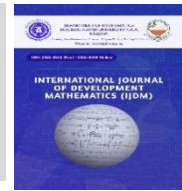




INTERNATIONAL JOURNAL OF DEVELOPMENT MATHEMATICS

ISSN: 3026-8656 (Print) | 3026-8699 (Online)

journal homepage: <https://ijdm.org.ng/index.php/Journals>



Contagious Bovine Pleuropneumonia Prediction Using Random Forest Technique

Musa A. Girei^{*} and Bala P. Bulus^a

^aDepartment of Computer Science, Faculty of Physical Sciences, Modibbo Adama University, Yola Nigeria

ARTICLE INFO

Article history:

Received 1 September 2024

Received in revised form 15 November 2024

Accepted 28 November 2024

Keywords:

Contagious bovine pleuropneumonia, economy, livestock, random forest

MSC 2020 Subject classification:

68T45

ABSTRACT

Contagious Bovine Pleuropneumonia (CBPP) is a highly infectious disease that poses a significant threat to cattle populations, particularly in sub-Saharan Africa, where livestock is a critical economic and food security resource. Early detection and intervention are essential to mitigate its impact on cattle productivity and rural livelihoods. This study aims to develop a predictive model for CBPP using the Random Forest (RF) algorithm, a robust and scalable machine learning technique renowned for its accuracy and ability to handle complex datasets. The research utilized a dataset of 847 samples collected from veterinary clinics in Yola, Adamawa State, and online sources. Key symptoms of CBPP, such as respiratory distress, nasal discharge, fever, and pleuritis, served as input features for the model. Preprocessing steps, including outlier detection, handling missing values, and data normalization, were applied to ensure the dataset's quality and reliability. The RF algorithm was trained and validated to classify CBPP cases effectively, with performance metrics such as accuracy, precision, recall, and specificity used to evaluate its effectiveness. The experimental results demonstrated an impressive prediction accuracy of 95.65%. Results demonstrated the RF model's high accuracy and robustness in predicting CBPP, offering a practical tool for early diagnosis and decision-making. The study further highlights the potential of deploying the model in a mobile application, enabling cattle breeders to perform preliminary diagnoses and seek timely veterinary intervention. This approach aligns with global efforts to integrate technology into agricultural practices, enhancing livestock disease management and supporting sustainable agricultural development. The findings contribute to the growing body of knowledge on applying machine learning to livestock disease prediction, showcasing Random Forest's potential to transform disease management strategies in the livestock sector.

1. Introduction

Agriculture is a cornerstone of Nigeria's economy, contributing approximately 21% to the country's gross domestic product (GDP) in the second quarter of 2023 (Statista, 2023). While crop production dominates, accounting for nearly 19% of the GDP, the livestock sector, though smaller, plays a vital role, contributing 1.27%. This underscores the importance of livestock as a driver of food security, income generation, and rural employment, making it integral to Nigeria's agricultural and economic development. Livestock, including cattle, sheep, and goats, provides essential resources such as food, raw materials, and economic opportunities. However, the oil boom of the 1970s led to a decline in livestock management, with significant resources diverted toward food imports during 1970–1974 (Zhang, Su, and Chen, 2021). Despite this, cattle remain crucial to northern Nigeria's agricultural systems, valued for their contributions to beef and dairy production. Beef, in particular, is a nutrient-rich food source, providing essential nutrients such as iron, zinc, and vitamin B12, which are critical for vulnerable populations like women and infants (Beef Cattle Research Council, n.d.). Beyond nutrition, the livestock subsector drives job creation, income generation, and the supply of raw materials like hides and skins (Valdes-Donoso *et al.*, 2017). However, the sector faces persistent challenges, including inadequate feeding practices, poor hygiene, and disease outbreaks that compromise cattle health and productivity. One of the most critical threats to cattle production is Contagious Bovine Pleuropneumonia (CBPP), a highly infectious and economically devastating disease. CBPP leads to significant losses in cattle productivity, quality of beef and milk, and

^{*}Corresponding author. Tel.: +2347034798964

E-mail address: adamumgirei@mau.edu.ng (Musa A. Girei)

<https://doi.org/10.62054/ijdm/0104.18>

broader economic impacts, particularly in regions like northern Nigeria, where cattle farming underpins rural livelihoods.

