

Article

High-Accuracy Chicken Breed Identification Using Microsatellite Genotype Data and AutoGluon Framework

Rajaonarison Faniriharisoa Maxime Toky ^{1,†}, Sutthisak Sukhamsri ^{2,*}, Sadeep Medhasi ^{1,†} , Trifan Budi ¹, Thitipong Panthum ¹, Worapong Singchat ^{1,3} and Kornorn Srikulnath ^{1,3,*} 

¹ Animal Genomics and Bioresource Research Unit (AGB Research Unit), Faculty of Science, Kasetsart University, Bangkok 10900, Thailand; rajaonarison.t@ku.th (R.F.M.T.); sadeep.m@chula.ac.th (S.M.); trifan.bu@ku.th (T.B.); thitipong.pant@ku.ac.th (T.P.); worapong.singc@ku.ac.th (W.S.)

² Department of Information Technology, Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna Tak, Tak 63000, Thailand

³ Biodiversity Center, Kasetsart University (BDCKU), Bangkok 10900, Thailand

* Correspondence: sutthisak@rmutl.ac.th (S.S.); kornorn.s@ku.ac.th (K.S.)

† These authors contributed equally to this work.

Simple Summary

Identifying chicken breeds correctly is important for conserving local breeds and improving breeding programs. However, many breeds look very similar, making visual identification difficult and sometimes inaccurate. In this study, genetic information from several chicken populations was used to train a machine learning model to recognize breed patterns. The model was tested and showed high accuracy in identifying most breeds. This demonstrates that computer-based methods can offer a practical and reliable tool for farmers, breeders, and conservation groups. As more genetic data becomes available, this approach is expected to become even more accurate and useful for protecting and managing valuable chicken breeds.

Abstract

The practical applications of breed identification are numerous and diverse, and they include breed conservation and breeding program design. However, distinguishing between breeds remains challenging and costly, especially for phenotypically similar chicken populations. Continued research is necessary to develop more accessible and optimized methodologies. To address these challenges, machine learning (ML) offers promising tools for analyzing complex genetic data. The capabilities of machine learning, especially the random forest (RF) model, to enhance various fields, including bioinformatics, have recently been demonstrated. In this study, microsatellite genotype data from 651 individuals across 30 chicken populations filtered from a larger initial dataset for consistency were used to classify breeds using an RF model. Cross-validation techniques, including 10-fold cross-validation and leave-one-out cross-validation, were employed to assess the performance of the model. The model performance was evaluated using metrics such as accuracy, Cohen's Kappa, 95% confidence interval, and F1-score. Results showed that the RF model achieved a 95.38% accuracy on the testing dataset. Accuracies of 91.44% and 90.99% were observed for 10-fold cross-validation and leave-one-out cross-validation, respectively. It is believed that larger datasets will significantly improve outcomes for other breeds. Because of its generalizability, the trained model can serve as a straightforward and modern method for chicken breed determination using machine learning. This study demonstrates that ML, particularly automated approaches like AutoGluon, provides a robust and accessible framework for chicken breed identification using cost-effective microsatellite data.



Academic Editor: Wojciech Makalowski

Received: 24 November 2025

Revised: 17 December 2025

Accepted: 19 December 2025

Published: 22 December 2025

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: breed determination; chicken breeds; machine learning; microsatellite; random forest

1. Introduction

Chicken breeds represent important global genetic resources for food security, cultural heritage, and environmental resilience, yet conserving their genetic diversity remains a major challenge worldwide [1,2]. Breed identification involves classifying an animal into a group characterized by a homogeneous phenotype. Accurate classification is vital in animal breeding, including livestock, poultry, and aquaculture, as it forms the basis for maintaining specific traits and performing essential operations such as selective breeding and resource management [1–3]. Various approaches, including morphological identification and molecular markers, have been developed for breed identification or confirmation [4–7]. Molecular marker identification involves methods such as Amplified Fragment Length Polymorphism (AFLP), microsatellite genotyping, Single nucleotide polymorphism (SNP) panels with microarray, and whole-genome sequencing. However, most of these approaches are costly, time-consuming, require advanced genomic expertise, and involve labor-intensive processes. For instance, SNP panels with microarrays and whole-genome sequencing are highly informative for breed identification owing to their examination of many genetic loci. However, their high cost makes them impractical for use in laboratories or communities in developing countries (Global South). Additionally, farmers and breeders, who benefit most from breed identification, often lack the necessary experience, making expense reduction a critical concern [8,9]. Legacy methods for breed identification remain difficult to access and costly for many users. Some uncertainties regarding the origins of certain chicken breeds, particularly when phenotypic similarity is involved, have recently been clarified [10]. Improved methods are needed to facilitate this process.

Microsatellite DNA, or short tandem repeats, are short, repetitive sequences of 1–6 base pairs that are tandemly repeated in the genomes of both prokaryotes and eukaryotes [11]. Owing to their co-dominant inheritance, high allelic diversity per locus—referred to as high multiplex ratio—and the ease and cost-effectiveness of amplification by PCR, microsatellites are widely used as genetic markers across all continents [12–14]. Prior research has indicated that these tandem repeats are closely associated with the genetic structure of populations [15]. The strong correlation between microsatellite markers and breed specificity makes them valuable tools for various genetic studies, including breed differentiation in animals, particularly chickens [16–19]. Genotypic data for microsatellite markers, which represent allele information at specific loci, are obtained by extracting DNA, amplifying microsatellite regions via PCR, and determining allelic sizes [11]. The genetic variation between individuals is organized into a genotypic data table for further analysis. When the genotypic data are large and sufficient, representing multiple breeds and localities, the accuracy of breed identification is strengthened. The Siamese chicken bioresource project, as described by Wattanadilokchatkun et al. [20], provides microsatellite genotype data across many chicken populations and breeds, serving as a valuable resource for breed identification. However, determining breeds using genotypic data involves clustering analysis and probability, which sometimes results in chickens from the same breed not being completely grouped but overlapping with other breeds. This may be due to high genetic variation within breeds. Developing methods to justify and confirm breed identities is necessary to improve reliability and practicality.

Over the past decade, advances in artificial intelligence and increased computational power have made machine learning (ML) widely accessible and effective for analyzing

complex biological data [7]. The choice of ML algorithms depends on the study objectives and data structure, with supervised learning being particularly suitable when labeled data are available [21–23]. In this context, Random Forest (RF) is well-suited for genotypic data because it can handle high-dimensional, categorical predictors and capture complex interactions among loci without requiring assumptions about data distribution. These properties make RF a robust approach for genotype-based classification tasks. Two main categories of supervised models are distinguished: regression models and classification models [24]. The most notable supervised learning algorithms include decision trees, K-nearest neighbors (KNN), support vector machines (SVM), and random forests (RF). In bioinformatics, various ML algorithms have been adopted to address complex classification tasks involving high-dimensional genetic data. Commonly used models include decision trees, SVM, KNN, ensemble methods such as RF, and gradient boosting frameworks like LightGBM and XGBoost. Recently, autoML platforms such as AutoGluon have emerged, offering streamlined approaches that combine multiple models and optimize hyperparameters [25,26]. Of these methods, the RF model has become a popular choice owing to its efficiency and high predictive accuracy across various data types, including those with large attribute spaces and complex structures [27,28]. The RF model, which was initially proposed by Breiman (2000, 2001, 2004) [29–32], can be applied to both regression and classification tasks, and is effective in a variety of practical applications [33–36]. The RF algorithm, which operates by constructing multiple decision trees, trains each tree independently to classify data instances, and the final classification is determined by majority vote across all trees [37]. This ensemble approach, which is used to enhance predictive accuracy and generalizability, is considered superior to single models or the decision tree [38,39].

Although microsatellite genotyping has been widely used in population genetics and breed characterization in various animal species, its application in ML-based classification of chicken breeds remains underexplored. Burócziová and Říha (2009) [40] provided an overview of the use of classification algorithms with microsatellite genotype data to identify genetic differences between horse breeds. However, few studies have systematically explored whether similar approaches can be applied to chicken breeds, particularly using ensemble models like RF. It is expected that using these genotype data to train RF models will achieve high performance in chicken breed prediction. Genotype data from multiple chicken populations were compiled and used to train RF-based classification models. These models were assessed with robust cross-validation techniques to ensure reliability, supporting the development of a more accessible, scalable, and accurate breed identification system with applications in conservation and breeding.

Therefore, this study aims to evaluate the performance of a Random Forest model and the AutoGluon—AutoML framework for classifying chicken breeds using microsatellite genotype data, providing a comparative analysis of their accuracy and practicality.

2. Materials and Methods

2.1. Microsatellite Marker Genotype Dataset

Data on chicken genotypes used in this study were obtained from <https://doi.org/10.5061/dryad.hhmgqnkm0> (accessed on 6 June 2024) [20] and “[Dataset] Supplementary Information: High-accuracy chicken breed identification using microsatellite genotype data and AutoGluon framework by Toky et al.” from the Kasetsart University Knowledge Repository (https://kukr.lib.ku.ac.th/kukr_es/dataset, accessed on 22 December 2025), comprising 651 individuals across 30 populations (11 indigenous chicken breeds, 3 local chicken breeds, and 16 populations of red junglefowl), which were also under the Siam Chicken Bioresource Project (<https://www.sci.ku.ac.th/scbp/>; accessed on 30 June 2024) (Table S1) [10,20,41,42]. The genotype matrix was composed of 28 bi-allelic microsatellite

loci. Twenty-eight microsatellite primer sets were selected based on the recommendations of the Food and Agriculture Organization for chicken biodiversity assessments [43]. To ensure robust model performance, populations with fewer than 30 individuals were excluded from the analysis; however, the number of training samples was sufficient, even after data splitting for testing purposes. The data was initially formatted using Microsoft Excel, after which the genotype data were imported into a Pandas DataFrame using Pandas version 2.2.2, and a manual class weighting strategy was implemented to address class imbalance during model training [44]. This was achieved by using the `compute_class_weight` function from the `sklearn.utils.class_weight` module, which ensured appropriate weighting of classes with fewer samples to mitigate potential bias during training. A comprehensive list of the populations included in the final analysis is provided in Table 1. The proportion of data reserved for testing was optimized through experimentation with various splits, ranging from 10% to 30%. The optimal testing split was identified as 15%, which produced the best-performing model. Feature scaling was also applied to both training and testing datasets to minimize inaccuracies in predictions. This scaling, which prevented larger magnitude datasets from disproportionately influencing the results, was achieved using `StandardScaler` from the `scikit-learn` library version 1.5.1 [45,46]. The RF model was used as the baseline, while the AutoGluon framework, which incorporated hyperparameter tuning across various models, including ensemble methods, tree-based algorithms, linear models, neural networks, and KNN, was applied to enhance predictive performance for chicken breed classification.

Table 1. Distribution of chicken population genotype data after exclusion of less than 30 individuals, and random split from the dataset division into 85% training and 15% testing data.

