Research Article

www.dergipark.org.tr/en/pub/mediterranean

MEDITERRANEAN AGRICULTURAL SCIENCES
(2025) 38(2): 85-90
DOI: 10.29136/mediterranean.1622702

Development and evaluation of machine learning models for predicting chicken meat production

Kemal ESKIOGLU, Berkant Ismail YILDIZ, Demir OZDEMIR

Akdeniz University, Faculty of Agriculture, Department of Agricultural Biotechnology, 07058, Antalya, Türkiye

Corresponding author: D. Ozdemir, e-mail: dozdemir@akdeniz.edu.tr
Author(s) e-mail: kemaleskioglu94@gmail.com, berkantyildizz@gmail.com

ARTICLE INFO

Received: January 18, 2025
Received in revised form: April 25, 2025
Accepted: May 20, 2025

***Keywords*:**

Chicken meat production
Machine learning models
Predictive modeling
Agricultural data analysis
Sustainable poultry production

ABSTRACT

Chicken meat production plays a vital role in the global food industry by providing a cost-effective and sustainable protein source for a rapidly growing population. Given strategic importance, accurately estimating production levels is essential for improving operational efficiency, optimizing resource use, and responding to market demand. In recent years, data-oriented methods have become integral to modern agriculture, with machine learning models emerging as powerful tools for modeling agricultural outputs. This study aims to develop and compare predictive models for chicken meat production in Türkiye using four machine learning algorithms: Linear Regression (LR), Random Forest (RF), k-Nearest Neighbors (k-NN), and Extreme Gradient Boosting (XGBoost). The models were trained on a comprehensive dataset spanning 62 years (1961–2022), incorporating meteorological, agricultural, and economic variables. Key predictors such as feed use, weight gain, and environmental factors were included. The dataset was carefully prepared to ensure robust model training and validation. Model performance was evaluated using multiple metrics, including the coefficient of determination (R^2), MAE, MSE, and RMSE. Results indicated that the Linear Regression model achieved the highest R^2 value and the lowest error rates among all algorithms. These findings underscore the potential of AI-based approaches to enhance decision-making and resource management in the poultry sector. With further integration of diverse data sources and advanced learning techniques, this framework can contribute to the development of more efficient, adaptive, and sustainable poultry production systems.

1. Introduction

Chicken meat has become one of the most widely consumed sources of animal protein worldwide due to its low cost, short production cycles, and high feed conversion efficiency (FAO 2013). The growing global population and shifting dietary habits are further increasing the demand for this economical and efficient source of protein. However, the increasing demand has made not only the quantity of production but also its sustainability a critical concern. Today, ensuring sustainability in the poultry sector necessitates optimizing production processes, utilizing resources efficiently, and responding promptly to market demands (Mottet and Tempio 2017). In this context, the accurate and timely prediction of chicken meat production is of critical importance for both internal decision-making within the industry and the development of national food security strategies.

Predictive modeling has become a fundamental component of modern agricultural practices, providing data-driven insights that significantly contribute to the optimization of production systems and improvement of resource management. Among these, machine learning (ML) offers significant advantages over traditional statistical methods by identifying complex patterns within large datasets (Ahmed and Hussain 2022). Machine learning (ML) has been successfully applied across various agricultural domains, including yield prediction, early detection of plant diseases, and livestock management (Gorczyca and Gebremedhin 2020). However, the use of these methods in the context of chicken meat production remains limited, presenting a significant opportunity for innovation. Chicken meat production is influenced by a wide range of interacting variables, including feed intake, growth rate, climatic conditions, and economic dynamics (Vandana et al. 2021; Istiak and Khaliduzzaman 2022; Liu et al. 2024). The nonlinear and interdependent relationships among these factors make accurate prediction particularly challenging. Nevertheless, studies that integrate meteorological and agricultural data have demonstrated that ML algorithms are highly suitable for addressing this complexity, offering a powerful alternative to traditional modeling techniques (Yıldız et al. 2024).

