International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Application of Artificial Neural Network Model for Predicting the Effect of Environmental Factors on Poultry Egg Production Rate

Okwori Anthony Okpe[1], Anthony Obogo Otiko[2], Kolawole Olamide Michael[3], Ekoro, Ekoro Igo (PhD)[4], Adoga Stephen[1]

[1] *Computer Science, Federal University Wukari, Wukari, Nigeria*
[2] *Computer Science, University of cross-river state, Calabar, Nigeria*
[3] *Computer Science, Arthur Jarvis University, Calabar, Nigeria*
[4] *Computer Science, Education, University of Education and Entrepreneurship, Akamkpa, Nigeria*

ABSTRACT

In Nigeria, the egg production rate in poultry farms is highly affected by various environmental factors such as the quality of air, quality of water, humid nature of the environment among others. To improve egg productivity, there is an urgent need for proper understanding of the effect of these various environmental factors and proactively control them. This paper aims at developing an Artificial Neural Network (ANN) model for predicting the effect of environmental factors on poultry egg production rates. The ANN model was developed using an egg production rate prediction dataset obtained from Kaggle machine learning repository. This model was achieved through three distinct developmental phases including data preprocessing, model building and model evaluation respectively. The ANN model was implemented using Python programming language including TensorFlow and Scikit-Learn APIs to meet the complex requirements of artificial neural networks and machine learning in general. Some of the Python modules such as NumPy and pandas enable the processing and integration of datasets, TensorFlow is utilized for complex numerical computations while Matplotlib for visualization. The model was evaluated using confusion matrix and its related matrixes such as accuracy, precision, recall, f1-score as well as area under the receiver operating characteristic cure (AUC) and found that the ANN is highly efficient in predicting the effect of the various environmental factors in egg production rate in poultry with its prediction accuracy of 85% and AUC of 96% respectively.

Keywords: Egg, Environmental Factors, Machine Learning, Scikit-learn, Neural Network

1. Introduction

Poultry farming is a significant component of the global agricultural industry, with eggs being one of the primary products. The ability to accurately predict egg production rates is crucial for efficient farm management, optimizing resources, and meeting market demands. Traditionally, predicting egg production has relied on manual observation and basic statistical methods, which are often time-consuming and prone to errors. With the advancement of technology, particularly in artificial intelligence (AI) and machine learning (ML), there is a growing interest in applying these technologies to improve agricultural processes. The application of Artificial Intelligence (AI) is revolutionizing several sectors, including agriculture, by providing novel approaches to streamline operations and improve output [1] Within the domain of chicken egg production, artificial intelligence (AI), offers substantial prospects to tackle current obstacles and enhance productivity as it encompasses technologies such as machine learning, neural networks, and deep learning [2] models that can imitate specific cognitive processes through a well-defined pattern. Furthermore, in the field of chicken egg production, artificial intelligence (AI) has the potential to improve forecasting, disease detection, and general management procedures [3]. The successful deployment of artificial intelligence depends on the incorporation of many data sources, encompassing environmental factors (such as temperature and humidity), biological data (including rates of egg production and intake of feed), and operational data (such as practices of farm management). Artificial neural network algorithms enhance prediction accuracy and operational efficiency by acquiring intimate knowledge of intricate interactions inside extensive datasets. Artificial neural network models facilitates data-driven decision-making, therefore diminishing dependence on intuition and enhancing the dependability of management decisions [4]. In the context of poultry farming, artificial neural network model analyzes vast amounts of data, such as environmental factors, feed intake, and historical production records, to provide more accurate predictions of egg production rates. The need for accurate and reliable prediction models in poultry farming is becoming increasingly important due to the rising global demand for eggs. The global poultry industry plays a critical role in food security by supplying a significant portion of the world's protein through egg production. However, one of the persistent challenges in poultry farming is the accurate prediction of egg production rates. Traditional methods, which often rely on manual observation and basic statistical techniques, are limited in their ability to account for the numerous variables that influence egg production, such as feed quality, hen health, genetic factors and environmental conditions and are often time-consuming and prone to errors [5] leading to inefficiencies and reduced profitability. Inaccurate predictions can result in either overproduction, leading to wasted resources and financial losses, or underproduction, failing to meet market demands and losing potential revenue. This highlights the need for more sophisticated and accurate predictive tools that can handle the complexity and variability of egg production. Although some traditional machine learning models has been used however these models often face issues with model complexity, accuracy, and adaptability to new data. Efficient implementation of artificial neural network model in poultry management is necessitated as it offers significant improvements using it's multiple layers and sophisticated architectures to capture and analysis complex patterns and relationships within extensive datasets to enhance forecasting precision and flexibility and effectively model intricate interactions between the various environmental factors thereby bridging the gaps between prediction accuracy and model adaptability, leading to more efficient and cost-effective solutions in poultry management. This research aims at developing Artificial Neural Network model for predicting egg production rates in poultry farm using environmental factors in order to enhance the efficiency of poultry farm operations and contribute to sustainable farming practices, thereby supporting better farm management and contributing to the overall efficiency and sustainability of poultry farming.