Population	Aberrations	Number of Individuals	Training Dataset	Testing Dataset
Betong	Bt	30	26	4
Chaiyaphum (<i>G. g. spadiceus</i>)	Chaiya Ggs	30	24	6
Chaiyaphum (<i>G. g. gallus</i>)	Chatha Ggg	30	25	5
Fighting chicken	fight	30	26	4
Huai Yang Pan (<i>G. g. spadiceus</i>)	HYP Ggs	30	26	4
Khao Kho (<i>G. g. spadiceus</i>)	KK Ggs	30	23	7
Khok Mai Rua (<i>G. g. gallus</i>)	KMR Ggg	30	25	5
Mae Hong Son	MHS	70	59	11
Petchaburi (<i>G. g. spadiceus</i>)	Petch Ggs	30	29	1
Roi Et (<i>G. g. gullus</i>)	RE Ggg	30	27	3
Sa Kaeo (<i>G. g. gullus</i>)	SK Ggg	30	24	6
Si Sa Ket (<i>G. g. gullus</i>)	SSK Ggg	30	25	5
Uthai Thani (Samae Dam)	UT	33	29	4
Total		433	368	65

2.2. Baseline Model: RF Model

Training the model: The models were developed and evaluated using Python 3.12.4 programming language, which was installed via the Conda 24.5.0 package on the SUSE Linux Enterprise Server 15 SP4 operating system. All the requisite Python modules were incorporated into the environment to facilitate this analysis [47,48]. Both linear and non-linear feedforward neural networks were tested for multiclass classification tasks using PyTorch version 2.3.1.post100 [49,50]. The neural network architectures were constructed using the `nn` submodule from the primary PyTorch package, `torch`. In the case of non-linear models, the hidden layers were configured to use the rectified linear unit activation function, thereby introducing non-linearity. By contrast, the linear models used solely the

nn.Linear class. A variety of combinations of hyperparameters, optimizers, loss functions, and training epochs were investigated.

Preliminary Model Comparisons: The neural network models, however, attained a maximum accuracy of 73% on this dataset. Random Forest (RF) was selected as a baseline classification model because the dataset comprised multiple categorical and mixed-type predictors rather than a single continuous variable. RF effectively captures complex interactions among categorical variables through an ensemble of decision trees trained on different subsets of samples and features [31]. The RF model was implemented using the RandomForestClassifier class from the scikit-learn.ensemble module of scikit-learn version 1.5.1, and an accuracy of 75.51% was achieved with the use of only 15 decision trees, indicating slightly superior performance [51,52]. The performance of the RF model was further optimized by conducting experiments with varying numbers of decision trees. Since the RF algorithm can handle categorical data, the two allelic values at each bi-allelic locus were concatenated using an underscore as a separator [53,54]. This approach ensured that the model interpreted the data as distinct categorical combinations rather than as a single continuous value, while robust performance was guaranteed using a previously described class weighting [55]. After comparisons against Gini, the criterion for determining the optimal splits at each node was set to entropy, calculated using Equation (1):

$$Entropy(S) = -\sum_{i=1}^C p_i \log_2(p_i) \quad (1)$$

where p_i represents the proportion of class i in the subset, and C is the number of classes. To ensure training reproducibility and consistency, the random_state was fixed at the arbitrary value of 42, which guaranteed the reproducibility of random processes, including data and feature subset selection for each decision tree, as described by Liaw and Wiener (2002) [53] and Breiman (2001) [31].

The performance and behavior of ML models are significantly influenced by the hyperparameters used in their configuration [56]. Proper tuning of these parameters is crucial, as it can significantly enhance the generalizability of the model to unseen data. To optimize the hyperparameters of the model, a tuning procedure was conducted using the RandomizedSearchCV function from the sklearn.model_selection module, which involved exploring a range of hyperparameter configurations through a randomized search of five distinct random combinations of parameters. The number of decision trees in RF was varied from 10 to 1000, with the maximum depth of each tree set to “None,” allowing for unrestricted growth. The search was executed over 50 iterations, during which a three-fold cross-validation strategy was used to assess the performance of each hyperparameter combination. This process was undertaken to identify the optimal hyperparameters that maximize the accuracy of the model.

Model evaluation using cross-validation techniques: The model performance was evaluated using two cross-validation methods, repeated 10-fold and leave-one-out, which provided a comprehensive understanding of its robustness and reliability [47,57]. In repeated 10-fold cross-validation (R10FCV), the dataset was divided into ten equal partitions, or folds. The model was trained on nine of the folds and tested on the remaining fold. This process was repeated ten times, with each fold serving as a test set once. The performance metrics were subsequently averaged across the ten iterations to obtain the total estimate of the model performance. With the leave-one-out cross-validation (LOOCV), each sample in the dataset was used as a test set once, with the remaining $n-1$ samples used for training. This approach results in n iterations, where n is the total number of samples. This ensures that each sample is tested individually while the model is trained on all other samples. The model is trained and evaluated on each data point of the dataset. This approach enables a comprehensive evaluation of the model, although high computational costs are

incurred for large datasets. R10FCV was implemented using the StratifiedKFold class from the model_selection module, with n_splits set to 10 ($k = 10$). This ensured that each fold was representative of the overall class distribution. LOOCV was performed using the LeaveOneOut class, which provided a straightforward method for performing this type of cross-validation. These cross-validation techniques provided insights into the generalizability of the model to unseen data.

Model testing and performance assessments: In addition to cross-validation analyses, the predictive performance of the model was evaluated on manually defined test splits (85% training and 15% testing) using the “predict” method of the trained classifier instance, which generated predictions based on the test data. The precision of these predictions was evaluated using the torch.eq function, which performed an element-wise comparison between the predicted values and true labels, resulting in a tensor that indicated where the predictions matched the true labels (Figure S1). The accuracy of the model was then calculated based on the proportion of correct predictions relative to the total number of predictions as follows:

$$Accuracy = \left(\frac{\text{Correct predictions}}{\text{Total number of predictions}} \right) \times 100 \quad (2)$$

In evaluating classification models, the term “agreement” describes the degree to which two outputs produce the same classification. The Cohen’s Kappa metric was employed to assess the agreement between predicted and true classifications, accounting for the chance agreement that may occur [58]. Kappa’s value is calculated as follows:

$$K = \frac{P(A) - P(E)}{1 - P(E)} \quad (3)$$

where $P(A)$ is the observed agreement and $P(E)$ is the expected agreement by chance. It provides a measure of the improvement in model performance relative to what might be expected by chance alone, thus offering a more comprehensive understanding of model performance [59].

For each class, precision is defined as the ratio of true positives (TP) to the sum of TP and false positives (FP), while recall is the ratio of TP to the sum of TP and false negatives (FN) [60,61]. TP are correct classifications, FP are incorrect classifications to other classes, and FN are instances belonging to other classes but predicted to be among the current considered class. Precision and recall are calculated using Equations (4) and (5), respectively:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

The harmonic means of precision and recall, also known as the F1-score, is calculated using the following equation:

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

The F1-score serves as a critical metric for assessing model performance, particularly in scenarios involving class imbalances [62]. As the F1-score provides a comprehensive understanding of TP, FP, and FN, it is essential in classification projects where false positives and false negatives can have serious consequences [63,64].

2.3. Optimized Model: AutoGluon Approach

The methodological workflow began with data preprocessing to ensure data quality and balance. Missing values within the dataset were addressed using the MISSINGNO technique, which provides a systematic visualization and imputation to handle incomplete entries. To address class imbalance, particularly prevalent in categorical classification tasks, the synthetic minority oversampling technique (SMOTE) method was applied to generate synthetic samples for minority classes, thus enhancing the representativeness of the training data with an AutoGluon approach in the final step (Figure 1).

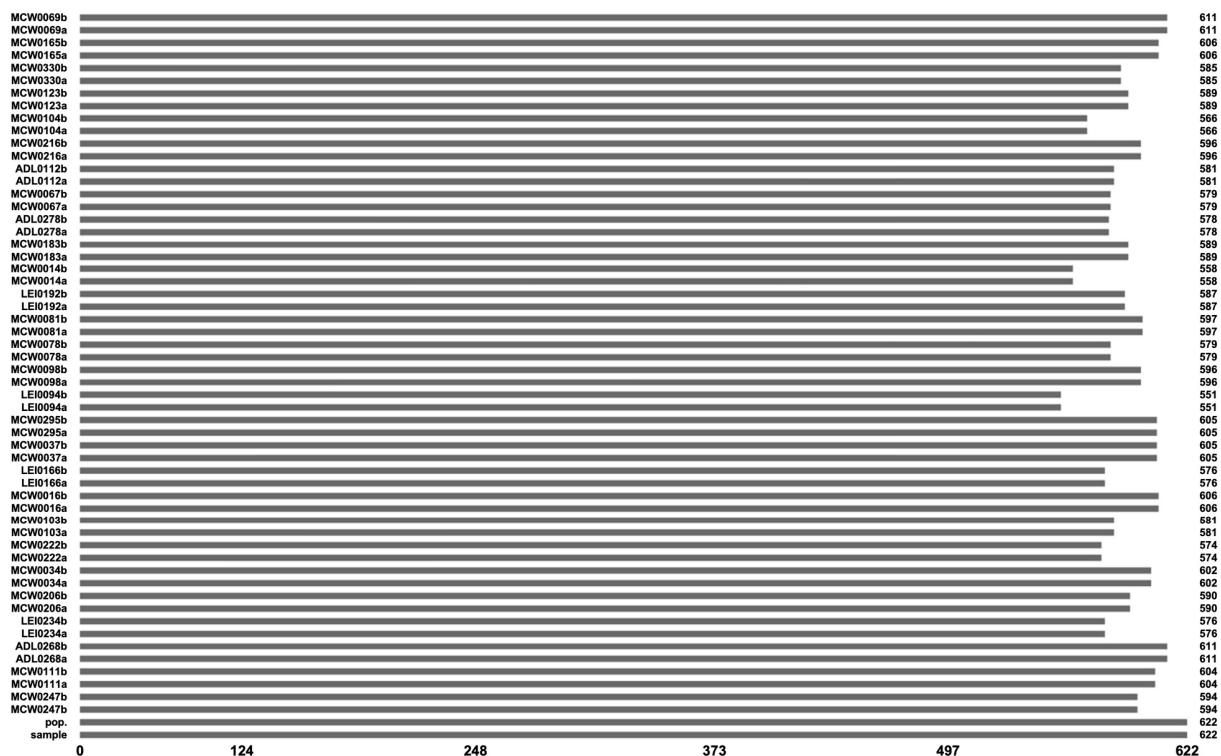


Figure 1. Visualization of nullity by column of the dataset.

Machine-learning data preprocessing: The dataset was composed of genetic data collected from various chicken breeds to facilitate breed classification using ML techniques. It included 622 samples and 58 columns, which consisted of 56 genetic marker parameters, a unique identifier that was labeled “sample,” and a target column that was labeled “pop”. To ensure a comprehensive understanding of data quality before further processing, a thorough examination of missing values was conducted. The MISSINGNO tool was used to generate a visual overview of the distribution of missing data, which enabled the identification of features or samples with incomplete information. Following the results obtained from the MISSINGNO library, varying degrees of missing data were observed across different genetic markers in the dataset. The proportion of missing values ranged from approximately 1.7% to 11.4%, with a significant concentration found in genetic markers belonging to the locus-specific extended information (LEI) sequence group. This finding shows that certain microsatellite markers within the LEI sequence exhibited a higher tendency for incomplete data compared to other genetic markers, such as Microsatellite Chicken Washington (MCW) or avian disease and leukosis (ADL) sequences. Notably, lower levels of missing data were displayed by MCW and ADL markers, indicating their robustness in genetic studies. To ensure that missing data does not introduce bias or negatively affect ML models, it is essential to address these gaps to maintain dataset reliability, improve data preprocessing steps, and ensure robust breed classification predictions.

based on genetic markers. To handle missing values in the dataset, the KNN imputation method was employed. Missing values were estimated based on the similarity between samples, with K samples identified as the most similar to the sample with missing data using a selected distance metric, such as Euclidean distance. The missing values were then replaced with the average numerical values from the nearest neighbors, ensuring that the imputed values preserved the genetic structure and patterns within the dataset.