This study aims to develop prediction models for chicken meat production using four machine learning algorithms—Linear Regression (LR), Random Forest (RF), k-Nearest Neighbors (k-NN), and Extreme Gradient Boosting (XGBoost)—and to evaluate their predictive performance. The models were trained on a comprehensive dataset spanning 62 years (1961-2022) from Türkiye. This dataset includes not only agricultural variables such as chicken meat production and the number of live chickens, but also meteorological indicators such as average temperature and precipitation, as well as economic indicators such as producer prices and gross domestic product. The performance of

© Akdeniz University Faculty of Agriculture

the models was evaluated using widely accepted metrics, including the coefficient of determination (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). In this context, the study aims to contribute to sustainability goals in chicken meat production by leveraging historical and multidimensional data, and to provide a scientific foundation for the development of data-driven decision support systems.

2. Materials and Methods

2.1. Material

This study utilized a dataset comprised of 21 variables relevant to chicken meat production, encompassing agricultural, meteorological, demographic, and economic factors. The variables included chicken meat production (tonnes), live chicken count (1000 heads), chicken meat producer price ($ per tonne), population (units), rural population (units), urban population (units), agricultural area (1000 ha), pasture and meadow area (1000 ha), food price inflation (%), temperature (°C), rainfall (mm), gross national product (GNP million $), gross domestic product (GDP, million $), GDP per capita ($), barley production (tonnes), corn production (tonnes), barley price (tonnes $), corn price ($ per tonne), wheat price ($ per tonne), chicken import quantity (1000 heads), and chicken export quantity (1000 heads). The dataset spans the years 1961-2022 and was employed as the training set for developing predictive models.

Agricultural data, including chicken meat production, live chicken count, agricultural area, pasture and meadow area, barley production, and corn production, were retrieved from the Republic of Turkey Ministry of Agriculture and Forestry (TMAF 2024). Meteorological data, comprising temperature and rainfall, were sourced from the General Directorate of Meteorology under the Ministry of Environment, Urbanization, and Climate Change (GDM 2024). Demographic and economic data, such as chicken meat producer price, GNP, GDP, GDP per capita, barley and corn production and prices, wheat price, and chicken import and export quantities, were obtained from the United Nations Food and Agriculture Organization (FAO 2024a).

The statistical properties of these variables, including their mean, standard deviation, maximum, and minimum values, are summarized in Table 1, providing insights into the data distribution and variability within the dataset.

2.2. Data preparation

The dataset underwent a comprehensive preprocessing stage to ensure data quality and enhance model performance. Missing values were imputed using mean value calculations, while outliers were identified and removed based on predefined threshold values. To standardize the scale of all features and improve the accuracy of machine learning algorithms, feature normalization was applied. The data was then partitioned into training and test subsets, with 70% allocated to training and the remaining 30% reserved for testing.

2.3. Machine learning models

This study utilized four machine learning algorithms: Linear Regression (LR), Random Forest (RF), k-Nearest Neighbors (k-NN), and Extreme Gradient Boosting (XGBoost). To optimize the performance of these algorithms, hyperparameter tuning was conducted using the Grid Search method, which systematically evaluates combinations of hyperparameters to identify the optimal configuration for each model.

Table 1. Statistical properties for attributes

Attributes	Mean	Standard Deviation	Minimum	Maximum
Chicken meat production (tonnes)	709492	716027.9	60000	2417995
Live chicken count (1000 heads)	156150.4	122168.3	26116	391394
Chicken meat producer price ($ per tonne)	2022.367	327.8872	847.3	3271.8
Population (units)	56014.39	17363.83	28255	85279.55
Rural population (units)	21823.53	1349.229	19121.82	24723.16
Urban population (units)	33622.59	17580.66	9025.07	65453.23
Agricultural area (1000 ha)	38753.7	1176.671	36517	41223
Pasture and meadow area (1000 ha)	12415.62	1797.074	10000	14617
Food price inflation (%)	19.35789	12.1942	3.896166	77.86761
Temperature (°C)	12.05032	1.530993	9.58	15.1
Rainfall (mm)	606.1389	70.49718	460.69	793.8
GNP (million $)	349669.4	283621.1	904.19	941689.7
GDP (million $)	369093.9	294729.9	23609.87	957799
GDP per capita ($)	5286.164	3465.499	649.337	12507.8
Barley production (tonnes)	6258981	1868969	2900000	9551000
Corn production (tonnes)	2727806	1921058	800000	8500000
Barley price ($ per tonne)	202.8625	49.39973	100.1	370.3
Corn price ($ per tonne)	228.0594	44.37527	138	375.3
Wheat price ($ per tonne)	245.6938	55.0746	126.9	438.7
Chicken import quantity (1000 heads)	1495.148	1389.77	2	6150
Chicken export quantity (1000 heads)	8022.277	12665.12	1	62335

© Akdeniz University Faculty of Agriculture

2.4. Performance evaluation

The predictive performance of the models was assessed separately on the training and test datasets. Model predictions were compared to the actual values in the test set, and performance was evaluated using four key metrics: the Coefficient of Determination (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Square Error (RMSE).

