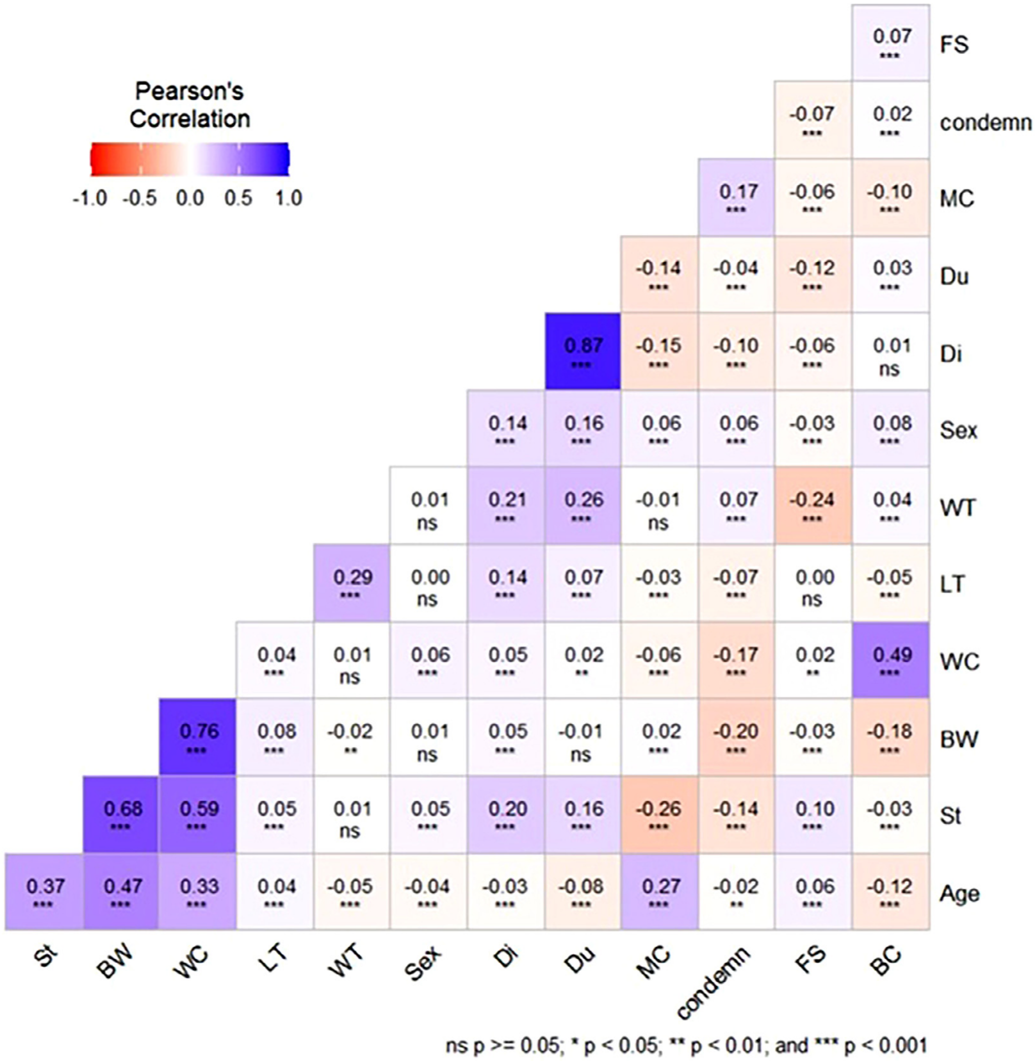

Figure 2. The correlation matrix between the studied variables. Age (bird age), BW (mean body weight), BC (bird/crate), MC (mortality and culling rate), Di (distance of transport), Du (duration of transport), FS (flock size), Sex, WC (weight/crate), LT (lairage time), Se (season), St (rearing stocking density), Ti (time of transport), WT (feed withdrawal time) and condemn (%condemnation).

density (**St**), and age were the most significant predictors of a high condemnation rate for the CT model, and BW, WC, feed withdrawal time (**WT**), and St for the RF models. These were not the same as the 4 most significant predictors of the LASSO model, which were bird per crate (**BC**), sex, WC, and mortality and culling rate (**MC**), as illustrated in Figure 5. Only 1 predictor, WC, was among the top 4 most influential predictors across the 3 models. Although they were not included in the LASSO model, BW and WC were the 2 most significant predictors for the CT and RF models. This inconsistency underlines the varied outcomes of feature selection in the models.