2. Related Literature

Recent studies have demonstrated the potential of deep learning models in agricultural applications. Using multiple regression, Bayesian networks, and artificial neural networks, [6] forecasted the total egg production in European quails based on previously observed traits. The motivation for their work emanated from the necessity to make efficient management decisions in forecasting total egg production (TEP) in commercial plants. Their work aimed at examining multiple modelling approaches for predicting Total Egg Production (TEP) in meat type quails based on phenotypes including weight, weight gain, egg production, and egg quality measures. The researchers implemented multiple linear regression and artificial neural network (ANN) with Bayesian network (BN) and a stepwise approach used for variables selection. The author's model was developed to predict TEP using phenotypic data on 30 variables from two lines of quail, each consisting of 180 and 205 measurable points. Adequate prediction accuracies were achieved only when the model incorporates partial data of egg production. In related development, [7] utilized the Random Forest classification algorithm to forecast the occurrence of daily oviposition events in broiler breeders that were equipped with a precision feeding system. Their target was to tackle the difficulties associated with forecasting egg-laying events in group-housed chicken, where the variability caused by hormones and the environment complicated the prediction of follicle maturation and egg production. Data on the eating activity and real-time body weight of 202 free-run Ross and 708 broiler breeder hens were gathered from week 21 to 55 of their research. The data was analyzed to derive 34 characteristics associated with feeding behavior, which were then merged with daily oviposition data. Applying recursive feature elimination with 5-fold cross-validation resulted in the selection of 28 features for constructing the Random Forest model. The algorithm had a remarkable accuracy rate of almost 85% in forecasting the likelihood of a hen depositing an egg on a certain day. A recent study conducted by [8] investigated the application of artificial neural networks (ANN) in forecasting the likelihood of oviposition occurrences in broiler breeder hens that are fed with precision. The study emphasizes the need of precisely determining the daily incidence of egg-laying in individual birds, a critical aspect of efficient bird control, particularly in identifying non-laying birds that may be impacted by inadequate nutrition or illnesses. The researchers constructed a machine learning model using a precision feeding system that documents specific feeding behaviors, with the objective of forecasting daily occurrences of egg-laying. Their work accurately forecasts oviposition events at precise time intervals within a day. According to the authors reports, the feed-forward artificial neural network (ANN) model exhibited exceptional classification performance, as evidenced by an area under the receiver operating characteristic (ROC) curve of 0.94. [9] employed a multi-trait animal model and a selection methodology to assess and enhance the genetic capacity of body weights and egg production characteristics in Thai Native Synthetic Chickens. The primary objective of their study was to improve the growth and egg production of purebred local chickens by overcoming their natural constraints. The analysis examined data spanning from 2015 to 2020, encompassing a range of body weight and egg production measurements. The findings indicated substantial heritability estimates for both growth and egg production characteristics, showing favorable genetic associations between body weight at specific gestational ages and egg production attributes, especially at 240 days. [10] evaluated the effects of environmental control strategies (ECSs) on egg production in cage-free farm settings. The study employed sophisticated machine learning methods, namely a Random Forest (RF) model, to forecast daily variations in hen-day egg production (HDEP) by considering factors such as temperature, air quality, and hygrothermal conditions. Upon doing a variable importance study, the researchers have verified that temperature emerged as the most influential predictor, with hen's age and relative humidity following suit. The Random Forest (RF) model exhibited very good accuracy, with R^2 values of 0.94 and 0.78 for the training and testing datasets, respectively. The scenario analysis of their findings demonstrated that a 5% rise in temperature could have an adverse impact on egg production, underscoring the crucial need of environmental surveillance in cage-free systems and overall poultry farm management system.