Coefficient of Determination (R^2): This metric indicates the proportion of variance in the dependent variable explained by the independent variables. R^2 values range from 0 to 1, with values closer to 1 reflecting a better fit of the model to the data.

Mean Absolute Error (MAE): MAE represents the average magnitude of absolute differences between predicted and actual values, offering a measure of the accuracy of predictions. A lower MAE indicates superior model performance.

Mean Squared Error (MSE): MSE calculates the average of the squared differences between predicted and actual values, penalizing larger errors more significantly. It provides a measure of the model's overall prediction error.

Root Mean Square Error (RMSE): RMSE, the square root of MSE, directly reflects the magnitude of prediction errors in the original data units, offering a clear and interpretable metric of accuracy.

The use of these diverse metrics allows for a comprehensive evaluation of model performance, capturing various aspects of prediction quality. The formulas used to calculate these metrics are outlined below to ensure clarity and reproducibility in the evaluation process.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (4)$$

$(n) = Total\ number\ of\ observations$; $(yi) = (i) - th\ observation\ of\ true\ values$; $(\hat{y}_i) = (i) - th\ observation\ of\ predicted\ values$

2.5. Software

The analyses were conducted using the Python programming language, version 3.12.2 (Python Software Foundation 2024), leveraging libraries such as Pandas (1.3.0) for data manipulation, Numpy (version 1.21.0) for numerical operations, Matplotlib (version 3.4.2) for data visualization, and Scipy (version 1.10.0) for scientific computations (Yıldız et al. 2024).

3. Results and Discussion

The forecast models developed using LR, RF, k-NN and XGBoost algorithms were evaluated using agricultural, meteorological and economic data affecting chicken meat production. As a result of the analyses, the forecast accuracy of the models was compared with R^2, MAE, MSE and RMSE metrics. In the obtained findings, LR provided the highest accuracy with R^2= 0.9907 and the lowest error values (MAE= 51324.28; MSE= 5932206846.61; RMSE= 77020.82). RF showed high accuracy with R^2= 0.9855 but its error values (MAE= 63976.58; MSE= 9284599433.44; RMSE= 96356.63) were higher than LR. XGBoost provided similar accuracy with R^2= 0.9836 but fell behind RF in terms of performance with MAE= 60,364.73, MSE= 10468242376.31 and RMSE= 102,314.43. k-NN had the lowest accuracy rate with R^2= 0.9585 and showed the highest error values (MAE= 100259.51; MSE= 26524968103.15; RMSE= 162864.88). These results show that LR is superior to other algorithms in terms of error rates and accuracy (Table 2).

The agreement analysis between the predicted and actual values reveals that the estimates generated by the LR algorithm closely align with the y= x line at a 45-degree angle, as illustrated in the agreement graph (Figure 1). This strong alignment highlights the high accuracy of the LR predictions and the model's excellent fit to the dataset. Supported by low error metrics, including MAE, MSE, and RMSE, these results confirm that the LR algorithm effectively captures the underlying relationships within the data and is the most suitable model for this specific dataset.

In this study, the Linear Regression (LR) algorithm stood out with lower error rates and higher accuracy compared to the other models, which is consistent with findings from similar studies in the literature. For instance, in a study conducted by Yıldız et al. (2024) to predict honey production in Türkiye, meteorological and agricultural data were jointly evaluated, and the LR algorithm emerged as the most successful model with an R^2 value of 0.97. This result indicates that the LR algorithm can achieve high accuracy even in multivariate data structures across different production systems.

Studies conducted specifically in the poultry sector emphasize the effectiveness of machine learning algorithms in handling complex multivariate biological datasets. Yıldız et al. (2025) developed models integrating internal and external quality parameters to predict egg quality in Japanese quails, where the Random Forest and Gradient Boosting algorithms stood out with accuracy rates exceeding 97%. Similarly, in a study conducted by Sehirli and Arslan (2022), the aim was to classify egg quality without relying on Haugh Unit (HU) data; the logistic regression algorithm was identified as the most successful model, achieving an accuracy of 98.6% and an MCC value of 0.96. In the same study, the Random Forest algorithm was reported to exhibit high performance in the classification of egg quality based on HU data.