For the balanced dataset, RF with RUS was selected due to its great performance in predicting the high condemnation rate, as shown in Figure 6. The top 4 predictors in each model differed from those in the imbalanced dataset. In the LASSO model for the balanced dataset, the 4 most important predictors were BC, sex, WC, and MC. The 4 most significant predictors in the CT model were BW, MC, WC, and distance of transport (**Di**). In the RF model, BW, WT, MC, and lairage time (**LT**) were the 4 most important predictors for the balanced dataset. MC was the only predictor that was among the top 4 variables in all 3 models. In the LASSO and CT models, WC was 1 of the top 4 predictors, while BW was in the CT and RF models. The first 4 ranks of the CT and RF models did not include any categorical variables,

Table 5. Model performances with the original datasets.

	Parameters[1]						
Models	ACC	Se	Sp	PPV	NPV	F1	AUC
Train data							
LASSO regression	0.92	0.06	0.99	0.74	0.92	0.11	0.72
Classification tree	0.92	0.15	0.99	0.64	0.93	0.24	0.57
Random forests	0.93	0.17	0.99	0.96	0.93	0.30	0.99
Test data							
LASSO regression	0.92	0.04	0.99	0.79	0.92	0.09	0.71
Classification tree	0.92	0.13	0.99	0.63	0.93	0.22	0.56
Random forests	0.93	0.11	0.99	0.78	0.93	0.20	0.80

[1]Accuracy (ACC), sensitivity (Se), specificity (Sp), positive predicted value (PPV), negative predicted value (NPV), F1 (F-measure), AUC (Area under the ROC curve).

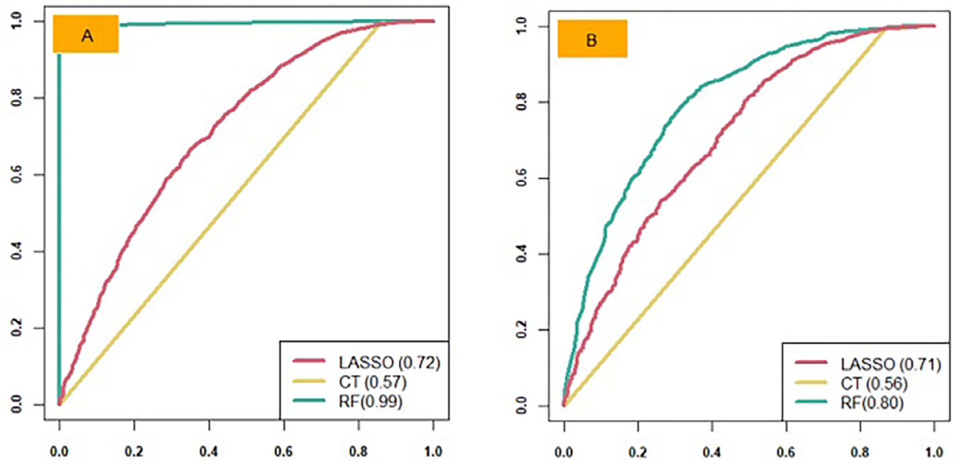

Figure 3. Receiver operator characteristic curve of the LASSO regression (LASSO), classification tree (CT), and random forests (RF) models for the original data sets. A= train dataset, B = test dataset.

but the second rank of the LASSO model only included the categorical variable of sex.

Ranking Variable Importance

Variable importance scores play a crucial role in understanding the importance of individual variables in predictive modeling and classification tasks. These scores provide information on the extent to which a particular variable influences the classification status of observations within a dataset. The most important features can improve model interpretability and guide feature selection. According to Khalilia et al., (2011), the variable important score shows how strongly a variable is correlated with the categorization status. In the current study, the RF algorithm employing the RUS tuning method was the most effective for both classifying and forecasting a high condemnation rate. It was discovered that the 4 most important variables for forecasting a high condemnation rate were BW, WT, MC, and LT.

The first rank of important variable for predicting a high condemnation rate was BW. Pirompud et al. (2023) reported a significant negative relationship between condemnation rate and mean body weight in

Table 6. Model performance of the balanced datasets with 4 sampling techniques.