The rapid spread of diseases like CBPP underscores the urgency for early detection and intervention. As Noone *et al.* (2021) highlight, timely diagnosis of cattle diseases is essential to sustaining livestock health and ensuring economic stability. This research adopts the Random Forest (RF) algorithm, a widely recognized machine learning technique, to analyze CBPP symptoms and enhance early detection. RF's robustness, accuracy, and ability to handle complex datasets make it particularly suitable for this application, providing reliable predictions for proactive disease management. The primary objective of this study is to develop and evaluate a Random Forest-based classification model for predicting CBPP in cattle. The model aims to empower cattle breeders with a tool for early diagnosis, enabling them to take preventive measures before seeking veterinary intervention. By leveraging historical veterinary data and CBPP symptom patterns, this approach not only enhances disease prediction accuracy but also minimizes livestock losses, thereby improving food security and economic resilience.

This study contributes to the growing body of knowledge on applying machine learning to livestock disease management. The Random Forest model developed in this research demonstrates how predictive tools can enhance animal health and agricultural productivity. Furthermore, this project proposes the model deployment in an Android mobile application that integrates the RF model, allowing cattle breeders to conduct preliminary diagnoses of CBPP from the convenience of their homes. This system aligns with global efforts to incorporate technology into agricultural practices, supporting sustainable development and bolstering resilience in the livestock sector.

1.1 Contagious Bovine Pleuropneumonia

Contagious bovine pleuropneumonia (CBPP) is a highly infectious respiratory disease in cattle, caused by the bacterium *Mycoplasma mycoides* subspecies *mycoides*. Due to its highly contagious nature, it poses a serious risk to cattle populations, spreading rapidly through direct contact, contaminated equipment, aerosols, and carrier animals (Alhaji *et al.*, 2020). The stealthy progression of CBPP and the difficulties in controlling its spread present a significant challenge, particularly for cattle-rearing regions in sub-Saharan Africa. The symptoms of CBPP can vary widely among infected cattle, with some showing mild signs while others exhibit severe clinical symptoms. Common signs include respiratory distress, nasal discharge, systemic symptoms, pleuritis, pneumonia, and, in some cases, abortions in pregnant cows. Diagnosis generally requires a clinical examination by a veterinarian, supplemented by serological tests and postmortem evaluations. Given the variability in symptoms, machine learning techniques offer a promising solution to support the clinical diagnosis of CBPP, improving accuracy in classification and prediction. Machine learning, a branch of artificial intelligence focused on developing predictive and inferential algorithms, provides data-driven insights and decisions based on comprehensive data analysis (Zhang *et al.*, 2020).

2. Literature Review

Sikiru, *et al.* (2023) analyzed Africa's livestock production amidst climate change using 40 years of meteorological and livestock production data. By 2050, the average surface temperature in Africa is projected to rise by 2.2 °C, with livestock populations expected to grow despite environmental challenges. West Africa was identified as the most vulnerable region. Small ruminants such as sheep and goats showed greater resilience, evidenced by their increasing populations. Strategies were suggested to sustain livestock production, but there is a need to explore predictive modeling techniques, such as random forests, to improve the accuracy of resilience projections and livestock adaptation measures under varying climate scenarios. Uzonwanne, Francis, and Nwokoye (2023) examined the impact of livestock production on Nigeria's GDP from 1981 to 2021 using an auto-redistributed lag (ARDL) model. The study found a significant positive relationship between livestock production and economic growth, with a bidirectional influence. While the findings emphasized economic benefits, they did not explore regional variations or how climatic factors could affect livestock production's contribution to GDP. Random forest models could address this gap by integrating environmental and economic factors for more granular predictions. Tengfei *et al.* (2024) explored animal agriculture's impact on host-microbe interactions and its role in zoonotic bacterial infections and antimicrobial resistance (AMR). They proposed strategies like integrated surveillance programs, novel treatments, and education reforms to mitigate these challenges within the One Health framework. Although the study highlighted the interconnected impacts of animal agriculture, it lacked a quantitative analysis of risk factors. Random forest techniques could be employed to analyze large datasets and identify key predictors for zoonotic infections and AMR trends.