Table 2. Performance metrics of algorithms on the test set

Algorithms	R^2	MAE	MSE	RMSE
LR	0.9907	51324.28	5932206846.61	77020.82
RF	0.9855	63976.58	9284599433.44	96356.63
k-NN	0.9585	100259.51	26524968103.15	162864.88
XGBoost	0.9836	60364.73	10468242376.31	102314.43

R^2, coefficient of determination; MAE, Mean Absolute Error; MSE, Mean Squared Error; RMSE, Root Mean Square Error.

© Akdeniz University Faculty of Agriculture

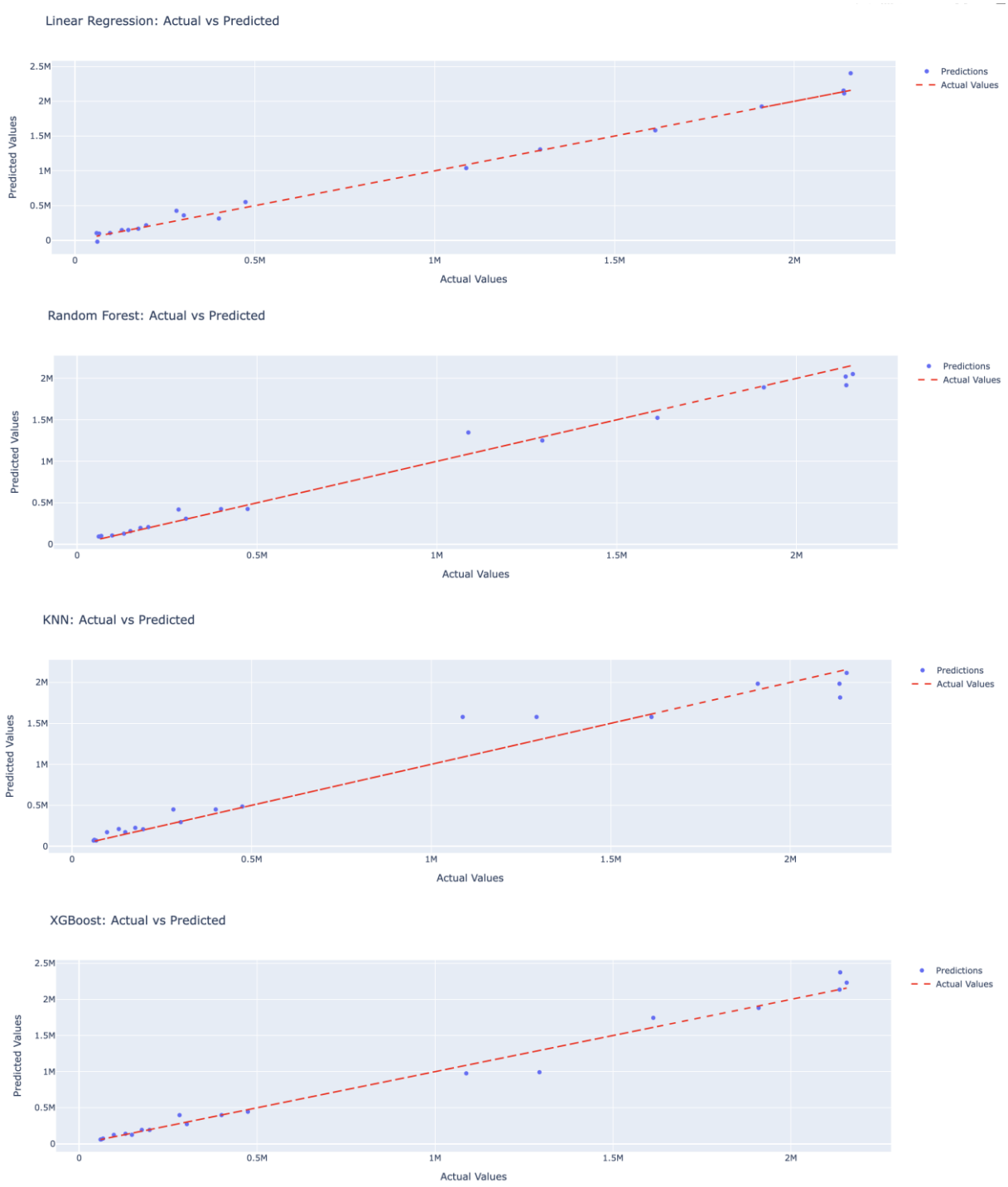


Figure 1. Analysis of the alignment between predicted and actual values.