		Parameters[2]						
Models	Sampling technique[1]	ACC	Se	Sp	PPV	NPV	F1	AUC
Train dataset								
LASSO regression	ROS	0.99	1.00	0.99	0.99	1.00	0.99	1.00
	RUS	0.78	1.00	0.76	0.26	1.00	0.42	0.98
	BOTH	0.98	0.99	0.98	0.79	1.00	0.88	1.00
	ROSE	0.87	0.42	0.91	0.29	0.94	0.34	0.77
Classification tree	ROS	0.99	1.00	0.99	0.99	1.00	0.99	1.00
	RUS	0.79	1.00	0.77	0.27	1.00	0.43	1.00
	BOTH	0.97	0.99	0.97	0.77	0.99	0.87	1.00
	ROSE	0.87	0.40	0.92	0.30	0.94	0.34	0.78
Random forest	ROS	0.99	1.00	0.99	0.99	1.00	0.99	0.99
	RUS	0.78	1.00	0.76	0.26	1.00	0.42	0.99
	BOTH	0.97	0.99	0.98	0.80	0.99	0.88	0.99
	ROSE	0.87	0.44	0.90	0.28	0.95	0.34	0.81
Test dataset								
LASSO regression	ROS	0.92	0.21	0.98	0.57	0.93	0.30	0.72
	RUS	0.73	0.67	0.74	0.18	0.96	0.29	0.72
	BOTH	0.91	0.34	0.96	0.44	0.94	0.39	0.71
	ROSE	0.86	0.38	0.90	0.25	0.94	0.31	0.69
Classification tree	ROS	0.92	0.24	0.98	0.59	0.94	0.34	0.57
	RUS	0.75	0.70	0.75	0.20	0.97	0.31	0.57
	BOTH	0.91	0.35	0.96	0.42	0.94	0.38	0.57
	ROSE	0.87	0.35	0.91	0.25	0.94	0.29	0.58
Random forest	ROS	0.92	0.24	0.98	0.59	0.94	0.34	0.80
	RUS	0.74	0.70	0.75	0.20	0.97	0.31	0.80
	BOTH	0.91	0.36	0.96	0.45	0.94	0.40	0.81
	ROSE	0.86	0.41	0.90	0.27	0.94	0.32	0.79

[1]ROS = random over sampling, RUS = random under sampling, BOTH = both sampling, ROSE = random over sampling example.
[2]Accuracy (ACC), sensitivity (Se), specificity (Sp), positive predicted value (PPV), negative predicted value (NPV), F1 (F-measure), AUC (Area under the curve).

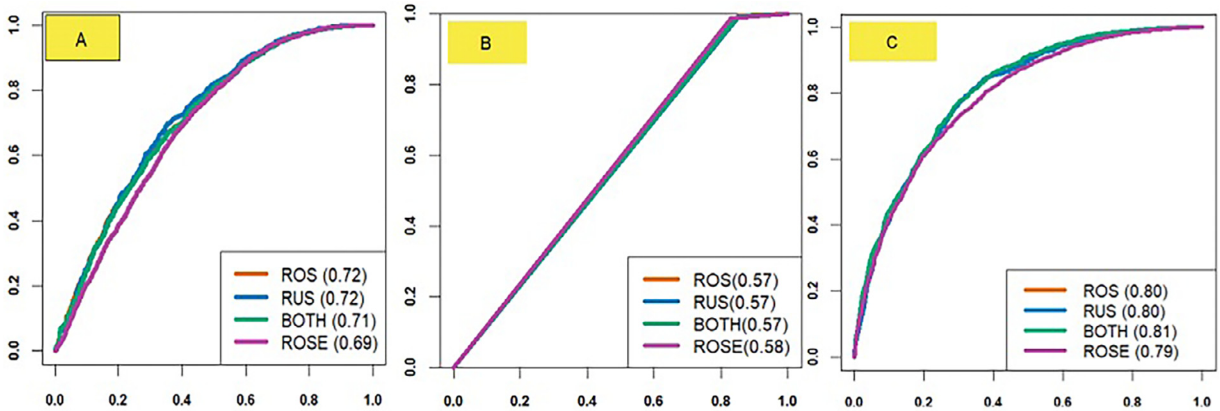

Figure 4. Receiver operator characteristic curve of the LASSO regression (A), classification tree (B), and random forests (C) models of test balance datasets with different sampling techniques.