Golden, Rothrock, and Mishra (2019) applied random forest (RF) and gradient boosting machine (GBM) models to predict *Listeria* spp. prevalence in farming environments using meteorological variables. While both models performed well for feces samples, GBM outperformed RF for soil samples. The study demonstrated the effectiveness

of these models for food safety predictions but did not explore their scalability to broader zoonotic disease prediction. Extending random forest models to integrate diverse environmental and host factors could enhance disease risk management in livestock farming. Noone *et al.* (2021) introduced IoT and data mining for diagnosing rare cattle diseases, offering a cost-effective solution for animal healthcare in resource-limited settings. While promising, the approach was not tested against machine learning models like random forests, which could potentially improve diagnostic accuracy by analyzing complex interrelationships in health data. Zhang and Chen (2020) highlighted the potential of machine learning (ML) techniques for analyzing and predicting animal epidemics. Their study provided an overview of ML applications but lacked a detailed evaluation of individual techniques. Random forest models, with their capacity for handling imbalanced datasets, could offer significant improvements in epidemic predictions by identifying critical features associated with disease spread. Sana, Bhanushikha, and Roshan (2023) utilized ML techniques, including random forests, to classify animal diseases and predict zoonotic potential. However, they did not report the performance metrics of each technique. A deeper exploration of random forest models' strengths, such as feature importance rankings and handling high-dimensional datasets, could address this gap and improve classification accuracy.

Atıl and Akıllı (2016) and Khanikar, Phookan, and Gogoi (2024) emphasized the role of artificial neural networks (ANNs) in livestock studies. ANN models were shown to outperform cluster analysis in dairy cattle classification (Atıl and Akıllı, 2016) and demonstrated high accuracy in predicting breeding values, milk yield, and meat quality (Khanikar *et al.*, 2024). While ANN has been highly effective, its reliance on extensive training data and computational resources highlights the need for complementary models like random forests, which offer better interpretability and efficiency in feature selection. World Bank (2021), used techniques such as random forest to compute the probability of swine movements in two regions.

2.1 Gap and Random Forest Application

While existing studies demonstrate the potential of machine learning in livestock management, significant gaps persist. There is limited exploration of Random Forest models for assessing livestock resilience under varying climatic conditions, analyzing disease risks, and quantifying antimicrobial resistance trends. Furthermore, comparative evaluations of Random Forest performance relative to other machine learning models in livestock disease diagnosis and epidemic prediction remain scarce. This research addresses these gaps by leveraging Random Forest models to identify critical features and deliver robust, interpretable predictions. These insights can inform decision-making in livestock production and disease management, offering a practical approach to improve resilience and health outcomes.

2.1 Random Forest

Random Forest (RF) is an ensemble learning technique employed for both classification and regression tasks. It involves constructing multiple decision trees during the training process, and the final output is determined by the mode of the classes for classification or the mean prediction for regression. RF's effectiveness in predicting diseases makes it a highly suitable model (World Bank, 2021). As noted by Noone *et al.* (2021), RF is a user-friendly and scalable machine learning algorithm that often delivers excellent results without requiring hyperparameter tuning. One of the primary drawbacks of decision trees is their tendency to overfit, which occurs when the tree memorizes information rather than generalizing. Random Forest overcomes this issue as it is a type of ensemble learning that applies the algorithm multiple times to different or similar models.

The training process for RF involves the following steps:

1. Data Sampling: First, samples of data are extracted.
2. Random Feature Selection: For each tree, a random subset of features is selected, which helps ensure diversity in the features and reduces correlation between the trees.
3. RF Training: A RF is built using the chosen features, continuing until criteria such as maximum tree depth or the minimum number of data points per leaf node are met.
4. Iteration: Steps 1 to 3 are repeated multiple times to create a forest of decision trees.
5. Prediction: To make predictions on new data, all the trees in the forest are used, and the final prediction is obtained by averaging the outputs of all the trees.

The Random Forest algorithm is a variation of the traditional decision tree model. It determines the splitting criterion for each node based on the Gini impurity or entropy of the dataset. The Gini impurity is calculated as follows:

$$\text{Gini impurity: } \text{Gini}(p) = 1 - \sum (p_i^2) \quad (1)$$

where p represents the vector of class probabilities, and p_i is the probability of a particular class.

$$\text{Entropy: } H(p) = - \sum (p_i \log_2(p_i)) \quad (2)$$

Similarly, here p represents the vector of class probabilities, and p_i is the probability of a given class. Random forests tend to outperform single decision trees because they are less susceptible to overfitting and can capture more complex decision boundaries.

Random Forest is an excellent choice for CBPP prediction due to its ability to handle high-dimensional, noisy, and non-linear data while providing interpretable results. Compared to SVM and Decision Trees, RF offers superior accuracy, scalability, and robustness, making it ideal for predicting CBPP in complex, real-world datasets.