The optimization of the KNN imputation process involved identifying the optimal number of K samples with the smallest genetic distance, which were used to estimate the missing data. The selection of K is crucial, as a smaller K can lead to overfitting, while a larger K can cause excessive smoothing, thereby reducing the variability of the dataset. The KNN imputation method estimates the missing value $\hat{x}_{i,j}$ using the values of the same feature from the KNN of the sample. The $K = 3$ value was chosen as it was considered optimal for data imputation. The imputed value $\hat{x}_{i,j}$ is computed as follows:

$$\hat{x}_{i,j} = \frac{1}{k} \sum_{m=1}^k x_{m,j} \quad (7)$$

where $x_{m,j}$ is the value of the j -th feature for the m -th nearest neighbor and k : is the number of nearest neighbors to consider.

Following the imputation of missing data, attention was directed toward class imbalance in the target variable. In classification tasks, the target variable is often highly skewed, as shown in the class distribution (Figure 2), which visually represents the number of instances available for each chicken breed in the dataset.

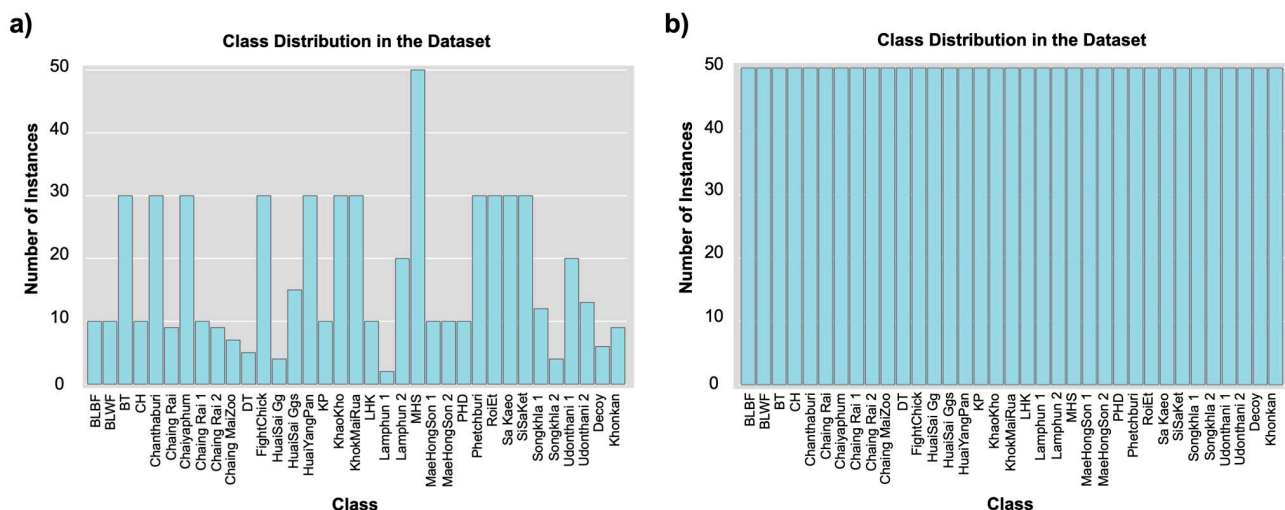


Figure 2. Class distribution before and after dataset balancing. (a) Original dataset showing unequal sample counts across chicken breed classes. (b) Balanced dataset after applying the Synthetic Minority Over-sampling Technique (SMOTE), where class sizes were standardized to approximately 50 instances per class.

A closer examination revealed that the dataset exhibited significant class imbalance, which could affect the performance of any predictive model applied to it. Certain classes, such as “MHS,” contained a disproportionately high number of instances, while other breeds, including “Lamphun_1” and “HuauSai_Gg,” had significantly fewer samples. This disparity may lead to biased ML models that favor well-represented breeds and underperform on underrepresented classes, distorting the predictive accuracy. To ensure unbiased data before developing classification models, a well-balanced dataset is essential. Data imbalance may result in overfitting in dominant classes, where patterns are primarily learned from breeds with abundant data, while those with fewer samples are overlooked.

To mitigate this issue, SMOTE was used to synthesize the imbalanced data. This method generated synthetic samples for the minority class by interpolating between existing minority class samples instead of duplicating existing ones. For a given minority class sample x_i , SMOTE selects one of its KNN x_{zi} and generates a new synthetic sample x_{new} as follows:

$$x_{new} = x_i + \lambda \cdot (x_{zi} - x_i) \quad (8)$$

where

x_i : A minority class sample.

x_{zi} : One of the K-nearest neighbors of x_i

λ : A random number between 0 and 1, which determines the position of the new sample along the line segment connecting x_i and x_{zi} .

The result of applying SMOTE to the dataset was oversampling, which led to a balanced distribution with 50 instances in each class (Figure 3). The final step in data preprocessing was feature standardization, which is critical for ensuring that the data is in an optimal format for training ML models.

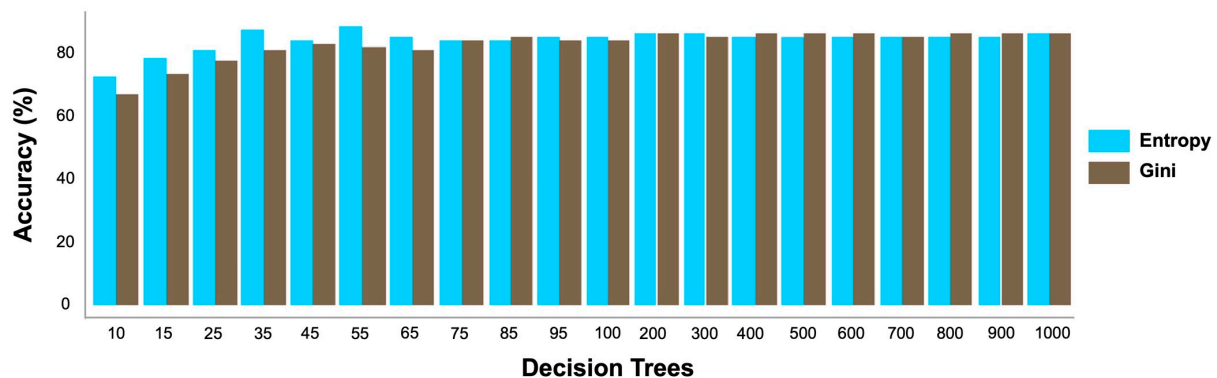


Figure 3. Performance testing of the random forest model on the whole microsatellite genotype chicken dataset with 30 populations. Evolution following the increase of decision trees. Comparison of entropy and Gini functions as criteria to define data split in trees.

One of the most used techniques for standardization is the StandardScaler, which is a powerful tool in ML and data analysis. The features of the dataset were transformed using the StandardScaler to have a mean of 0 and a standard deviation of 1, which was achieved by subtracting the mean of each feature from the data points and dividing by the standard deviation. Mathematically, this transformation is represented as

$$z = \frac{x - \mu}{\sigma} \quad (9)$$

where x is the original feature value, μ is the mean of the feature, σ is the standard deviation of the feature, and z is the standardized feature value.

This standardization process ensured that all features were centered around zero and had a consistent scale, which is particularly important for ML algorithms that are sensitive to the magnitude of input data. Without standardization, features with larger ranges could dominate the learning process, leading to biased or suboptimal model performance.

2.4. Optimized Model: AutoGluon

After data preprocessing, the model was developed using the AutoGluon framework, which automates the entire ML. This framework generates and selects features and identifies the most relevant attributes for the classification task. It trains a diverse set of base models, including tree-based algorithms (LightGBM, XGBoost, CatBoost), lin-

ear models, KNN, and neural networks, applying automatic hyperparameter tuning and cross-validation to ensure robust performance. Following the base-model training, AutoGluon used a multilayer stacking ensemble strategy, combining predictions from multiple models through meta-models to improve generalization and reduce overfitting. The final model was selected based on the F1-score, which is especially suitable for imbalanced classification tasks.

The analysis pipeline, code, and hyperparameters of the final optimized models are available in the Supplementary Information.

3. Results

The RF model was tested on microsatellite genotype data of chicken breeds and demonstrated superior performance compared to other evaluated models, including linear and non-linear feedforward neural networks, KNN, and ensemble models implemented via the AutoGluon framework. The model was initially configured with only the number of decision trees and impurity criterion, using the complete unfiltered dataset. The number of trees was varied from 10 to 1000, and an accuracy of 75.51% was achieved with as few as 15 trees (Figure 3). A notable increase in accuracy was observed up to 55 trees, after which performance plateaued at around 81%. Entropy performed better than the Gini index, particularly between 10 and 65 trees, and was therefore selected as the impurity measure (Figure 3), which showed bootstrap sampling and node generation based on minimum entropy. Populations with fewer than 30 individuals were then excluded to ensure sufficient training data; consequently, 13 chicken populations comprising 433 individuals were used to tune the model to optimize performance.

3.1. Performance Evaluation of the Hyperparameter-Tuned Random Forest Model

Hyperparameters were tuned using 85% of the training data to optimize the performance of the RF model. During tuning, the number of cross-validation folds was varied from 2 to 9, with the best model performance obtained when 3 samples were included per fold. Tuning the hyperparameters of the filtered dataset resulted in an accuracy of 95.38%, which was a significant increase from that of the baseline model. Both models were evaluated using 15% of the testing genotype data. The trained model was used to estimate membership probabilities for each class through a voting mechanism (Table 2). In the training phase, each decision tree was built from a randomly selected subset of the training data. For every test sample, predictions were generated by all trees, which assigned probabilities for class membership. The final prediction was based on the average of these probabilities, with the class receiving the highest average being selected (Table 2). This approach helped reduce overfitting and increased model robustness, and ensured that the final output was the most consistent and probable among the ensemble. A significant disparity was noted between the achieved accuracy (95.38%, Table 3) and the No-Information Rate (16.92%), which underscored the ability of the model to learn from the data. Even the lower bound of the 95% confidence interval (CI = 0.9028, Table 3) reflected strong performance in the worst-case scenario. The Cohen's Kappa value was 0.9492, which shows a high level of agreement between predicted and actual classifications, affirming the reliability of the model beyond random chance (Table 3). The confusion matrix further evidenced the performance of the model, with most values aligned along the diagonal and few off-diagonal errors, indicating minimal misclassifications (Figure 2). This distribution confirmed the robustness of the model in distinguishing between chicken breeds. The F1-score, which represented the harmonic mean of precision and recall, was calculated for each class. Eight out of thirteen classes achieved an F1-score of 1.00, while two classes ("Petch Ggs" = 0.67 and "KMR Ggg" = 0.75) received lower scores, resulting in

a macro average of 0.93, which highlighted the overall effectiveness of the classification model (Table 4).

Table 2. Functionality of the random forest decision according to the probability of class membership. Illustrations of 10 prediction cases among the testing data processed with the trained model.