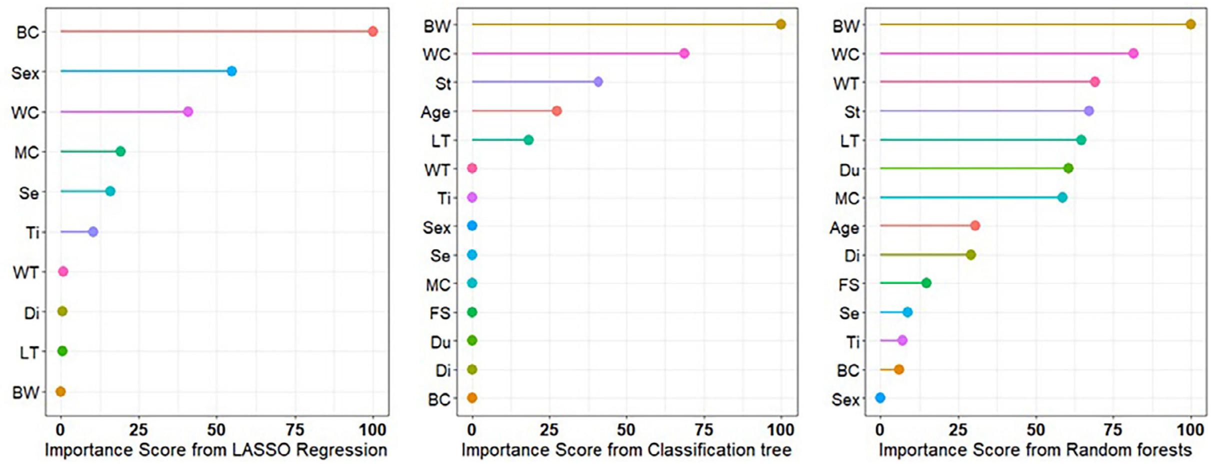

Figure 5. Variable importance plots from the original dataset show the predictors of high condemnation (%) of broilers raised without antibiotics from the LASSO regression, classification tree (CT), and random forests (RF) models. BC (bird/crate), BW (mean body weight), WC (weight/crate), St (rearing stocking density), WT (feed withdrawal time), MC (mortality and culling rate), age, LT (lairage time), Du (duration of transport), Se (season), Ti (time of transport), FS (flock size), and sex.

broilers reared without an antibiotic program. In agreement with Hashimoto et al. (2013) they reported a significant negative correlation between body weight and condemnation rate (r = −0.195). Kanabata et al. (2023) found a positive correlation between high body weight and condemnation, especially in ascites. It is hypothesized that this correlation is the result of excessive oxygen consumption caused by metabolic overload due to weight gain. However, they did not find a significant correlation between body weight and overall total

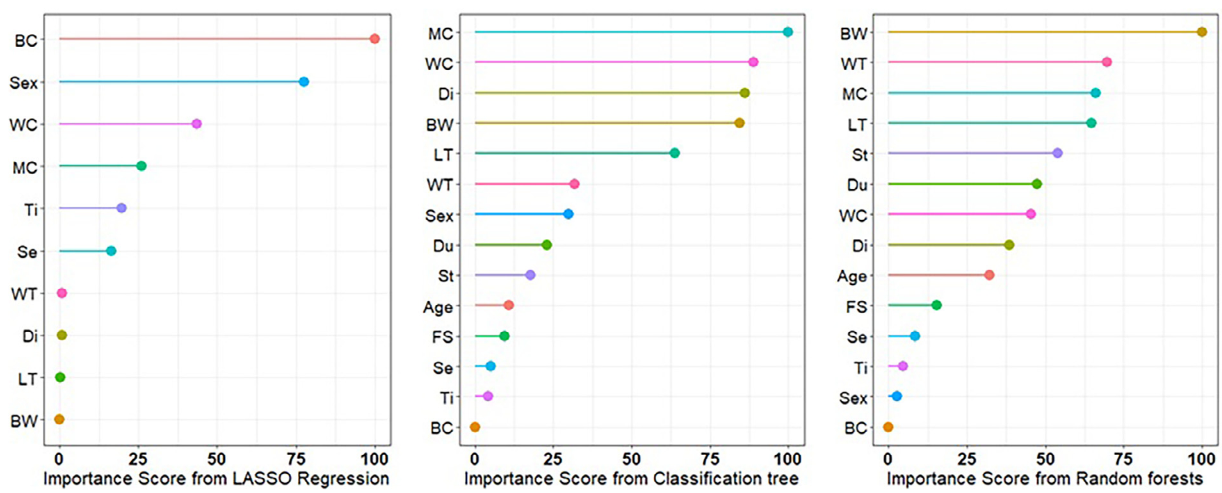

Figure 6. Variable importance plots from the RUS balance dataset show the predictors of high Condemnation (%) of broilers raised without antibiotics from the LASSO regression, classification tree (CT), and random forests (RF) models. BC (bird/crate), BW (mean body weight), WC (weight/crate), St (rearing stocking density), WT (feed withdrawal time), MC (mortality and culling rate), age, LT (lairage time), Du (duration of transport), Se (season), Ti (time of transport), FS (flock size), and sex.