The widespread integration of machine learning algorithms into animal production systems provides a broader context in which to interpret the results of the current study. For instance, Alsahaf et al. (2018) employed four different algorithms—Random Forest (RF), Extremely Randomized Trees (ET), Gradient Boosting Machine (GBM), and XGBoost (XGB)—to estimate the age at which pigs reached 120 kg slaughter weight, concluding that GBM and XGB yielded the lowest error metrics. Similarly, Srivastava et al. (2021) assessed the predictive capabilities of RF, XGB, and Support Vector Machine (SVM) in estimating carcass weight (CWT), marbling score (MS), backfat thickness (BFT), and eye muscle area (EMA) in Hanwoo cattle. Their findings indicated that EGB performed best for CWT and MS, whereas SVM achieved the lowest mean squared error for BFT and EMA.

In the literature, it has been shown that these algorithms have versatile application areas such as not only production estimation, but also optimization of operational processes and analysis of consumer behavior. Lyu et al. (2023) showed that real-time predictions provide a practical solution by optimizing

© Akdeniz University Faculty of Agriculture

weighing methods in broiler production with a machine learning-based model. Chiras et al. (2023) emphasized the importance of machine learning methods in supporting strategic decision processes in the food sector by analyzing consumer behavior.

The feature importance analysis used in this research was conducted to determine which variables have a greater impact in the prediction models. In particular, the feature importance scores calculated using the RF algorithm revealed the main factors affecting chicken meat production. According to the results of this analysis, Gross Domestic Product (GDP), urban population and national income per capita (GDP per capita) stood out as the most important variables (Figure 2). While the increase in GDP provides economic stability and production capacity, thus enabling more investment in the sector, the increase in per capita income enables consumers to turn to more and higher quality protein sources. In addition, the growth of the urban population has increased the demand for chicken products by increasing fast-paced consumption habits. These findings highlight the critical impact of economic and demographic factors on chicken meat production (FAO 2013; FAO 2024b).

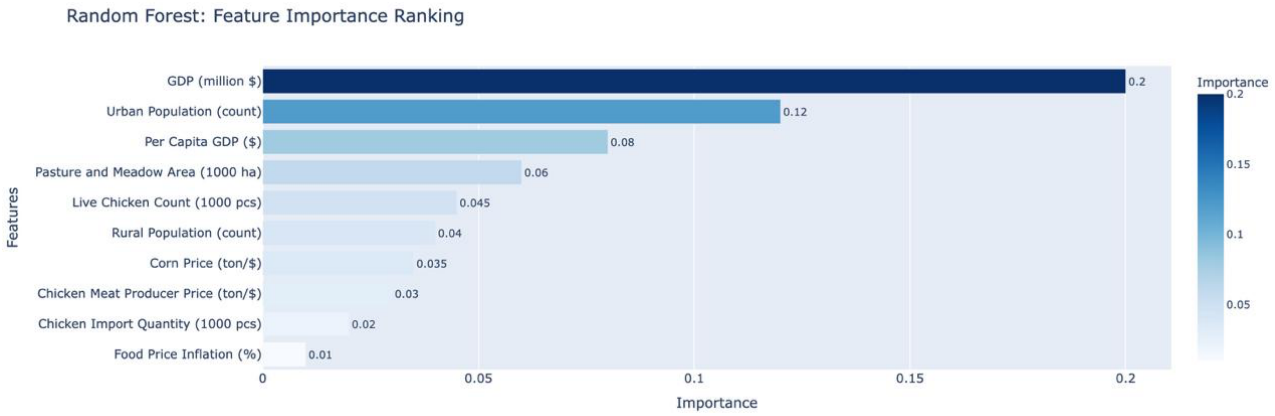


Figure 2. Analysis of feature importance for chicken meat production prediction.