True Label	First Class		Second Class		Third Class		Final Prediction
	Membership Probability	Population	Membership Probability	Population	Membership Probability	Population	
KK Ggs	52.81	KK Ggs	14.16	KMR Ggg	7.64	SK Ggg	KK Ggs
MHS	97.98	MHS	0.90	UT	0.45	KMR Ggg	MHS
SK Ggg	44.72	SK Ggg	15.96	Chaiya Ggs	10.56	Chatha Ggg	SK Ggg
fight	36.40	fight	13.71	KMR Ggg	12.36	SSK Ggg	fight
KMR Ggg	16.40	fight	14.16	Petch Ggs	12.58	Chatha Ggg	fight
fight	50.34	fight	13.71	RE Ggg	8.76	KMR Ggg	fight
MHS	92.81	MHS	3.60	UT	0.90	Chatha Ggg	MHS
SSK Ggg	48.54	SSK Ggg	19.33	RE Ggg	17.53	fight	SSK Ggg
fight	26.97	fight	22.47	KMR Ggg	13.93	SSK Ggg	fight
Chaiya Ggs	63.60	Chaiya Ggs	10.11	KK Ggs	8.76	KMR Ggg	Chaiya Ggs

Table 3. Performance assessments through overall accuracy, 95% confidence interval, Cohen's Kappa value, and No-information rate after training random forest models across data validation techniques.

Method	Accuracy (%)	Accuracy Std	95% CI	Kappa	NIR
Fixed data split	95.38	-	(0.9028, 1.0000)	0.9492	0.1692
R10FCVT	91.44	0.0408	(0.8904, 0.9384)	0.9065	0.1617
LOOCV	90.99	0.2866	(0.8830, 0.9369)	0.9016	0.1617

Table 4. Classification report showing precision, recall, and F1-score, with a random forest model after hyperparameter tuning and training with fixed data splits.

Population	Precision	Recall	F1-Score
Bt	1.00	1.00	1.00
Chaiya Ggs	1.00	0.83	0.91
Chatha Ggg	1.00	1.00	1.00
fight	0.80	1.00	0.89
HYP Ggs	1.00	1.00	1.00
KK Ggs	0.88	1.00	0.93
KMR Ggg	1.00	0.60	0.75
MHS	1.00	1.00	1.00
Petch Ggs	0.50	1.00	0.67
RE Ggg	1.00	1.00	1.00
SK Ggg	1.00	1.00	1.00
SSK Ggg	1.00	1.00	1.00
UT	1.00	1.00	1.00
macro average	0.94	0.96	0.93
weighted average	0.97	0.95	0.95

3.2. Cross-Validation Results

A further evaluation of the performance of the RF model in classifying chicken breeds was conducted using cross-validation techniques. The overall accuracies and Kappa values facilitated a direct comparison between R10FCV and LOOCV (Table 3). The principal difference between these methods was in the manner of testing the splits. LOOCV used one sample per fold, whereas R10FCV used $k = 10$ samples per fold. Both methods were

effective as high model performance was consistently achieved on the chicken genotype dataset, with accuracies significantly exceeding the No-Information Rate of 16.17% (Table 3). This performance showed that the RF model can learn meaningful patterns from the data, regardless of the validation method used [65]. Kappa values of 0.9065 for R10FCV and 0.9016 for LOOCV were obtained, indicating that reliable classifications were made well beyond random chance. Corresponding accuracies of 91.44% and 90.99% for R10FCV and LOOCV, respectively, further validated the effectiveness of the model. R10FCV was slightly more effective (Table 3). The standard deviation for LOOCV (0.2866) was higher. This variation led to greater fluctuations in accuracy estimates during final averaging. However, the minimal difference in accuracy (0.45%) between the two methods proved the consistency and robustness of the RF model across different validation approaches. Confusion matrices generated from each method were examined (Figure 3). Minor variations were observed between them; thus, the model maintained stable predictive performance, and it was capable of accurately identifying chicken breeds regardless of dataset size or validation method used.

3.3. Loci Importance in Prediction Process

The feature importance analysis was conducted using an RF model that had been trained on a fixed training set. The model identified the contribution of each locus (feature) to prediction accuracy (Agarwal et al., 2023) [66]. The analysis revealed that loci ADL0112, MCW0216, and MCW0111 had the highest importance scores, collectively contributing 20.40% of the predictive power of the model (Table S2), and the remaining 79.60% of prediction accuracy arose from other loci, highlighting the distributed nature of model performance. This finding showed the importance of retaining all loci, as their combined contributions enhanced prediction accuracy and reflected the complexity of the genotype data (Figure 4).

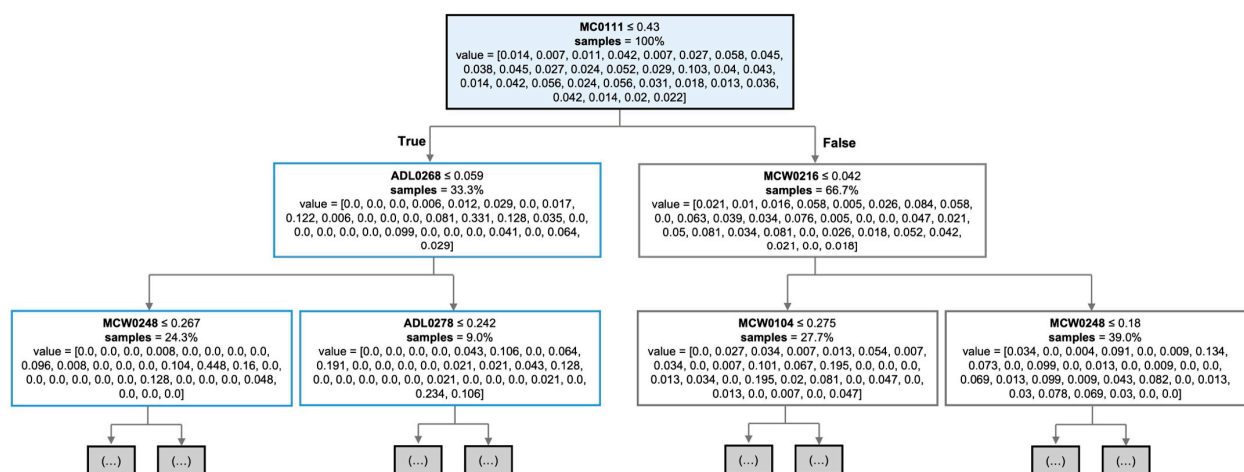


Figure 4. Sample of decision trees, split using entropy function. The top level shows the bootstrap sampling (random sampling) of the total dataset. The lower level shows the beginning of split using the entropy function.

3.4. Using AutoGluon to Optimize Predictive Models in Chicken Breed Classification

The selection of the most effective predictive model for chicken breed classification was facilitated by AutoGluon, an advanced automated ML (AutoML) framework developed by Amazon Web Services. This framework automates model selection, feature engineering, hyperparameter optimization, and ensembling strategies. As an open-source AutoML library, various ML tasks, including tabular data classification, regression, image classification, object detection, and text prediction, were supported by AutoGluon [67]. Owing to its

adaptability, AutoGluon is particularly well-suited for analyzing complex genetic datasets, where classification performance is influenced by intricate relationships between variables. The advantages of AutoGluon include its ability to automate model selection, allowing for the systematic evaluation of multiple ML models to identify the most suitable approach for a given dataset. We used the TabularPredictor of AutoGluon, designed for structured datasets containing both numerical and categorical variables such as genetic markers [68]. A wide range of ML algorithms, including RF, gradient boosting models (XGBoost and LightGBM), multilayer perceptrons, KNN, and SVM, were integrated into the framework [69]. These models were assessed using key performance metrics, including accuracy, precision, recall, and F1-score, which ensured that model selection was guided by empirical performance rather than arbitrary choices. Multilayer stacking, an advanced ensembling technique, was also applied to enhance predictive accuracy by combining multiple base models [70]. It involved training models independently and using a metamodel to learn from their predictions, which improved generalization and reduced overfitting. The prediction results were obtained through a robust 10-fold cross-validation strategy, ensuring reliability and generalizability (Table 5). The evaluation results showed that ensemble learning techniques, particularly WeightedEnsemble_L3 and WeightedEnsemble_L2, achieved the highest accuracy scores of 0.992 and 0.991, respectively (Table 5). These results highlight the effectiveness of stacked ensembling, where multiple base models were combined to enhance predictive performance. The superior accuracy achieved by these ensembles showed that generalization errors were reduced, improving robustness against dataset variability. Strong predictive performance was exhibited by ExtraTreesGini_BAG_L1 (0.989), NeuralNetFastAI_BAG_L2 (0.988), and LightGBMXT_BAG_L2 (0.988) (Figure 5).

Table 5. Performance Evaluation of AutoGluon Models for Chicken Breed Classification.

Model	Score (Accuracy)	Prediction Time (s)	Fit Time (s)	Pred Time Marginal (s)	Fit Time Marginal (s)	Stack Level	Fit Order
WeightedEnsemble_L3	0.992000	3.831986	277.421132	0.000882	0.280986	3	17
WeightedEnsemble_L2	0.991429	0.367359	2.854856	0.000486	0.125387	2	13
ExtraTreesGini_BAG_L1	0.989143	0.110687	1.107739	0.110687	1.107739	1	9
NeuralNetFastAI_BAG_L2	0.988571	3.831105	277.140146	0.204727	7.725666	2	14
ExtraTreesEntr_BAG_L1	0.988000	0.163950	0.799778	0.163950	0.799778	1	10
LightGBMXT_BAG_L2	0.988000	4.605415	361.044465	0.979038	91.629985	2	15
LightGBMXT_BAG_L1	0.987429	1.184527	14.656198	1.184527	14.656198	1	4
RandomForestGini_BAG_L1	0.986286	0.086773	0.875183	0.086773	0.875183	1	6
LightGBM_BAG_L2	0.986286	3.823396	302.657472	0.197019	33.242992	2	16
CatBoost_BAG_L1	0.985714	0.070553	189.957427	0.070553	189.957427	1	8
RandomForestEntr_BAG_L1	0.985143	0.092236	0.821953	0.092236	0.821953	1	7
NeuralNetFastAI_BAG_L1	0.972571	0.072332	4.908889	0.072332	4.908889	1	3
NeuralNetTorch_BAG_L1	0.970857	0.174236	27.011140	0.174236	27.011140	1	12
LightGBM_BAG_L1	0.970857	1.020726	19.746374	1.020726	19.746374	1	5
XGBoost_BAG_L1	0.965714	0.387963	9.517267	0.387963	9.517267	1	11
KNeighborsDist_BAG_L1	0.925143	0.109600	0.003412	0.109600	0.003412	1	2
KNeighborsUnif_BAG_L1	0.860571	0.152794	0.009120	0.152794	0.009120	1	1

Parameter Descriptions: Model: The machine learning model used for classification. Score (Accuracy): The classification accuracy of the model. Prediction Time (s): Time taken to generate predictions. Fit Time (s): Time taken to train the model. Pred Time Marginal (s): Additional time taken for predictions compared to previous models. Fit Time Marginal (s): Additional time required for model training. Stack Level: The level at which the model is stacked in the ensemble process. Fit Order: The order in which the model was fitted during the AutoGluon pipeline.

An accuracy of 0.985 was achieved by CatBoost_BAG_L1, with a prediction time of 0.070 s. Accuracies of 0.988 were obtained by NeuralNetFastAI_BAG_L2 and LightGBMXT_BAG_L2, with fit times of 277 and 361 s, respectively. XGBoost_BAG_L1 and RandomForestGini_BAG_L1 achieved accuracies of 0.965 and 0.986. Higher accuracies of 0.992 and 0.991 were yielded by WeightedEnsemble_L3 and WeightedEnsemble_L2, respec-

tively. Lower accuracies of 0.925 and 0.860 were obtained by KNeighborsDist_BAG_L1 and KNeighborsUnif_BAG_L1 (Figure 6).