condemnation. The influence of mean body weight and condemnation rate might be varied among studies might be due to broiler genotypes, rearing systems, causes of condemnation, or tools to analyze e.g. conventional statistics or ML algorithms.

Feed withdrawal time ranked as the second most significant variable in predicting a high rate of condemnation. Feed withdrawal time includes the time in the house without feed, the catching time, the transport time, and the time in the lairage area before processing, while water should be provided until catching (Monleon, 2012). Feed withdrawal 8 to 12 h before slaughter is the most effective method to minimize carcass contamination and reduce carcass weight loss (Lyon et al., 1991; Northcutt et al., 1997). The condemnation rate in broilers is influenced by the feed withdrawal time. This could be due to the carcasses being contaminated by the remaining contents of the crop and digestive tract if the withdrawal time is insufficient (Northcutt et al., 1997). The long duration of food deprivation led to the rupture of many inner layers of the mucosa and submucosa, making the intestine sensitive and facilitating contamination with feces. The process of evisceration led to the penetration of feces into the abdominal cavity and carcass, as documented by Lopez, (2010). Excessive feed deprivation for more than 13 h can lead to discoloration of the empty crop, proventriculus, and gizzard due to bile accumulation. The intestines may become weak and thin, with detached mucosa, increasing the likelihood of ruptures and increasing the risk of microbial contamination (Monleon, 2012).

The third crucial variable for predicting a high condemnation rate was MC. Several factors, including rearing density, litter moisture, ammonia levels, ventilation, temperatures, and disinfection, contributed to the clinical signs of reduced feed intake and performance in colibacillosis—a disease caused by E. coli. As a result, there was a high mortality rate during rearing (Chauvin et al., 2011; Whiting et al., 2007). Some surviving broilers showed gastrointestinal pathology that could spread to other organs, leading to respiratory infections and potential septicemia through the colonization of internal organs (Muchon et al., 2019). These broilers were more susceptible to mortality during transportation, which increased the condemnation rate due to viscera unfit for human consumption (Cockram and Dulal, 2018; Lupo et al., 2010).

Lairage time was the fourth significant variable that predicted a high rate of condemnation. Lupo, et al. (2010) found a negative correlation between lairage time and condemnation rate, but a positive correlation with mortality rate. Because they don't have much room in the crate during lairage, heat builds up and has a big impact on the temperature environment. All of the birds in the load face difficulties due to inadequate ventilation in the lairage location. A high death rate may result if the temperature inside the crates is higher than the ambient temperature and the body temperature increases by more than 1°C (Bayliss and Hinton, 1990; Jacobs et al., 2017; Mitchell and Kettlewell, 1998; Nijdam et al., 2004). The observed increase in mortality rate instead of condemnation rate can be explained by the fact that more birds die under unfavorable lairage conditions, which means a lower condemnation rate for the remaining birds. According to Petracci et al. (2005), an effective environmental management system in the processing plant's holding area explains why the lairage period had little effect on the mortality rate or the carcass condemnation rate. This result agrees with that of Pirompud et al. (2023), who found no effect of lairage time on the condemnation rate in broilers reared without antibiotics.

In addition to the 4 main variables, this study identified other significant factors contributing to high condemnation rates in broilers reared without antibiotics. These factors include rearing stocking density, duration of transport, weight per crate, distance of transport, and age of the birds. Rearing stocking density plays a crucial role in the overall welfare of broilers as it affects their health and stress levels during the production cycle. In addition, the duration and distance of transport can affect the welfare of the animals and potentially lead to physiological stress and injury, which in turn can result in condemnation. Weight per crate is another critical factor, as overcrowding or inappropriate loading density can exacerbate stress and physical trauma during transportation and affect the overall condition of the birds on arrival. The age of the birds is also an important factor, as older birds may be more susceptible to stress and health problems during transportation and processing. Addressing these risk factors through proper management practices, optimal transport conditions and adherence to recommended stocking densities can contribute to a significant reduction in condemnation rates and thus minimize economic losses in broiler production.