3. Research Methodology

The process for achieving the proposed Artificial Neural Network model has been stratified into three major phases including dataset description and preprocessing, model development, and performance evaluation. Every phase has a crucial function in guaranteeing the efficiency of the target egg production rate prediction system.

3.1 Dataset Description and Preprocessing

The dataset for predicting egg production rates in poultry was sourced from the Kaggle machine learning repository. The dataset contains comprehensive data related to various environmental factors influencing egg production in poultry farms. This dataset includes five features with 1500 instances that are critical to understanding and forecasting the rate of egg production. The target class in the dataset has three class labels of which the high label having 501 records, medium with 500 records, and the low label with 499 records. The distribution shows that the dataset is highly balanced. In dataset preprocessing phase, missing values were addressed using mean value imputation, feature normalization, removal of noise or outliers, feature scaling, label encoding and relabeling of the target class was effectively carried out. The dataset was then partitioned into two subsets: a training set and a test set, utilizing an 80:20% division. The training set, constituting of 80% of the total data (equivalent to 1200 from 1500 total instances), was employed to train the ANN algorithm, enabling the model to discern patterns and relationships within the data. The test set, representing the remaining 20% (making of 300 instances), was solely utilized to evaluate the model's generalization capacity and assess its performance on previously unseen data. This methodology guarantees that the trained model is not only optimized for the training data but also proficient in making accurate predictions on novel, real-world inputs.

3.2. Model Development

Informed by research into the nervous system and the brain, artificial neural network (ANN) is a sophisticated technology developed to represent intricate connections in data, such as forecasting egg production rates in poultry. In order to replicate the form and function of real brain networks, these networks employ a simplified set of rules. In particular, Artificial Neural Network (ANN) replicates the electrical activity of the brain by linking several layers of perceptron-like processing units, which correspond to neurons. The proposed ANN for predicting egg production rate utilizes perceptron's arranged in many layers, consisting of an input layer, many hidden layers, and an output layer as show on Figure 3.4. The output generated by each perceptron in a given layer is used as the input for the perceptron's in the following layer. The layered nature of the network allows it to effectively represent intricate and non-linear relationships present in the data. The input values received by each perceptron in the network are multiplied by the associated connection weights, represented as W_m, which replicate the synaptic connections observed in the brain. The weights in question serve as indicators of the magnitude of the link, and their tuning during the training phase is crucial for the acquisition of knowledge by the model. The sum of these weight-adjusted inputs corresponds to the combined input to a perceptron as shown in equation 3.1.

$$y = \sum w_{ij} x_i \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots 3.1$$

The summing of these values yields a singular input value for the perceptron, which subsequently undergoes a change implemented by a transfer function, often a non-linear function known as the activation function. The inclusion of an activation function in the model induces non-linearity, therefore facilitating the network's ability to acquire knowledge and represent intricate patterns within the data, such as the variables that impact egg production rates most.

The iterative process of input collection, transformation, and output generation occurs throughout the layers of the network. The ANN improves its predictions by iteratively refining them layer by layer until it generates the final output, denoted as O_n which in this instance represents the projected rate of egg production. The network is trained by iteratively calibrating the weights according to the discrepancy between the expected and actual values, hence progressively enhancing its precision in forecasting egg production rates.

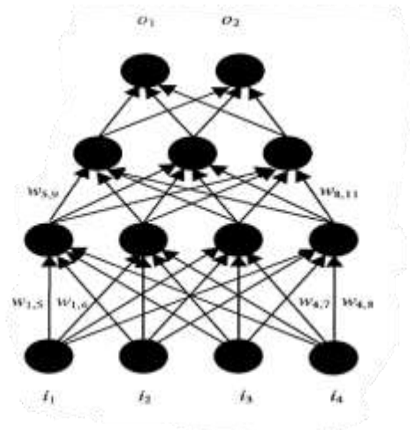

Figure 3.1. Artificial Neural Network Architecture.