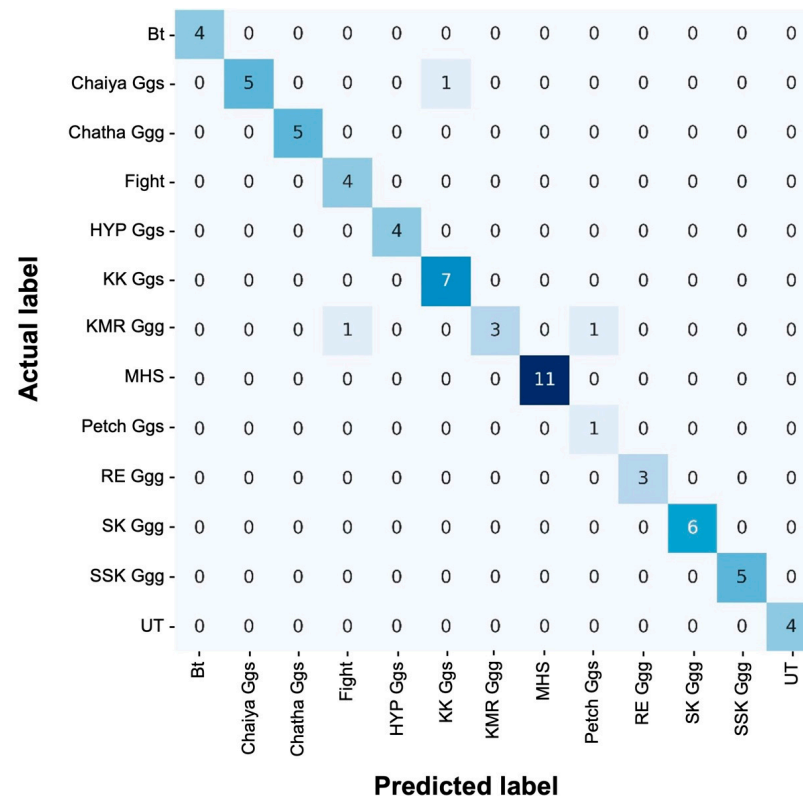


Figure 5. Confusion matrix of random forest trained on 85% of the total dataset and evaluated on 15% of the dataset. Darker blue shades indicate higher numbers of samples, while lighter shades represent lower counts in each cell.

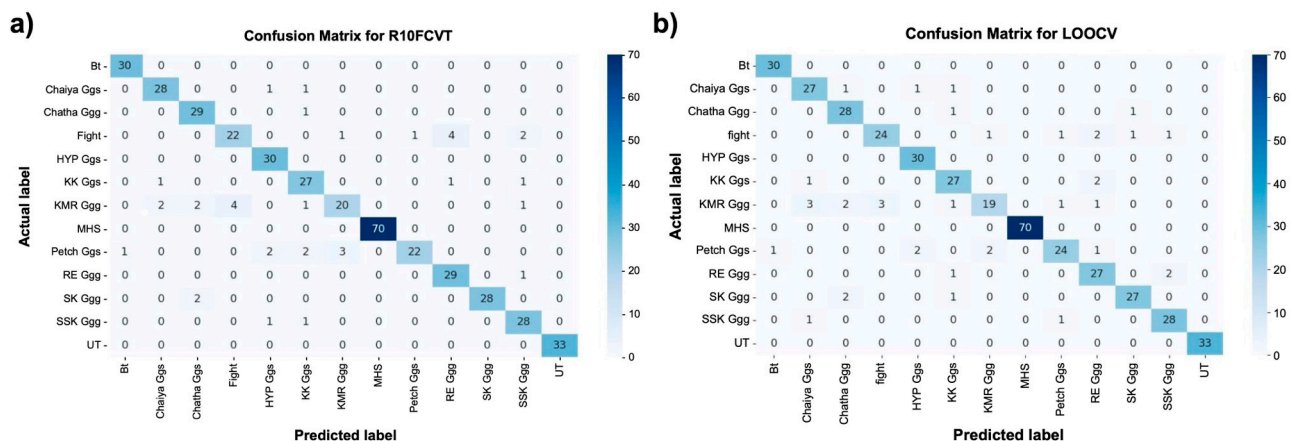


Figure 6. Confusion matrices of the random forest model evaluated using different cross-validation strategies. (a) Confusion matrix obtained from repeated 10-fold cross-validation (R10FCV), where 90% of the dataset was used for training and 10% for testing in each fold. (b) Confusion matrix obtained from leave-one-out cross-validation (LOOCV), in which $n - 1$ individuals were used for training and one individual was used for testing in each iteration. Darker blue shades indicate higher numbers of samples, while lighter shades represent lower counts in each cell.

4. Discussion

Breed determination is used in many applications, including authenticating mono-breed products and supporting breeding programs [71,72]. For this purpose, a variety of approaches have been proposed, including phenotypic identification, genetic testing,

and pedigree analysis [73–75]. The growing use of computer-vision methods has been observed in animal breed confirmation, shifting from purely phenotypic human recognition to computational approaches [76,77]. In this study, the RF and AutoGluon models were used with microsatellite markers, which are frequently applied in population studies and breed classification [40,78,79]. This approach is aligned with previous studies, which have demonstrated the efficacy of this model in the domain of genetics [72,80–82]. The high accuracy demonstrated in this study provides further evidence of the potential of ML in breed prediction, as supported by classification tasks in related fields [83–85]. By contrast, AutoGluon, which automates the entire ML—including data preprocessing, feature engineering, model selection, hyperparameter tuning, and ensemble learning—is capable of training and evaluating diverse models, such as tree-based methods, linear models, neural networks, and KNN. These models were combined through advanced stacking techniques to enhance performance. Consequently, AutoGluon provides a more flexible, automated, and potentially more accurate approach compared to the standalone RF model.

4.1. Performance Metrics, Impurity Criteria, and Marker Optimization for RF Models

The F1-score, which ranges between 0 and 1 with 1 indicating perfect classification, was used as a performance measure [86]. The high macro average F1-score observed reflects an effective balance between precision and recall across all classes, demonstrating that the algorithm performs robustly and effectively handles diverse class distributions in this classification task [60]. Entropy and Gini—widely used impurity measures in decision-tree algorithms—typically produce comparable results [87]. While Gini is more computationally straightforward, entropy provides a more nuanced measure of information gain, which can lead to slightly improved performance in complex classification tasks [61,88]. The use of entropy in this study aligns with the goal of maximizing model accuracy, demonstrating its practical classification superiority. R10FCV is considered the best evaluation method, despite the similar performance observed with LOOCV [89,90]. The results presented here support this, showing slight performance improvement and greater stability with R10FCV. LOOCV is time-consuming during training [90]. Another frequently used evaluation technique with an equivalent objective is the out-of-bag (OOB) estimate [91]. In contrast to cross-validation, which explicitly divides the dataset into training and testing subsets (folds), the OOB method provides internal validation during training by using approximately 37% of each bootstrap sample [92]. Both OOB prediction and cross-validation can be used to evaluate classifier performance, providing robust estimates. However, studies have shown that cross-validation is generally preferred over OOB because it better facilitates the selection of optimal hyperparameters, such as m_{try} in RF models [93,94]. This preference is especially important when fine-tuning model performance, as cross-validation allows more explicit control over data splitting and evaluation. Rasoarahona et al. (2023) [95] proposed an efficient method for selecting a reduced set of microsatellite panels that can be used similarly to the full list for population characterization and origin determination. This methodology demonstrated its efficacy for individual identification in certain animal breeds [96–98]. Applying this method in future experiments with chicken data, within the current RF approach, will enable the RF model to operate with fewer independent features. The performance of the resulting model will be evaluated to determine whether it remains comparable to or surpasses the current performance.

4.2. Model Performance, Computational Considerations, and Locus Importance

AutoGluon was applied to optimize predictive models for chicken breed classification, demonstrating its effectiveness in automating model selection and hyperparameter tuning. The evaluation metrics showed that ensemble techniques, particularly Weighted-

dEnsemble_L3 and WeightedEnsemble_L2, achieved the highest accuracy scores of 0.992 and 0.991, respectively, highlighting the strength of stacked ensembling in improving predictive performance. The superior accuracy of these ensembles suggests that the aggregation of diverse models reduces generalization errors and enhances robustness against data variability. Of the individual models, ExtraTreesGini_BAG_L1 (0.989), NeuralNetFastAI_BAG_L2 (0.988), and LightGBMXT_BAG_L2 (0.988) demonstrated strong predictive performance. The presence of ExtraTrees algorithms among the top models aligns with existing reports of their efficiency in high-dimensional, non-linear data [99]. NeuralNetFastAI models, which are based on deep-learning architectures, and LightGBM models, which are known for their effectiveness in structured data, also achieved high accuracy, reinforcing their utility in breed classification [26]. From a computational efficiency perspective, this experiment was conducted on an Apple M1 system. CatBoost_BAG_L1 achieved 0.985 accuracy with a prediction time of 0.070 s, indicating its suitability for real-time inference. NeuralNetFastAI_BAG_L2 and LightGBMXT_BAG_L2, despite high accuracy, had significantly longer fit times of 277 and 361 s, respectively. This demonstrates a trade-off between model performance and computational cost, which is important in large-scale applications with limited resources. Models such as XGBoost_BAG_L1 (0.965 accuracy) and RandomForestGini_BAG_L1 (0.986 accuracy) performed well, but their longer fit times suggest that ensemble methods like WeightedEnsemble_L3 offer a better balance between accuracy and training efficiency. KNN models (KNeighborsDist_BAG_L1: 0.925, KNeighborsUnif_BAG_L1: 0.860) demonstrated lower accuracy, which highlights their limitations in handling high-dimensional genetic datasets due to their sensitivity to noise and lack of feature selection mechanisms. The complexity of the dataset, characterized by 28 distinct bi-allelic loci, posed a challenge during model training because effective learning requires a substantial number of samples [100,101]. The limited genotype data for certain populations restricted the scope of analysis, indicating that more extensive datasets are necessary in future studies. The stochastic nature of RF algorithms poses a challenge in identifying key loci. However, this method had a significant advantage of RF over alternative ML methods and has been effective in other studies in identifying informative loci panels [28,102–105]. Future studies should compare multiple classification methods across different tasks to determine the most informative loci for chicken breed prediction.

4.3. Computational Trade-Offs in Model Selection

While ensemble and deep-learning models achieved the highest predictive accuracies, the associated computational costs varied significantly [106]. Although LightGBMXT_BAG_L2 and NeuralNetFastAI_BAG_L2 demonstrated strong classification performance, their long-fit times may limit their practicality in time-sensitive or resource-constrained settings [68]. In contrast, CatBoost_BAG_L1 offered a favorable balance between accuracy and computational efficiency, making it suitable for real-time inference. Therefore, both predictive performance and resource requirements should be considered when selecting models for large-scale genetic classification tasks, especially in field applications or under limited computational infrastructure [107].

4.4. Contributions to the Identification and Conservation of Breeds

Effective breed recognition is crucial for breeding program success and conservation of genetic resources, but it remains challenging, especially for non-experts [108–110]. While SNP-based ML models have shown promise in breed classification [72,111], this study explores an alternative method using microsatellite genotypes, which may offer higher informativeness per genetic locus in certain species, such as chicken, and may be less costly if fewer samples are tested [95]. Although image recognition is a modern phenotypic

classification method, it often depends on high-quality images and preprocessing [112]. This study proposes a novel ML-based approach that uses the correlation between microsatellite markers and population-specific traits. The results demonstrate that methods like RF can effectively support breed identification, enriching tools available in animal genetics. However, the accuracy of ML-based approaches using genetic data is still limited by the small genotypic library of animal breeds, which does not represent the full genetic diversity of chickens. A large genotypic library with many individuals across different localities is necessary to improve the reliability of breed identification.