It is important to note that factors of lesser importance, such as the number of birds per crate, transport time, sex and season, should not be disregarded. While certain variables such as sex and season are unchangeable, factors such as the number of birds per crate and transport time can be adjusted and thus offer potential opportunities for risk mitigation.

Overall, the machine learning model identifies and ranks the factors associated with condemnation rates, providing crucial insights for risk assessment. The poultry industry can leverage these insights to target and mitigate high-risk factors that lead to higher condemnation rates. In addition, by applying new data under different scenarios of factor combinations, the model can predict future condemnation rates. These predictions provide valuable insight into the future and allow the industry to anticipate outcomes and implement more effective strategies.

Limitations

It should be highlighted that although machine learning algorithms and sampling techniques are the focus of this work, it only covers 3 algorithms and 4 sample techniques for parameter optimization. A more comprehensive investigation of the factors influencing increased condemnation rates could significantly improve decision-making for broiler producers. Expanding the comparison to include additional results derived from different algorithms and sampling methods would contribute to a more nuanced and informed decision-making process aimed at reducing condemnation rates. The observed variability in results between different ML algorithms highlights the crucial role played by factors such as the number of predictors, variable types, and datasets. Researchers are advised to carefully consider these factors both in the selection of ML methods and in the interpretation of study results.

In summary, this research not only advances machine learning predictive models but also provides insightful information for broiler production decision-makers. Stakeholders can avoid financial losses by implementing targeted farm management and monitoring measures by accurately predicting high condemnation status. The methodology used in the study provides flexibility in analyzing data from different broiler producers, which increases its relevance in commercial contexts.

ACKNOWLEDGMENTS

We sincerely thank Sungroup Company, Thailand, for generously providing the data required for our analysis and for their unwavering support in this endeavor. In addition, our sincere appreciation goes to the Faculty of Industrial Education and Technology at King Mongkut Institution of Technology Ladkrabang for granting us access to the research facilities, which greatly contributed to the success of our research.

DISCLOSURES

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome. As corresponding author, I confirm that the manuscript has been read and approved for submission by all named authors.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.psj.2024.104270.

REFERENCES

Ahmad, M. W., J. Reynolds, and Y. Rezgui. 2018. Predictive modeling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees, and regression trees. J. Clean. Prod. 203:810–821.

Avizheh, M., M. Dadpas and, E. Dehnavi, and H. Keshavarzi. 2023. Application of machine-learning algorithms to predict calving difficulty in Holstein dairy cattle. Anim. Prod. Sci. 63:1095–1104.

Bayliss, P., and M. Hinton. 1990. Transportation of broilers with special reference to mortality rates. Appl. Anim. Behav. Sci. 28:93–118.

Blockeel, H., L. Devos, B. Frénay, G. Nanfack, and S. Nijssen. 2023. Decision trees: From efficient prediction to responsible AI. *Front.* Artif. Intell. 6:1124553.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 2017. Classification and regression trees. Chapman & Hall/CRC., Routledge, Boca Raton.

Buzdugan, S. N., Y. M. Chang, B. Huntington, J. Rushton, J. Guitian, P. Alarcon, and D. P. Blake. 2020. Identification of production chain risk factors for slaughterhouse condemnation of broiler chickens. Prev. Vet. Med. 181:105036.

Chauvin, C., S. Hillion, L. Balaine, V. Michel, J. Peraste, I. Petetin, C. Lupo, and S. L.e Bouquin. 2011. Factors associated with mortality of broilers during transport to slaughterhouse. Animal 5:287–293.

Cockram, M. S., and K. J. Dulal. 2018. Injury and mortality in broilers during handling and transport to slaughter. Can. J. Anim. Sci. 98:416–432.

Demir, S., and E. K. Şahin. 2022. Evaluation of oversampling methods (OVER, SMOTE, and ROSE) in classifying soil liquefaction dataset based on SVM, RF, and Naïve Bayes. Avrupa Bilim ve Teknoloji Dergisi 34:142–147.

Düpjan, S., and M. S. Dawkins. 2022. Animal welfare and resistance to disease: interaction of affective states and the immune system. Front. Vet. Sci. 9:929805.