SVM can be computationally expensive for large datasets, especially with non-linear kernels. RF scales more efficiently to large datasets with faster training and prediction times. In CBPP prediction, where the prevalence of disease (positive class) may be low compared to the healthy population, RF can adjust weights or resample data during training to handle class imbalance better than SVM. RF provides interpretable outputs, such as feature importance scores, which are easier to understand compared to SVM's abstract decision boundaries.

3. Materials and Methodology

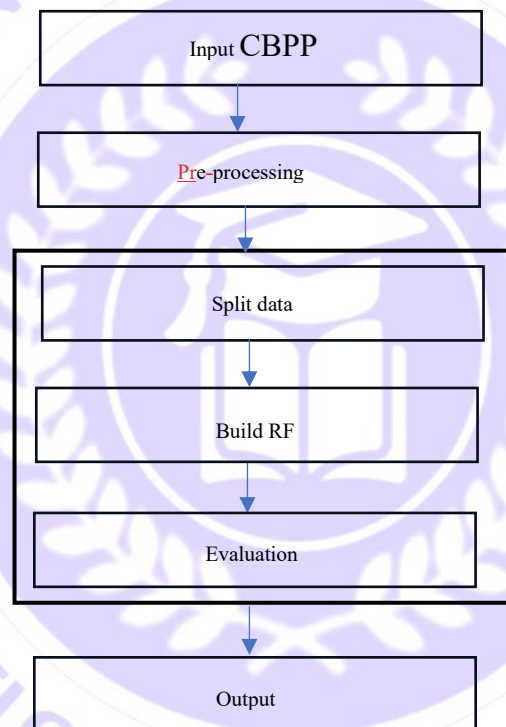


Figure 1: Model development cycle.

3.1 Data Extraction

To develop the model, data were sourced from various platforms, including the internet and veterinary clinics in Yola, Adamawa State. A total of 847 datasets were collected, which serve as essential input for making predictions using a machine learning algorithm based on historical data.

3.2 Input CBPP

In building a classification model, defining the input is crucial. For this study, the input consisted of CBPP symptoms, which were utilized by the prediction model to make accurate predictions. The model identified various symptoms, including respiratory distress, nasal discharge, fever, loss of appetite, weakness and lethargy, abortion, pleuritis, and pneumonia.

1. Preprocessing

Preprocessing in machine learning involves preparing and cleaning raw data before training the model. The objective is to improve data quality by addressing issues such as noise, missing values, and irrelevant information. The following preprocessing steps were applied: Outlier Detection and Handling: Outliers were identified and addressed through careful inspection and visualization of the dataset. Outliers are data points that significantly differ from the other observations in a dataset. They are unusually high or low values that can distort machine learning models if not properly handled. In our case, this occur due to typo in our dataset as some data were wrongly keyed.

Removing Duplicates: Duplicate entries were identified and removed from the dataset.

Data Type Conversion: Categorical variables were encoded using one-hot encoding. In this encoding, '0' represented the absence of symptoms, while '1' indicated the presence of symptoms. The target class was encoded as either "Yes" or "No."

One-Hot Encoding is a technique used to convert categorical data into a numerical format for machine learning models, where each category (or class) is represented as a vector with a single **1** in the position corresponding to the category, and **0s** in all other positions. This method ensures that categorical features, which cannot be used directly by machine learning models, are converted into a format that can be processed effectively.

2. Building the RF Classifier

To build, train, and evaluate the model after data collection, the random forest (RF) package was installed and utilized. R version 4.1.3 was employed to run the model. The random forest library was imported to construct the RF classifier. The dataset, comprising predictor variables and the target variable, was loaded into R in CSV format. Once the dataset was imported, the RF model was implemented to predict CBPP disease based on user input. The system takes user inputs and, using test data, predicts the likelihood of CBPP.

3.3 Data Split into Training and Test Sets

The dataset was divided into two parts: 70% for training and 30% for testing. The training set was used to train the RF model, while the test set was utilized for prediction and evaluation, assessing the model's performance on unseen data. This gives a larger test set, improving the reliability of the model evaluation, while still leaving enough data for training. With only 253 samples in the test set, we are more confident that the model evaluation is robust.

3.4 Training and Testing

Of the total 847 datasets, 592 were used as the training data to build and train the RF model. The remaining 253 datasets were used as new inputs to evaluate the model's performance.