5. Conclusions

This study investigated the effectiveness of the RF model for chicken breed classification using microsatellite genotypes, and strong performance was demonstrated across multiple metrics (Figure 7). The accuracy of the model was significantly improved by AutoGluon's automated ensembling, which outperformed traditional models. Weighted ensembles, particularly those combining ExtraTrees, LightGBM, and deep-learning models, achieved the best results, especially for complex genetic data. However, computational time remains a concern for high-accuracy models. These findings highlight the potential of AutoGluon as an AutoML tool for genetic classification, offering automation and robust performance for biological and agricultural research. Future efforts should focus on optimizing computational efficiency to enable real-time applications. Additionally, future studies could integrate multi-omics data, including approaches applied in recent single-cell transcriptomic analyses of avian performance [113,114], to further enhance the predictive accuracy and biological interpretability of breed classification models.

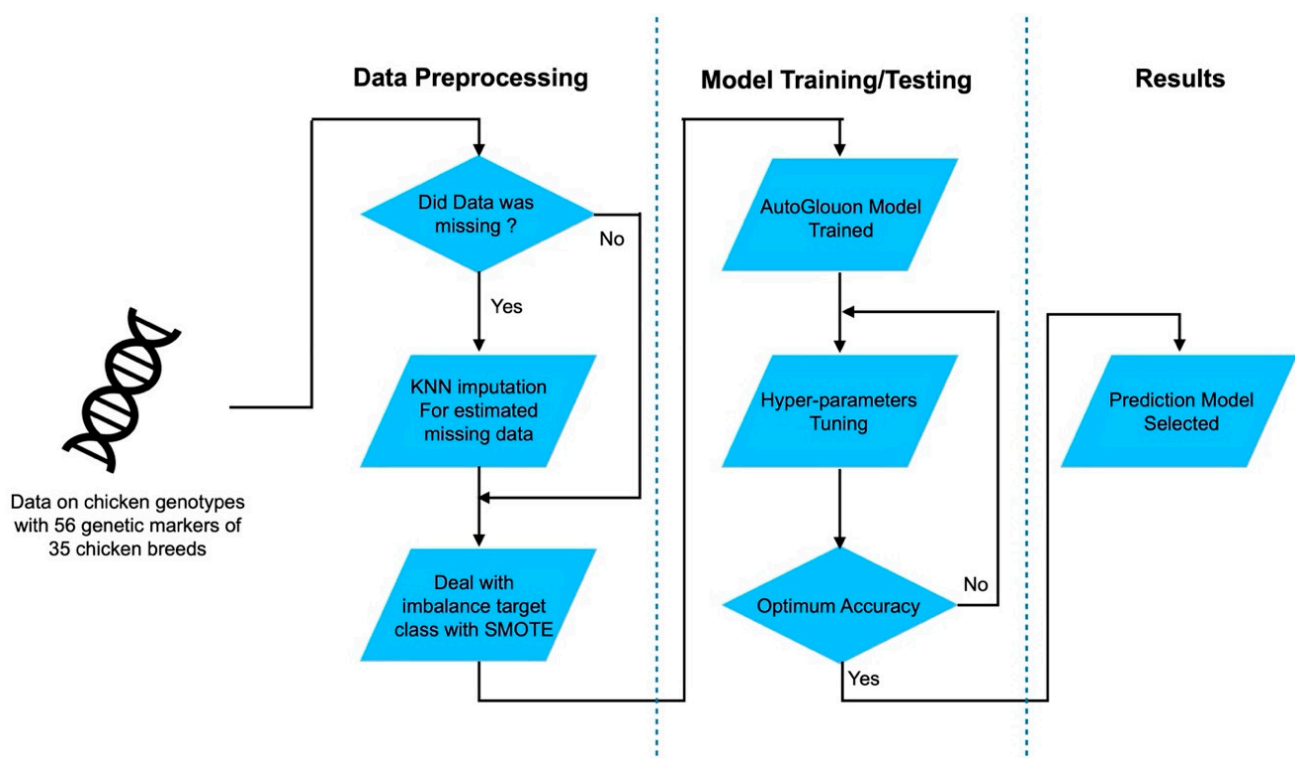


Figure 7. The workflow outlines data preprocessing (handling missing data with k-nearest neighbors imputation and class imbalance with synthetic minority oversampling technique), model training and evaluation using AutoGluon with hyperparameter tuning, and selection of the optimal prediction model.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/biology15010021/s1>, Table S1: Initial materials including 30 populations with number of available individuals. Table S2: Ranking of 28 loci importances in the breed prediction made with the random forest model built with a fixed training dataset. Figure S1: Correlations between predicted and observed breed classifications used as additional accuracy metrics. References [10,20,41,42,115–122] are cited in the supplementary materials.

Author Contributions: Conceptualization, R.F.M.T., S.S., W.S. and K.S.; data curation, K.S.; methodology, R.F.M.T., S.M. and W.S.; software, S.S.; validation, R.F.M.T., S.M., T.B., T.P., W.S. and K.S.; formal analysis, R.F.M.T. and S.S.; investigation, W.S. and K.S.; writing—original draft preparation, R.F.M.T., S.S., W.S. and K.S.; writing—review and editing, R.F.M.T., S.S., S.M., T.B., T.P., W.S. and K.S.; Funding acquisition, W.S. and K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Kasetsart University Research and Development Institute funds, [FF(KU)25.64 and FF(KU)51.67] awarded to W.S. and K.S. The Program Management Unit for Human Resources and Institutional Development and Innovation (PMU-B) under the Program of National Postdoctoral and Postgraduate System (Contract No. B137660130) awarded to R.F.M.T., W.S., and K.S., the Office of the Ministry of Higher Education, Science, Research and Innovation; and the Thailand Science Research and Innovation through the Kasetsart University Reinventing University Program 2022 awarded to S.M. and K.S.; the Betagro Group, (No. 6501.0901.1/68) awarded to K.S., the National Research Council of Thailand (NRCT) grant, (contract No. NRCT.MHESI/105/2564) awarded to W.S. and K.S. No funding source was involved in the study design, collection, analysis, data interpretation, report writing, or the decision to submit the article for publication.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Acknowledgments: The authors would like to thank the National Science and Technology Development Agency (NSTDA) Supercomputer Center (ThaiSC) for supporting us with server analysis services.

Conflicts of Interest: No potential conflict of interest relevant to this article was reported.

References

1. Weigend, S.; Romanov, M.N.; Rath, D. Methodologies to Identify, Evaluate and Conserve Poultry Genetic Resources. In Proceedings of the XXII World's Poultry Congress, Istanbul, Turkey, 8–13 June 2004; World's Poultry Science Association (WPSA)—Turkish Branch: Istanbul, Turkey, 2004; p. 84.
2. Gjedrem, T.; Baranski, M. *Selective Breeding in Aquaculture: An Introduction*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010; Volume 10.
3. Felius, M.; Theunissen, B.; Lenstra, J. Conservation of cattle genetic resources: The role of breeds. *J. Agric. Sci.* **2015**, *153*, 152–162. [CrossRef]
4. Dalvit, C.; De Marchi, M.; Dal Zotto, R.; Gervaso, M.; Meuwissen, T.; Cassandro, M. Breed assignment test in four Italian beef cattle breeds. *Meat Sci.* **2008**, *80*, 389–395. [CrossRef] [PubMed]
5. Felix, G.A.; Soares Fioravanti, M.C.; Cassandro, M.; Tormen, N.; Quadros, J.; Soares Juliano, R.; Alves do Egito, A.; de Moura, M.I.; Piovezan, U. Bovine breeds identification by trichological analysis. *Animals* **2019**, *9*, 761. [CrossRef]
6. Peng, W.; Yang, H.; Cai, K.; Zhou, L.; Tan, Z.; Wu, K. Molecular identification of the Danzhou chicken breed in China using DNA barcoding. *Mitochondrial DNA Part B* **2019**, *4*, 2459–2463. [CrossRef] [PubMed]
7. Ghosh, P.; Mustafi, S.; Mukherjee, K.; Dan, S.; Roy, K.; Mandal, S.N.; Banik, S. Image-based identification of animal breeds using deep learning. In *Deep Learning for Unmanned Systems*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 415–445.
8. Addisu, H.; Hailu, M.; Zewdu, W. Indigenous chicken production system and breeding practice in North Wollo, Amhara Region, Ethiopia. *Poult. Fish. Wildl. Sci.* **2013**, *1*, 108.
9. Desta, T.T.; Wakeyo, O. Breeding practice of indigenous village chickens, and traits and breed preferences of smallholder farmers. *Vet. Med. Sci.* **2024**, *10*, e1517. [CrossRef]