Ghosh, P., S. Azam, M. Jonkman, A. Karim, F. J. M. Shamrat, E. Ignatious, S. Shultana, A. R. Beeravolu, and F. De Boer. 2021. Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques. IEEE Access 9:19304–19326.

Google Developer. 2022. Imbalanced data. Accessed Feb. 2023. https://developers.google.com/machine-learning/data-prep/construct/sampling-splitting/imbalanced-data

Hashimoto, S., K. Yamazaki, T. Obi, and K. Takase. 2013. Relationship between severity of footpad dermatitis and carcass performance in broiler chickens. J. Vet. Med. Sci. 75:1547–1549.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. Random forests. The elements of statistical learning: Data mining, inference and prediction. 2nd edition Springer, Newyork, NY.

Hortêncio, M. C., L. R. M. Costa, M. V. P. de Souza, W. D. de Freitas, B. B. Fonseca, M. J. B. Silva, and M. V. C. Cossi. 2022. Time series evaluation of condemnation at poultry slaughterhouses enable to export in Southeastern Brazil (2009−2019): a tool for optimizing resources in the poultry production chain. BMC Vet. Res. 18:1–10.

Hosmer, D. W. Jr, S. Lemeshow, and R. X. Sturdivant. 2013. Applied logistic regression. (p. 398). John Wiley & Sons, New York, NY, 398.

Ilyrek, K. 2023. Area under the curve (AUC): A measure of performance ROC curve and AUC: Evaluating model performance. Accessed Feb. 2023. https://medium.com/@ilyurek/roc-curve-and-auc-evaluating-model-performance-c2178008b02

Jacobs, L., E. Delezie, L. Duchateau, K. Goethals, and F. A. Tuyttens. 2017. Broiler chickens dead on arrival: Associated risk factors and welfare indicators. Poult. Sci. 96:259–265.

Junghans, A., L. Deseniß, and H. Louton. 2022. Data evaluation of broiler chicken rearing and slaughter—an exploratory study. Front. Vet. Sci. 9:957786.

Kanabata, B., F. Souza, G. Biz, R. Pescim, and A. L. Soares. 2023. Relationship between pre-slaughter factors and major causes of carcass condemnation in a broiler slaughterhouse under federal inspection. Braz. J. Poult. Sci. 25:eRBCA-2022.

Khalilia, M., S. Chakraborty, and M. Popescu. 2011. Predicting disease risks from highly imbalanced data using random forest. BMC Med. Inform. Decis. Mak. 11:1–13.

Kuhle, S., B. Maguire, H. Zhang, D. Hamilton, A. C. Allen, K. Joseph, and V. M. Allen. 2018. Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: A retrospective cohort study. BMC Pregnan. Childb. 18:1–9.

Li, H.-b., Y.-j. Du, G. R. Kenmegne, and C.-w. Kang. 2024. Machine learning analysis of serum cholesterol's impact on knee osteoarthritis progression. Sci. Rep. 14:18852.

Lopez, E. C. 2010. Optimal feed withdrawal minimises losses at processing. Accessed Oct. 2022. https://www.poultryworld.net/home/optimal-feed-withdrawal-minimises-losses-at-processing/.

Lunardon, N., G. Menardi, and N. Torelli. 2014. ROSE: a package for binary imbalanced learning. R J 6:79–89.

Lupo, C., S. L.e Bouquin, L. Balaine, V. Michel, J. Peraste, I. Petetin, P. Colin, and C. Chauvin. 2009. Feasibility of screening broiler chicken flocks for risk markers as an aid for meat inspection. Epidemiol. Infect. 137:1086–1098.

Lupo, C., S. Bougeard, L. Balaine, V. Michel, I. Petetin, P. Colin, S. LeBouquin, and C. Chauvin. 2010. Risk factors for sanitary condemnation in broiler chickens and their relative impact: application of an original multiblock approach. Epidemiol. Infect. 138:364–375.

Lyon, C., C. Papa, and R. Wilson Jr. 1991. Effect of feed withdrawal on yields, muscle pH, and texture of broiler breast meat. Poult. Sci. 70:1020–1025.

McDonald. 2020. McDonald's data collection guidance. KWI template. Version 01. pp.1-25.

Magalhães, E. S., D. Zhang, C. Wang, P. Thomas, C. A. Moura, G. Trevisan, D. J. Holtkamp, C. Rademacher, G. S. Silva, and D. C. Linhares. 2023. Comparing forecasting models for predicting nursery mortality under field conditions using regression and machine learning algorithms. Smart Agri. Technol. 5:100280.