3.5 Model Performance Evaluation

Evaluating model performance is a critical step in understanding how well it is likely to perform on unseen data. Various metrics were used to assess the model:

True Positives (TP): Cases where cattle exhibited CBPP symptoms, and the classifier correctly predicted a positive outcome.

False Negatives (FN): Cases where the classifier predicted a negative outcome, but the actual outcome was positive.

The following performance metrics were derived from the confusion matrix:

- i. Accuracy: $(TP + TN) / (TP + TN + FP + FN)$
- ii. Precision: $TP / (TP + FP)$
- iii. Recall (Sensitivity or True Positive Rate): $TP / (TP + FN)$
- iv. Specificity (True Negative Rate): $TN / (TN + FP)$

4. Results

4.1 RF Model Prediction

The model was tested using the 'predict()' function, with test data serving as new inputs for the model. Table 1 provides the row and column totals of the confusion matrix for the classifier.

Table 1. Total rows and columns of the confusion matrix.

253	Prediction		
	NO	YES	Total
NO	77	0	77

YES	11	165	176
Total	88	165	253

From Table 1, a total of 253 test data points were used to evaluate the binary classification model. The model is designed to classify cattle as either "YES" or "NO" for CBPP (Contagious Bovine Pleuropneumonia). The classifier correctly predicted "NO" for 77 instances, meaning that these cattle showed no symptoms of CBPP, confirming the prediction as accurate. In this context, "NO" is considered the positive class. The model also predicted "YES" for 165 cases, indicating that these cattle exhibited symptoms of CBPP. However, the actual count of positive cases was 176. The model misclassified 11 cases as "NO" when they were actually "YES." These 11 instances represent true negative predictions for "YES" and false predictions for "NO." Table 2 provides the sensitivity, specificity, positive prediction value, and negative prediction value for the confusion matrix.

Table 2: Model sensitivity, specificity, positive prediction value, and negative prediction

Terms	Value%
Sensitivity:	0.8750
Specificity:	1.0000
Pos Pred:	1.0000
Neg Pred:	0.9375

4.2 Model Accuracy

The accuracy of the RF model on the training data is 94.58%, with the optimal accuracy achieved when 'mtry = 2'. Figure 2 illustrates the graph showing the optimal 'mtry' value.

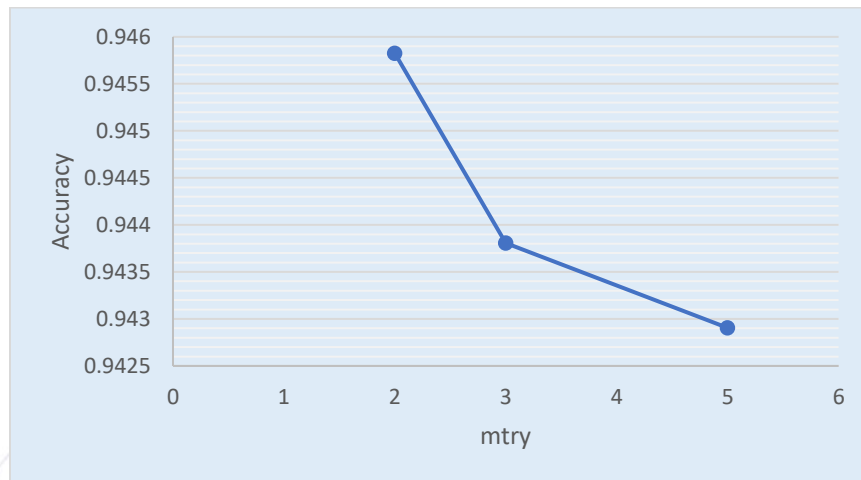


Figure 2. Graph of the optimal value of mtry.

4.3 Interpretation of Results

Accuracy is a primary metric for evaluating the overall performance of a machine learning model. It assesses the model's ability to correctly classify instances in both the training and test datasets. In this study, the model achieved an accuracy of 94.58% on the training data and 95.65% on the test dataset. This indicates a 1.07% improvement in predictive performance when evaluated on unseen data. As detailed in Tables 1 and 2, the sensitivity (or recall for the "NO" class) measures the proportion of actual "NO" instances correctly identified by the model. The sensitivity value of 87.50% indicates that the model successfully detected a significant majority of true "NO" cases. The specificity represents the proportion of actual "YES" instances that were correctly classified. The model achieved a perfect specificity of 100.00%, meaning all actual "YES" instances were correctly identified.