10. Tanglertpaibul, N.; Budi, T.; Nguyen, C.P.T.; Singchat, W.; Wongloet, W.; Kumnan, N.; Chalermwong, P.; Luu, A.H.; Noito, K.; Panthum, T. Samae Dam chicken: A variety of the Pradu Hang Dam breed revealed from microsatellite genotyping data. *Anim. Biosci.* **2024**, *37*, 2033. [\[CrossRef\]](#)
11. Vieira, M.L.C.; Santini, L.; Diniz, A.L.; Munhoz, C.d.F. Microsatellite markers: What they mean and why they are so useful. *Genet. Mol. Biol.* **2016**, *39*, 312–328. [\[CrossRef\]](#)
12. Weising, K.; Winter, P.; Hüttel, B.; Kahl, G. Microsatellite markers for molecular breeding. *J. Crop Prod.* **1997**, *1*, 113–143. [\[CrossRef\]](#)
13. McCouch, S.R.; Chen, X.; Panaud, O.; Temnykh, S.; Xu, Y.; Cho, Y.G.; Huang, N.; Ishii, T.; Blair, M. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol. Biol.* **1997**, *35*, 89–99. [\[CrossRef\]](#)
14. Guichoux, E.; Lagache, L.; Wagner, S.; Chaumeil, P.; Léger, P.; Lepais, O.; Lepoittevin, C.; Malausa, T.; Revardel, E.; Salin, F. Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* **2011**, *11*, 591–611. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Balloux, F.; Lugon-Moulin, N. The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* **2002**, *11*, 155–165. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Chang, C.-S.; Chen, C.; Berthouly-Salazar, C.; Chazara, O.; Lee, Y.; Chang, C.; Chang, K.; Bed'Hom, B.; Tixier-Boichard, M. A global analysis of molecular markers and phenotypic traits in local chicken breeds in Taiwan. *Anim. Genet.* **2012**, *43*, 172–182. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Abebe, A.S.; Mikko, S.; Johansson, A.M. Genetic diversity of five local Swedish chicken breeds detected by microsatellite markers. *PLoS ONE* **2015**, *10*, e0120580. [\[CrossRef\]](#)
18. Sartore, S.; Sacchi, P.; Soglia, D.; Maione, S.; Schiavone, A.; De Marco, M.; Ceccobelli, S.; Lasagna, E.; Rasero, R. Genetic variability of two Italian indigenous chicken breeds inferred from microsatellite marker analysis. *Br. Poult. Sci.* **2016**, *57*, 435–443. [\[CrossRef\]](#)
19. Fathi, M.; Al-Homidan, I.; Motawei, M.; Abou-Emera, O.; El-Zarei, M. Evaluation of genetic diversity of Saudi native chicken populations using microsatellite markers. *Poult. Sci.* **2017**, *96*, 530–536. [\[CrossRef\]](#)
20. Wattanadilokcahtkun, P.; Chalermwong, P.; Singchat, W.; Wongloet, W.; Chaiyes, A.; Tanglertpaibul, N.; Budi, T.; Panthum, T.; Ariyaphong, N.; Ahmad, S.F. Genetic admixture and diversity in Thai domestic chickens revealed through analysis of Lao Pa Koi fighting cocks. *PLoS ONE* **2023**, *18*, e0289983. [\[CrossRef\]](#)
21. Duc, T.L.; Leiva, R.G.; Casari, P.; Östberg, P.-O. Machine learning methods for reliable resource provisioning in edge-cloud computing: A survey. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 94. [\[CrossRef\]](#)
22. Ghahramani, Z. Unsupervised learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 72–112.
23. Cord, M.; Cunningham, P. *Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008.
24. Rokach, L.; Maimon, O. Top-down induction of decision trees classifiers—a survey. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **2005**, *35*, 476–487. [\[CrossRef\]](#)
25. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
26. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3149–3157.
27. Qi, Y. Random forest for bioinformatics. In *Ensemble Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 307–323.
28. Montesinos López, O.A.; Montesinos López, A.; Crossa, J. Random forest for genomic prediction. In *Multivariate Statistical Machine Learning Methods for Genomic Prediction*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 633–681.
29. Breiman, L. *Some Infinity Theory for Predictor Ensembles*; Technical Report 579; Statistics Department UCB: Berkeley, CA, USA, 2000.
30. Breiman, L. Randomizing outputs to increase prediction accuracy. *Mach. Learn.* **2000**, *40*, 229–242. [\[CrossRef\]](#)
31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
32. Breiman, L. *Consistency for a Simple Model of Random Forests*; Technical Report; University of California at Berkeley: Berkeley, CA, USA, 2004; Volume 670.
33. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Chen, X.; Ishwaran, H. Random forests for genomic data analysis. *Genomics* **2012**, *99*, 323–329. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Biau, G.; Scornet, E. A random forest guided tour. *Test* **2016**, *25*, 197–227. [\[CrossRef\]](#)
36. Rabiei, N.; Soltanian, A.R.; Farhadian, M.; Bahreini, F. the performance evaluation of the random forest algorithm for a gene selection in identifying genes associated with resectable pancreatic cancer in microarray dataset: A retrospective study. *Cell J.* **2023**, *25*, 347.
37. Oshiro, T.M.; Perez, P.S.; Baranauskas, J.A. How many trees in a random forest? In Proceedings of the International Workshop on Machine Learning and Data Mining in Pattern Recognition, Berlin, Germany, 13–20 July 2012; pp. 154–168.
38. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249. [\[CrossRef\]](#)

39. Ganaie, M.A.; Hu, M.; Malik, A.K.; Tanveer, M.; Suganthan, P.N. Ensemble deep learning: A review. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105151. [\[CrossRef\]](#)
40. Burócziová, M.; Říha, J. Horse breed discrimination using machine learning methods. *J. Appl. Genet.* **2009**, *50*, 375–377. [\[CrossRef\]](#)
41. Hata, A.; Nunome, M.; Suwanasopee, T.; Duengkae, P.; Chaiwatana, S.; Chamchumroon, W.; Suzuki, T.; Koonawootrittriron, S.; Matsuda, Y.; Srikulnath, K. Origin and evolutionary history of domestic chickens inferred from a large population study of Thai red junglefowl and indigenous chickens. *Sci. Rep.* **2021**, *11*, 2035. [\[CrossRef\]](#)
42. Singchat, W.; Chaibes, A.; Wongloet, W.; Ariyaphong, N.; Jaisamut, K.; Panthum, T.; Ahmad, S.F.; Chaleekarn, W.; Suksavate, W.; Inpota, M. Red junglefowl resource management guide: Bioresource reintroduction for sustainable food security in Thailand. *Sustainability* **2022**, *14*, 7895. [\[CrossRef\]](#)
43. FAO. Molecular genetic characterization of animal genetic resources. In *FAO Animal Production and Health Guidelines*; FAO: Rome, Italy, 2011; Volume 9.
44. Shwartz-Ziv, R.; Goldblum, M.; Li, Y.; Bruss, C.B.; Wilson, A.G. Simplifying neural network training under class imbalance. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 35218–35245.
45. Ahsan, M.M.; Mahmud, M.P.; Saha, P.K.; Gupta, K.D.; Siddique, Z. Effect of data scaling methods on machine learning algorithms and model performance. *Technologies* **2021**, *9*, 52. [\[CrossRef\]](#)
46. Alshaer, H. Studying the Effects of Feature Scaling in Machine Learning. Master's Thesis, North Carolina Agricultural and Technical State University, Greensboro, NC, USA, 2021.
47. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Montreal, QC, Canada, 20–25 August 1995; pp. 1137–1145.
48. Van Rossum, G. Python programming language. In Proceedings of the USENIX Annual Technical Conference, Santa Clara, CA, USA, 17–22 June 2007; pp. 1–36.
49. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic differentiation in PyTorch. In Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
50. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. [\[CrossRef\]](#)
51. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
52. Kramer, O. *Machine Learning for Evolution Strategies*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 20.
53. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
54. Hjerpe, A. Computing Random Forests Variable Importance Measures (vim) on Mixed Numerical and Categorical Data. Master's Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2016.
55. Xu, Z.; Dan, C.; Khim, J.; Ravikumar, P. Class-weighted classification: Trade-offs and robust approaches. In Proceedings of the International Conference on Machine Learning, Online, 13–18 July 2020; pp. 10544–10554.
56. Falkner, S.; Klein, A.; Hutter, F. Practical hyperparameter optimization for deep learning. In Proceedings of the ICLR 2018, Vancouver, BC, Canada, 30 April–3 May 2018.
57. Stone, M. Cross-validated choice and assessment of statistical predictions. *J. R. Stat. Soc. Ser. B (Methodol.)* **1974**, *36*, 111–133. [\[CrossRef\]](#)
58. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [\[CrossRef\]](#)
59. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Medica* **2012**, *22*, 276–282. [\[CrossRef\]](#)
60. Grandini, M.; Bagli, E.; Visani, G. Metrics for multi-class classification: An overview. *arXiv* **2020**, arXiv:2008.05756. [\[CrossRef\]](#)
61. Tangirala, S. Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 612–619. [\[CrossRef\]](#)
62. Murad, M.A.H.; Paul, M.K. A Hybrid Preprocessing Approach for the Classification of Class Imbalanced Data. In Proceedings of the 2023 6th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 7–9 December 2023; pp. 1–6.
63. Akay, M.F. Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Syst. Appl.* **2009**, *36*, 3240–3247. [\[CrossRef\]](#)
64. Lipton, Z.C.; Elkan, C.; Naryanaswamy, B. Optimal thresholding of classifiers to maximize F1 measure. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Nancy, France, 14–18 September 2014; pp. 225–239.
65. Bicego, M.; Mensi, A. Null/No Information Rate (NIR): A statistical test to assess if a classification accuracy is significant for a given problem. *arXiv* **2023**, arXiv:2306.06140. [\[CrossRef\]](#)
66. Agarwal, A.; Kenney, A.M.; Tan, Y.S.; Tang, T.M.; Yu, B. MDI+: A flexible random forest-based feature importance framework. *arXiv* **2023**, arXiv:2307.01932.

67. Zoph, B.; Vasudevan, V.; Shlens, J.; Le, Q.V. Learning transferable architectures for scalable image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8697–8710.
68. Erickson, N.; Mueller, J.; Shirkov, A.; Zhang, H.; Larroy, P.; Li, M.; Smola, A. Autogluon-tabular: Robust and accurate automl for structured data. *arXiv* **2020**, arXiv:2003.06505.
69. Li, Z.; Lu, T.; He, X.; Montillet, J.-P.; Tao, R. An improved cyclic multi model-eXtreme gradient boosting (CMM-XGBoost) forecasting algorithm on the GNSS vertical time series. *Adv. Space Res.* **2023**, *71*, 912–935. [[CrossRef](#)]
70. Hutter, F.; Kotthoff, L.; Vanschoren, J. *Automated Machine Learning: Methods, Systems, Challenges*; Springer Nature: Berlin/Heidelberg, Germany, 2019.
71. Putnová, L.; Štohl, R. Comparing assignment-based approaches to breed identification within a large set of horses. *J. Appl. Genet.* **2019**, *60*, 187–198. [[CrossRef](#)]
72. Kasarda, R.; Moravčíková, N.; Mészáros, G.; Simčič, M.; Zaborski, D. Classification of cattle breeds based on the random forest approach. *Livest. Sci.* **2023**, *267*, 105143. [[CrossRef](#)]
73. Quinteiro, J.; Sotelo, C.G.; Rehbein, H.; Pryde, S.E.; Medina, I.; Pérez-Martín, R.; Rey-Mendez, M.; Mackie, I. Use of mtDNA direct polymerase chain reaction (PCR) sequencing and PCR– restriction fragment length polymorphism methodologies in species identification of canned tuna. *J. Agric. Food Chem.* **1998**, *46*, 1662–1669. [[CrossRef](#)]
74. Oravcová, M. Pedigree analysis in White Shorthaired goat: First results. *Arch. Anim. Breed.* **2013**, *56*, 547–554. [[CrossRef](#)]
75. Jasielczuk, I.; Gurgul, A.; Szmatoła, T.; Radko, A.; Majewska, A.; Sosin, E.; Litwińczuk, Z.; Rubiś, D.; Ząbek, T. The use of SNP markers for cattle breed identification. *J. Appl. Genet.* **2024**, *65*, 575–589. [[CrossRef](#)]
76. Rudenko, O.; Megel, Y.; Bezsonov, O.; Rybalka, A. Cattle breed identification and live weight evaluation on the basis of machine learning and computer vision. *CMIS* **2020**, *2608*, 939–954. [[CrossRef](#)]
77. Khan, S.S.; Doohan, N.V.; Gupta, M.; Jaffari, S.; Chourasia, A.; Joshi, K.; Panchal, B. Hybrid deep learning approach for enhanced animal breed classification and prediction. *Trait. Du Signal* **2023**, *40*, 2087. [[CrossRef](#)]
78. Ciofi, C.; Funk, S.M.; Coote, T.; Cheesman, D.J.; Hammond, R.L.; Saccheri, I.J.; Bruford, M.W. Genotyping with microsatellite markers. In *Molecular Tools for Screening Biodiversity: Plants and Animals*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 195–201.
79. Mihailova, Y.; Rusanov, K.; Rusanova, M.; Vassileva, P.; Atanassov, I.; Nikolov, V.; Todorovska, E.G. Genetic diversity and population structure of Bulgarian autochthonous sheep breeds revealed by microsatellite analysis. *Animals* **2023**, *13*, 1878. [[CrossRef](#)] [[PubMed](#)]
80. Reif, D.M.; Motsinger, A.A.; McKinney, B.A.; Crowe, J.E.; Moore, J.H. Feature selection using a random forests classifier for the integrated analysis of multiple data types. In Proceedings of the 2006 IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology, Toronto, ON, Canada, 28–29 September 2006; pp. 1–8.
81. Holliday, J.A.; Wang, T.; Aitken, S. Predicting adaptive phenotypes from multilocus genotypes in Sitka spruce (*Picea sitchensis*) using random forest. *G3 Genes | Genomes | Genet.* **2012**, *2*, 1085–1093. [[CrossRef](#)]
82. Zhao, J.; Bodner, G.; Rewald, B. Phenotyping: Using machine learning for improved pairwise genotype classification based on root traits. *Front. Plant Sci.* **2016**, *7*, 1864. [[CrossRef](#)]
83. Saiprasath, G.; Babu, N.; ArunPriyan, J.; Vinayakumar, R.; Sowmya, V.; Soman, K. Performance comparison of machine learning algorithms for malaria detection using microscopic images. *Int. J. Curr. Res. Acad. Rev.* **2019**, *6*, 86–90.
84. Bezsonov, O.; Lebediev, O.; Lebediev, V.; Megel, Y.; Prochukhan, D.; Rudenko, O. Breed recognition and estimation of live weight of cattle based on methods of machine learning and computer vision. *East.-Eur. J. Enterp. Technol.* **2021**, *6*, 114.
85. Ghosh, P.; Mandal, S.N. PigB: Intelligent pig breeds classification using supervised machine learning algorithms. *Int. J. Artif. Intell. Soft Comput.* **2022**, *7*, 242–266. [[CrossRef](#)]
86. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [[CrossRef](#)]
87. Raileanu, L.E.; Stoffel, K. Theoretical comparison between the gini index and information gain criteria. *Ann. Math. Artif. Intell.* **2004**, *41*, 77–93. [[CrossRef](#)]
88. Mustafa, O.M.; Ahmed, O.M.; Saeed, V.A. Comparative analysis of decision tree algorithms using gini and entropy criteria on the forest covertypes dataset. In Proceedings of the International Conference on Innovations in Computing Research, Athens, Greece, 12–14 August 2024; pp. 185–193.
89. Ling, N.E.; Hasan, Y.A. Evaluation Method in Random Forest as Applied to Microarray Data. *Malays. J. Math. Sci.* **2008**, *2*, 73–81.
90. Tran, L.; He, K.; Wang, D.; Jiang, H. A cross-validation statistical framework for asymmetric data integration. *Biometrics* **2023**, *79*, 1280–1292. [[CrossRef](#)] [[PubMed](#)]
91. Ciss, S. Generalization Error and Out-of-Bag Bounds in Random (Uniform) Forests. Preprint, HAL Open Science, 2015. Available online: <https://hal.science/hal-01110524> (accessed on 15 May 2024).
92. Breiman, L. Out-of-Bag Estimation. 1996. Available online: <https://www.stat.berkeley.edu/breiman/OOBestimation.pdf> (accessed on 15 May 2024).