Mitchell, M., and P. Kettlewell. 1998. Physiological stress and welfare of broiler chickens in transit: Solutions not problems. Poult. Sci. 77:1803–1814.

Mlambo, F., C. Chironda, and J. George. 2022. Risk stratification of COVID-19 using routine laboratory tests: A machine learning approach. Infect. Dis. Rep. 14:900–931.

Monleon, R. 2012. Pre-processing handling in broilers. Accessed Jan. 2024. https://en.aviagen.com/assets/Tech_Center/Ross_Tech_Articles/RossTechNotePreProcessHandlingOct2012.pdf

Muchon, J. L., R. G. Garcia, É. R.d. S. Gandra, A. S.d. A. Assunção, C. M. Komiyama, F. R. Caldara, I. A. Nääs, and R. A.d. Santos. 2019. Origin of broiler carcass condemnations. Rev. Bras. Zootec. 48:e20180249.

Nagidi, J. 2020. Best ways to handle imbalanced data in machine learning. Accessed Feb. 2023. https://dataaspirant.com/handle-imbalanced-data-machine-learning/

Nijdam, E., P. Arens, E. Lambooij, E. Decuypere, and J. Stegeman. 2004. Factors influencing bruises and mortality of broilers during catching, transport, and lairage. Poult. Sci. 83:1610–1615.

Northcutt, J., S. Savage, and L. Vest. 1997. Relationship between feed withdrawal and viscera condition of broilers. Poult. Sci. 76:410–414.

Petracci, M., M. Bianchi, and C. Cavani. 2005. Preslaughter factors affecting mortality, liveweight loss, and carcass quality in broiler chickens. Proc. XVII Eur. Symp. Qual. Poul. Meat (pp. 251−255) 251–255.

Pirompud, P., P. Sivapirunthep, V. Punyapornwithaya, and C. Chaosap. 2023. Pre-slaughter handling factors affecting dead on arrival, condemnations, and bruising in broiler chickens raised without an antibiotic program. Poult. Sci. 102:102828.

Pirompud, P., P. Sivapirunthep, V. Punyapornwithaya, and C. Chaosap. 2024. Application of machine learning algorithms to predict dead on arrival of broiler chickens raised without antibiotic program. Poult. Sci. 103:103504.

Punyapornwithaya, V., K. Klaharn, O. Arjkumpa, and C. Sansamur. 2022. Exploring the predictive capability of machine learning models in identifying foot and mouth disease outbreak occurrences in cattle farms in an endemic setting of Thailand. Prev. Vet. Med. 207:105706.

R Core Team, R. 2023. R: A Language and environment for statistical computingR foundation for statistical computing, Vienna, Austria. Accessed Feb. 2023. https://www.R-project.org

Sampson, D. L., T. J. Parker, Z. Upton, and C. P. Hurst. 2011. A comparison of methods for classifying clinical samples based on proteomics data: a case study for statistical and machine learning approaches. PloS One 6:e24973.

Shahinfar, S., D. Page, J. Guenther, V. Cabrera, P. Fricke, and K. Weigel. 2014. Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. J. Dairy Sci. 97:731–742.

Sing, T., O. Sander, N. Beerenwinkel, and T. Lengauer. 2005. ROCR: Visualizing classifier performance in R. Bioinform 21:3940–3941.

Song, J., Y. Gao, P. Yin, Y. Li, Y. Li, J. Zhang, Q. Su, X. Fu, and H. Pi. 2021. The random forest model has the best accuracy among the four pressure ulcer prediction models using machine learning algorithms. Risk Manag. Healthc. Policy 14:1175–1187.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. J. Royal Statistical Soc. Series B: Stat. Methodol. 58:267–288.

Trevethan, R. 2017. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. Front. Public Health 5:307.

Vihinen, M. 2012. How to evaluate performance of prediction methods? Measures and their interpretation in variation effect analysis. Proc. BMC Genom. 13:1–10.

Whiting, T. L., M. E. Drain, and D. P. Rasali. 2007. Warm weather transport of broiler chickens in Manitoba. II. Truck management factors associated with death loss in transit to slaughter. Can. Vet. J. 48:148.

Wood, T. 2016. What is the F-score? Accessed Feb. 2023. http://deepai.org/machine-learning-glossary-and-terms/f-score.