4. Conclusion

In this study, which examined the applicability of machine learning models to predict chicken meat production, different models were developed using LR, RF and k-NN algorithms and comparative performance analysis was performed. The results revealed that the LR algorithm exhibited the best performance with higher accuracy (R^2) and lower error rates (MAE, MSE, RMSE) compared to other methods. These findings show that data-driven prediction models are a powerful tool for optimizing production processes, improving resource utilization and supporting decision-making processes in the poultry sector. The 62-year meteorological, agricultural and economic data used in the study increased the accuracy of machine learning models and demonstrated the feasibility of long-term predictions. However, the sectoral use of such models is directly related to the quality and diversity of available data sources. Inclusion of larger data sets and different attributes in the model can increase the accuracy and generalizability of future prediction models. The findings emphasize the importance of machine learning applications in the poultry sector. In industry, integrating such models into decision support systems can provide significant benefits, both economically and environmentally.

References

Ahmed MU, Hussain I (2022) Prediction of wheat production using machine learning algorithms in northern areas of pakistan. Telecommunications Policy 46: 102370.

Alsahaf A, Azzopardi G, Ducro B, Veerkamp RF, Petkov N (2018) Predicting slaughter weight in pigs with regression tree ensembles. Frontiers in Artificial Intelligence and Applications 310: 1-9.

Chiras D, Stamatopoulou M, Paraskevis N, Moustakidis S, Tzimitra-Kalogianni I, Kokkotis C (2023) Explainable Machine Learning Models for Identification of Food-Related Lifestyle Factors in Chicken Meat Consumption Case in Northern Greece. BioMedInformatics 3: 817-828.

GDM (2024) General Directorate of Meteorology, official climate statistics of Türkiye. https://www.mgm.gov.tr/veridegerlendirme. Accessed 4 November, 2024.

FAO (2013) Poultry development review. https://www.fao.org/4/i3531e/i3531e.pdf. Accessed 4 November, 2024.

FAO (2024a) FAOSTAT, demographic and economic and political stability data of Türkiye. https://www.fao.org/faostat/en/#country/223. Accessed 4 November, 2024.

FAO (2024b) Global and regional food consumption patterns and trends. Food and Agriculture Organization of the United Nations. https://www.fao.org/4/ac911e/ac911e05.htm. Accessed 12 November, 2024.

Gorczyca MT, Gebremedhin KG (2020) Ranking of environmental heat stressors for dairy cows using machine learning algorithms. Computers and Electronics in Agriculture 168: 105124.

Istiak MS, Khaliduzzaman A (2022) Poultry and egg production: An overview. In: Khaliduzzaman A (eds), Informatics in Poultry Production. Springer, Singapore, pp. 3-12.

Liu L, Wang Z, Wei B, Wang L, Zhang Q, Si X, Huang Y, Zhang H, Chen W (2024) Replacement of corn with different levels of wheat impacted the growth performance, intestinal development, and cecal microbiota of broilers. Animals 14: 1536.

Lyu P, Min J, Song J (2023) Application of machine learning algorithms for on-farm monitoring and prediction of broilers' live weight: A quantitative study based on body weight data. Agriculture 13: 2193.

Mottet A, Tempio G (2017) Global poultry production: current state and future outlook and challenges. World's Poultry Science Journal 73: 245-256.

© Akdeniz University Faculty of Agriculture

Python Software Foundation (2024) The Python language version 3.12.2.

Srivastava S, Lopez BI, Kumar H, Jang M, Chai HH, Park W, Park JE, Lim D (2021) Prediction of Hanwoo cattle phenotypes from genotypes using machine learning methods. Animals 11: 2066.

Sehirli E, Arslan K (2022) An application for the classification of egg quality and haugh unit based on characteristic egg features using machine learning models. Expert Systems with Applications 205: 117692.

TMAF (2024) Turkish Ministry Agriculture and Forestry, agricultural data information center. https://www.tarimorman.gov.tr/Konular/. Accessed 4 November, 2024.

Vandana GD, Sejian V, Lees AM, Pragna P, Silpa MV, Maloney SK (2021) Heat stress and poultry production: impact and amelioration. International Journal of Biometeorology 65: 163-179.

Yıldız Bİ, Eskioğlu K, Karabağ K (2024) Developing a machine learning prediction model for honey production. Mediterranean Agricultural Sciences 37: 105-110.

Yıldız Bİ, Eskioğlu K, Özdemir D, Akşit M (2025) Predicting quail egg quality using machine learning algorithms. Brazilian Journal of Poultry Science 27: eRBCA-2024.

© Akdeniz University Faculty of Agriculture