Algorithm 3.1. ANN Algorithm

Step 1: *pass the input with some weight to the input layer* $(x^1 x^2, \dots \dots x^6)$

Step 2: *Connect all the inputs to each neuron*

Step 3: *perform computation at the hidden layers (Eq 3.8)*

Step 4: *sum all input with their weight*

Step 5: *Get bias*

Step 6: *Get the threshold unit*

Step 7: *Repeat step 3-6 for each of the hidden layers*

Step 8: *Pass the result to an output layer*

To optimize model performance, various parameters such as Loss function, Max epochs, Optimizer, Learning Rate, Activation function and Batch Size were turned as shown in table 1. These parameter values determine the model's training and optimization properties, which in turn affect its capacity to learn and generalize from the input data.

Table 1. Egg Production Rate Model Parameter Setting.

Parameter	**Value**
Loss	Binary-Crossentropy
Max epochs	50
Optimizer	Adam
Learning Rate	0.001
Activation	Rectified Linear Unit (ReLU)
Batch Size	32

3.3. Data Visualizations

In an effort to analyze the dataset, various data visualization mechanisms were employed to show the graphical representation of the basic environmental factors. Figure 3.2 presents a histogram depicting the distribution of three principal variables: AQI (Air Quality Index), WQI (Water Quality Index), and Humidex (Humidity Index) based on feature importance.

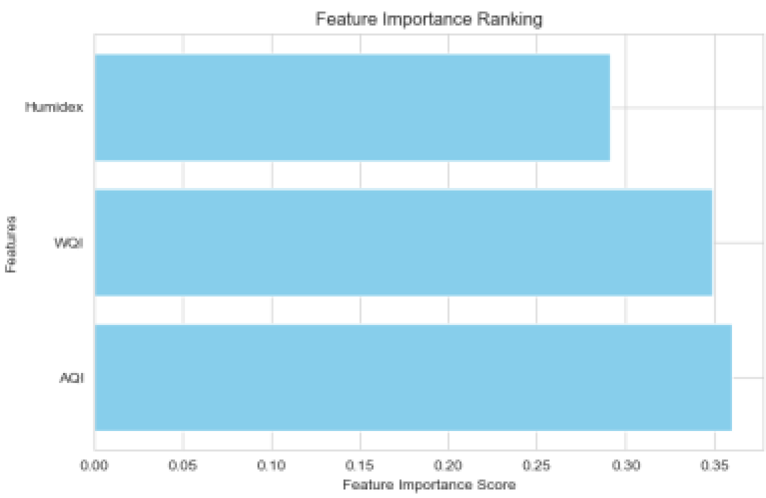

Figure 3.2. Feature Importance Ranking

Figure 3.3 presents a Kernel distribution Estimation (KDE) graph that illustrates the probability distribution of three significant environmental variables: Air Quality Index (AQI), Water Quality Index (WQI), and Humidex. The graph indicates that the AQI (depicted in blue) exhibits a wide distribution, with values ranging from nearly zero to significantly above 500. This signifies that air quality exhibits considerable variability within the dataset. The Water Quality Index (WQI), depicted in orange, exhibits a concentrated distribution, peaking within a moderate range, signifying that water quality is predominantly constant, with the majority of values aggregated between 0 and 100. This WQI is advantageous, as pristine water is essential for poultry health and reliable egg production. The Humidex (shown in green) has a pronounced, narrow peak, implying that humidity levels are closely clustered within a specific range, with little severe fluctuations. This distribution signifies a pleasant environment for chickens, as elevated humidity may result in heat stress and diminished egg production rates.