93. Kodovský, J. *Ensemble Classification in Steganalysis Cross-Validation and AdaBoost*; Tech. Rep. Digit. Data Embed. Lab. (DDE); Binghamton University: Binghamton, NY, USA, 2011.
94. Janitzka, S.; Hornung, R. On the overestimation of random forest's out-of-bag error. *PLoS ONE* **2018**, *13*, e0201904. [[CrossRef](#)]
95. Rasoarahona, R.; Wattanadilokchatkun, P.; Panthum, T.; Thong, T.; Singchat, W.; Ahmad, S.F.; Chaibes, A.; Han, K.; Kraichak, E.; Muangmai, N. Optimizing microsatellite marker panels for genetic diversity and population genetic studies: An ant colony algorithm approach with polymorphic information content. *Biology* **2023**, *12*, 1280. [[CrossRef](#)]
96. Jaito, W.; Singchat, W.; Patta, C.; Thatukan, C.; Kumnan, N.; Chalermwong, P.; Budi, T.; Panthum, T.; Wongloet, W.; Wattanadilokchatkun, P. Shared alleles and genetic structures in different Thai domestic cat breeds: The possible influence of common racial origins. *Genom. Inform.* **2024**, *22*, 12. [[CrossRef](#)]
97. Patta, C.; Singchat, W.; Thatukan, C.; Jaito, W.; Kumnan, N.; Chalermwong, P.; Panthum, T.; Budi, T.; Wongloet, W.; Wattanadilokchatkun, P. Optimizing Bangkaew dog breed identification using DNA technology. *Genes Genom.* **2024**, *46*, 659–669. [[CrossRef](#)]
98. Quainoo, D.K.; Chalermwong, P.; Muangsuk, P.; Nguyen, T.H.D.; Panthum, T.; Singchat, W.; Budi, T.; Duengkae, P.; Suksavate, W.; Chaibes, A. Genetic insights for enhancing conservation strategies in captive and wild Asian elephants through improved non-invasive DNA-based individual identification. *PLoS ONE* **2025**, *20*, e0320480. [[CrossRef](#)]
99. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **2006**, *63*, 3–42. [[CrossRef](#)]
100. Hall, M.A. *Correlation-Based Feature Selection for Machine Learning*; The University of Waikato: Hamilton, New Zealand, 1999.
101. Omary, Z.; Mtenzi, F. Machine learning approach to identifying the dataset threshold for the performance estimators in supervised learning. *Int. J. Infonomics* **2010**, *3*, 314–325. [[CrossRef](#)]
102. Trevor, H.; Robert, T.; Jerome, F. *The Elements of Statistical Learning*; Springer Series in Statistics; Springer: New York, NY, USA, 2001.
103. Kawakubo, H.; Yoshida, H. Rapid feature selection based on random forests for high-dimensional data. In Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA), Las Vegas, NV, USA, 16–19 July 2012; pp. 1–6.
104. Fox, E.W.; Hill, R.A.; Leibowitz, S.G.; Olsen, A.R.; Thornbrugh, D.J.; Weber, M.H. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. *Environ. Monit. Assess.* **2017**, *189*, 316. [[CrossRef](#)]
105. Sylvester, E.V.; Bentzen, P.; Bradbury, I.R.; Clément, M.; Pearce, J.; Horne, J.; Beiko, R.G. Applications of random forest feature selection for fine-scale genetic population assignment. *Evol. Appl.* **2018**, *11*, 153–165. [[CrossRef](#)]
106. Zanotti, M. The cost of ensembling: Is it always worth combining? *arXiv* **2025**, arXiv:2506.04677. [[CrossRef](#)]
107. Erickson, P.A.; Weller, C.A.; Song, D.Y.; Bangerter, A.S.; Schmidt, P.; Bergland, A.O. Unique genetic signatures of local adaptation over space and time for diapause, an ecologically relevant complex trait, in *Drosophila melanogaster*. *PLoS Genet.* **2020**, *16*, e1009110. [[CrossRef](#)] [[PubMed](#)]
108. Mtileni, B.; Muchadeyi, F.; Maiwashe, A.; Chimonyo, M.; Dzama, K. Conservation and utilisation of indigenous chicken genetic resources in Southern Africa. *World's Poult. Sci. J.* **2012**, *68*, 727–748. [[CrossRef](#)]
109. Chebo, C.; Betsha, S.; Melesse, A. Chicken genetic diversity, improvement strategies and impacts on egg productivity in Ethiopia: A review. *World's Poult. Sci. J.* **2022**, *78*, 803–821. [[CrossRef](#)]
110. Himel, G.M.S.; Islam, M.M.; Rahaman, M. Utilizing EfficientNet for sheep breed identification in low-resolution images. *Syst. Soft Comput.* **2024**, *6*, 200093. [[CrossRef](#)]
111. Liu, R.; Xu, Z.; Teng, J.; Pan, X.; Lin, Q.; Cai, X.; Diao, S.; Feng, X.; Yuan, X.; Li, J. Evaluation of six machine learning classification algorithms in pig breed identification using SNPs array data. *Anim. Genet.* **2023**, *54*, 113–122. [[CrossRef](#)]
112. Kumar, R.; Sharma, M.; Dhawale, K.; Singal, G. Identification of dog breeds using deep learning. In Proceedings of the 2019 IEEE 9th International Conference on Advanced Computing (IACC), Tiruchirappalli, India, 13–14 December 2019; pp. 193–198.
113. Leng, D.; Zeng, B.; Wang, T.; Chen, B.L.; Li, D.Y.; Li, Z.J. Single-nucleus and single-cell RNA sequencing of the chicken hypothalamic–pituitary–ovarian axis provides new insights into the molecular regulatory mechanisms of ovarian development. *Zool. Res.* **2024**, *45*, 1088–1107. [[CrossRef](#)]
114. Wang, T.; Leng, D.; Cai, Z.; Chen, B.; Li, J.; Kui, H.; Li, D.; Li, Z. Insights into left–right asymmetric development of the chicken ovary at the single-cell level. *J. Genet. Genom.* **2024**, *51*, 1265–1277. [[CrossRef](#)]
115. Munkong, P.; Suwanasopee, T.; Koonawoottritriron, S. Morphometric Analysis of Kai Dam Nil Kaset (Sart) Eggs: Implications for Production and Selection. *Khon Kaen Agric. J.* **2024**, *52*, 214.
116. Maneechot, N.; Tunim, S.; Wattanachant, C.; Khongsen, M.; Sukteab, P.; Phongphanich, P. Genetic Diversity of Native Chicken Populations and Red Jungle Fowl in Southern Thailand Based on Mitochondrial DNA D-loop Region. *Braz. J. Poult. Sci.* **2025**, *27*, 001–010. [[CrossRef](#)]
117. Mancinelli, A.C.; Menchetti, L.; Birolo, M.; Bittante, G.; Chiattelli, D.; Castellini, C. Crossbreeding to Improve Local Chicken Breeds: Predicting Growth Performance of the Crosses Using the Gompertz Model and Estimated Heterosis. *Poult. Sci.* **2023**, *102*, 102783. [[CrossRef](#)]

118. Siddiqui, S.A.; Rahmatullah, R.A.; Achyar, A.; Atifah, Y.; Ahmad, A.; Fitriani, A. Dong Tao Chickens in Vietnam—A Critical Review. *World's Poult. Sci. J.* **2024**, *80*, 1241–1263. [[CrossRef](#)]
119. Budi, T.; Luu, A.H.; Singchat, W.; Wongloet, W.; Rey, J.; Kumnan, N.; Chalermwong, P.; Nguyen, C.P.T.; Panthum, T.; Tanglertpaibul, N.; et al. Purposive Breeding Strategies Drive Genetic Differentiation in Thai Fighting Cock Breeds. *Genes Genom.* **2024**, *46*, 1225–1237. [[CrossRef](#)] [[PubMed](#)]
120. Phromnoi, S.; Yeamkong, S.; Mingchai, C. Phenotypic Characteristics and Morphology of Khiew-Phalee Chicken in Uttaradit Province. *Rajamangala Univ. Technol. Srivijaya Res. J.* **2023**, *15*, 37–48.
121. Somkuna, E.; Intaravicha, N.; Maksuwan, A. Effects of Natural Environmental Structure on Growth Performance of Indigenous Chicks (Lueng Hang Khao). *J. Vocat. Educ. Agric.* **2023**, *6*, 68–83.
122. Wongloet, W.; Singchat, W.; Chaiyes, A.; Ali, H.; Piangporntip, S.; Ariyaphong, N.; Budi, T.; Thienpreecha, W.; Wannakan, W.; Mungmee, A.; et al. Environmental and Socio-Cultural Factors Impacting the Unique Gene Pool Pattern of Mae Hong Son Chicken. *Animals* **2023**, *13*, 1949. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.