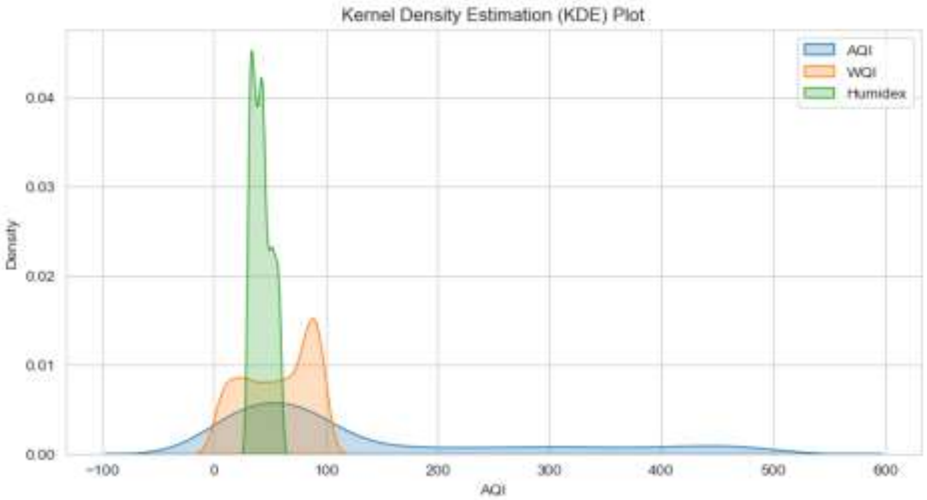

Figure 3.3. Kernel Density Estimation for Egg Production Rate.

Another data visualization tool used in this research is the bar chart in figure 3.4 that illustrates the class distribution for egg production with "0" (blue) denoting the high class and "1" (red) representing the low class. Although there were three target class labels in the original dataset, the medium class label was distributed among the high and low class label through dataset relabeling to convert it to two class labels. Thus, the y-axis indicates the frequency of occurrences in each class, whilst the x-axis signifies the two production categories. The chart indicates that both classes exhibit about equal distribution, with approximately 750 cases in each category. This equitable distribution indicates that the dataset is suitably organized for training machine learning models, as it mitigates the danger of class imbalance, which could result in biased predictions.

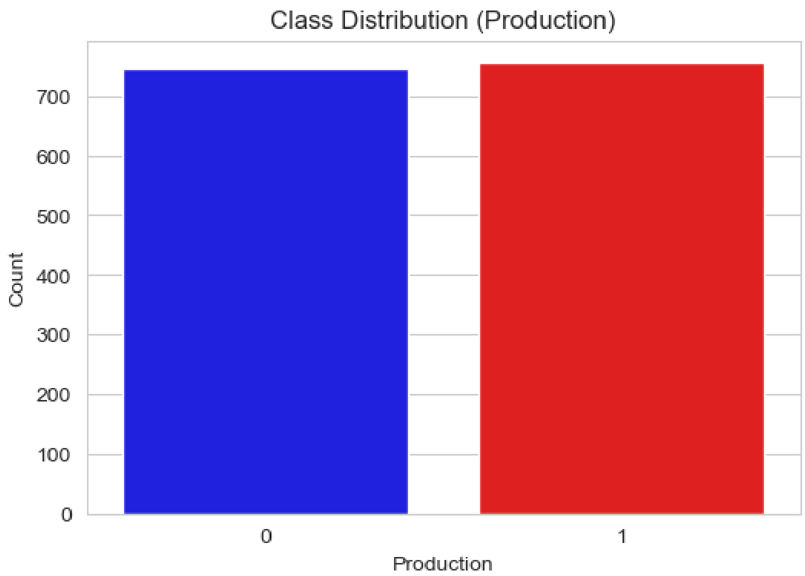

Figure 3.4. High and Low Prediction Rate Labels.

3.4. Performance Evaluation

The artificial neural network model developed was evaluated using some standard model evaluation metrics such as accuracy, precision, recall, F1-score and AUC-curve. These metrics values were computed using confusion metric and it True Positive (TP), False Negative (FN), False Positive (FP) and True Negative (TN) prediction values. The effective values of these metrics provide a comprehensive understanding of the model’s performance across different aspects of classification.

4. Results and Discussion

The performance evaluation of the artificial neural network-based egg production model is computed and presented in this section through confusion matrix that shows the values of the TP, TN, FP and FN prediction rates after successful model testing. The test dataset comprises of two class labels in the target class: 0 (low production) and 1 (high production), each containing 150 items. The model correctly identified 150 high-production cases (true positives) and 105 low-production cases (true negatives). However, it misclassified 45 low-production cases as high (false positives), while no high-production cases were incorrectly predicted as low (false negatives),

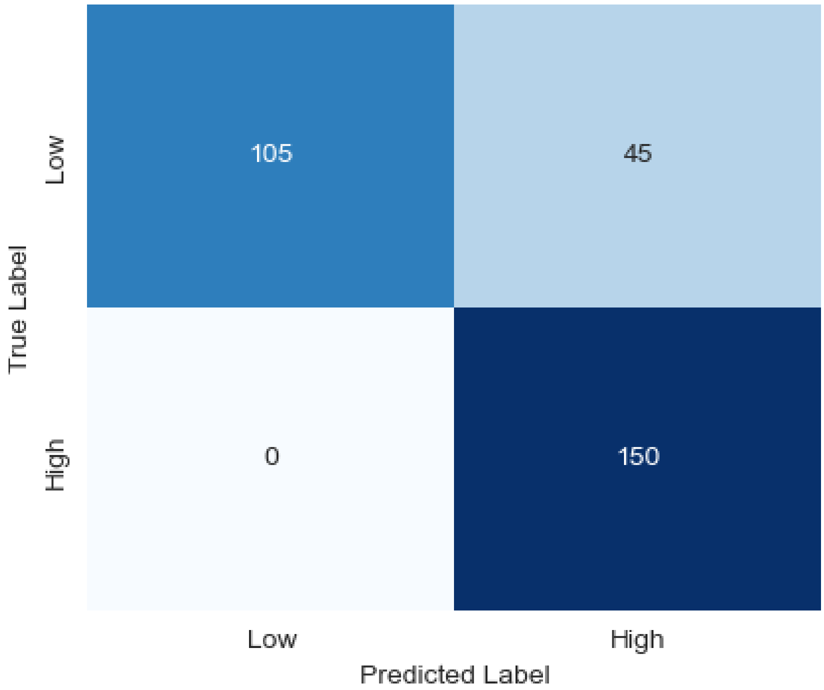

Figure 4.1. Confusion Matrix.

The confusion matrix values were used to compute the precision, recall, F1-score, and accuracy of the developed model to offers insights into the model's performance across many assessment measures. In class 0 (low production), the model attained a precision of 1.00, indicating that all predictions categorized as low output were accurate. The recall for class 0 is 0.70, signifying that 30% of actual low production events were inaccurately labeled as high output. The F1-score for class 0 is 0.82, indicating a balance between precision and recall. In contrast, for class 1 (high production), the recall is 1.00, indicating that the model accurately detected all high-production occurrences without any false negatives. Nonetheless, the precision for class 1 is 0.77, signifying that 23% of the forecasts categorized as high output were, in fact, low production instances. The F1-score for class 1 is 0.87, indicating a robust performance in identifying high-production occurrences. The macro average precision, recall, and F1-score are 0.88, 0.85, and 0.85, respectively, indicating that the model exhibits uniform performance across both classes. The weighted average values are equivalent, as anticipated due to the balanced dataset.

Classification Report:

	precision	recall	f1-score
0	1.00	0.70	0.82
1	0.77	1.00	0.87
accuracy			0.85
macro avg	0.88	0.85	0.85
weighted avg	0.88	0.85	0.85

Figure 4.2. classification report.

To determine the model's convergence, the Receiver Operating characteristic (ROC) Curve was generated. The ROC curve illustrates the model's classification efficacy by graphing the true positive rate (sensitivity) against the false positive rate (specificity) across various threshold levels. The Area Under the Curve (AUC) is 0.96, signifying exceptional classification performance, as it is nearly 1. The curve ascends steeply toward the top-left corner, indicating that the model attains a high true positive rate while preserving a low false positive rate. The model accurately differentiates between elevated and diminished egg production rates. The lack of a gradual slope or notable deviation from the top-left quadrant indicates that the model has attained steady learning and convergence, hence reducing misclassification errors. Moreover, the elevated AUC value indicates that the model generalizes effectively and exhibits minimal overfitting. The AUC of 0.96, equally shows that the model has robust discriminative capability, effectively distinguishing across production groups. The smooth curve signifies that the decision boundary is effectively optimized, exhibiting little bias-variance trade-off complications.

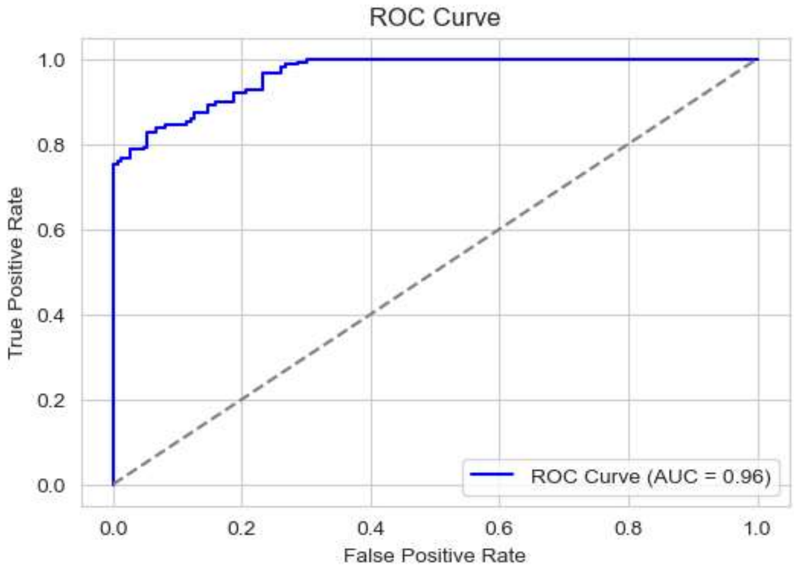

Figure 4.3. ROC Curve Report.

Conclusion and Recommendation

In this paper, the effect of Environmental Factors on Poultry Egg Production Rate was predicted using Artificial Neural Network (ANN) Model and environmental based egg production rate dataset obtained from Kaggle machine learning repository through an empirical experimental setup. The ANN model demonstrated strong performance in identifying high-production cases with an overall accuracy of 85%. It equally shows a high **convergence rate** as evident in its steep ROC curve and AUC value of 96%. Despite this strong performance, the model struggles moderately with precision in predicting low-production cases hence require further optimization to reduce false positives, which could be achieved through decision threshold adjustments, cost-sensitive learning as well as further fine-tuning of hyperparameters.

Conflicts of Interest

The authors declare no conflicts of interest

References

Patel, H., Samad, A., Hamza, M., Muazzam, A., & Harahap, M. K. (2023). Role of artificial intelligence in livestock and poultry farming. Sinkron: jurnal dan penelitian teknik informatika, 6(4), 2425-2429.

Siemens, G., Marmolejo-Ramos, F., Gabriel, F., Medeiros, K., Marrone, R., Joksimovic, S., & de Laat, M. (2022). Human and artificial cognition. Computers and Education: Artificial Intelligence, 3, 100107.

Soori, M., Arezoo, B., & Dastres, R. (2023). Artificial intelligence, machine learning and deep learning in advanced robotics, a review. Cognitive Robotics, 3, 54-70.

Depuru, B. K., Putsala, S., & Mishra, P. (2024). Automating poultry farm management with artificial intelligence: Real-time detection and tracking of broiler chickens for enhanced and efficient health monitoring. Tropical Animal Health and Production, 56(2), 75.

Ahmad, S., Mahmud, A., & Adnan, N. (2018). *An Overview of Poultry Production and its Prediction Methods*. Journal of Poultry Science, 55(4), 231-239.

Felipe, V. P., Silva, M. A., Valente, B. D., & Rosa, G. J. (2015). Using multiple regression, Bayesian networks and artificial neural networks for prediction of total egg production in European quails based on earlier expressed phenotypes. Poultry science, 94(4), 772-780.

You, J., van der Klein, S. A., Lou, E., & Zuidhof, M. J. (2020). Application of random forest classification to predict daily oviposition events in broiler breeders fed by precision feeding system. Computers and Electronics in Agriculture, 175, 105526.

You, J., Lou, E., Afrouziyeh, M., Zukiwsky, N. M., & Zuidhof, M. J. (2021). Using an artificial neural network to predict the probability of oviposition events of precision-fed broiler breeder hens. Poultry Science, 100(8), 101187.

Chomchuen, K., Tuntiyasawasdikul, V., Chankitisakul, V., & Boonkum, W. (2022). Genetic evaluation of body weights and egg production traits using a multi-trait animal model and selection index in Thai native synthetic chickens (Kaimook e-san2). Animals, 12(3), 335.-

Gonzalez-Mora, A. F., Rousseau, A. N., Larios, A. D., Godbout, S., & Fournel, S. (2022). Assessing environmental control strategies in cage-free aviary housing systems: Egg production analysis and Random Forest modeling. Computers and Electronics in Agriculture, 196, 106854.