The Positive Predictive Value (PPV) quantifies the accuracy of predictions labeled as "NO." The model demonstrated a PPV of 100.00%, signifying that all predicted "NO" cases were indeed "NO." Conversely, the Negative Predictive Value (NPV), which measures the accuracy of "YES" predictions, was 93.75%, indicating strong predictive reliability for "YES" instances as well. To balance performance across both classes, the Balanced Accuracy (BA), calculated as the average of sensitivity and specificity, was 93.75%. This metric further confirms the model's robust capability in handling imbalanced classifications, considering the "NO" class as the positive class. Overall, the model exhibits high performance, particularly in identifying "NO" instances, with excellent sensitivity, specificity, and precision. The strong results across key metrics support the potential application of this model in practical settings.

The limited evaluation of random forest (RF) models in livestock-related studies, such as disease diagnosis and epidemic prediction, underscores their potential. The model's 94.58% training accuracy and 95.65% test accuracy highlight the feasibility of deploying such models in mobile or web-based applications accessible to users. This would be especially valuable for analyzing zoonotic disease risks and antimicrobial resistance trends. Similar to other supervised learning techniques, RF is capable of identifying critical features and generating robust, interpretable predictions, making it a powerful tool for decision-making in livestock production, disease management, and food safety.

5. Conclusion

The performance of our random forest model demonstrates its potential as a valuable tool for cattle breeders, enabling initial at-home diagnoses before seeking veterinary assistance. This aligns with findings by Sana *et al.* (2023), which emphasize the utility of machine learning (ML) techniques in addressing emerging threats to livestock health. Many animals suffer due to the delayed diagnosis and limited understanding of illnesses by caretakers and experts. Timely identification of diseases is essential for safeguarding animal health, as delays often result in avoidable losses.

The application of machine learning methods, such as our random forest model, can aid in disease control by classifying conditions based on symptoms, similar to a veterinarian's diagnostic process. This approach not only empowers breeders to act swiftly but also enhances the efficiency and accuracy of veterinary services. By integrating such tools into practical solutions, breeders can save time, reduce losses, and contribute to improved overall livestock health management. The findings contribute to the growing body of knowledge on applying machine learning to livestock disease prediction, showcasing Random Forest's potential to transform disease management strategies in the livestock sector.

References

- Alhaji, N. B., Ankeli, P. I., Ikpa, L. T., and Babalobi, O. O. (2020). Contagious Bovine Pleuropneumonia: Challenges and prospects regarding diagnosis and control strategies in Africa. *Veterinary Medicine: Research and Reports*, 11, 71–85. <https://doi.org/10.2147/VMRR.S180025>
- Atıl, H., and Akıllı, A. (2016). Comparison of artificial neural network and K-means for clustering dairy cattle. *International Journal of Sustainable Agricultural Management and Informatics*, 2(1), 40-52.
- Beef Cattle Research Council (n.d). Nutritional qualities of beef. Retrieved January 18, 2024 from <https://www.beefresearch.ca/topics/nutritional-qualities-of-beef/>
- Golden, C., Rothrock, M., and Mishra, A. (2019). Comparison between random forest and gradient boosting machine methods for predicting *Listeria* spp. prevalence in the environment of pastured poultry farms. *Food Research International*. 122. 10.1016/j.foodres.2019.03.062.
- Gota, M., Ricardo, V. V., Fabyano, F. S., and Samodha, C. F. (2018). Big Data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. *Proceedings*
- Khanikar, D., Phookan A., and Gogoi, A. (2024). Artificial Neural Networks- An Introduction and Application in Animal Breeding and Production: A Review . *Agricultural Reviews*. 45(3): 480-487. doi: 10.18805/ag.R-2421.
- Noone, V. K., Sai, T. Y., Sandeep, K., Sangamesha, V., and Sumanth, R. (2021). Symptoms diagnosis using machine learning model random forest. *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, 9(6), 45–98. <https://doi.org/10.22214/ijraset.v9i6.5783>
- Sana, R., Bhanushikha, R., and Roshan, L. (2023). Animal disease prediction using machine learning techniques. *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, 11(6), 7–538.
- Singh, R. K., and Sharma, V. C. (2015). Ensemble approach for zoonotic disease prediction using machine learning techniques.
- Statista. (2023). Share of GDP by agricultural sector in Nigeria 2023. Retrieved December 23, 2023 from <https://www.statista.com/statistics/1207940/share-of-gdp-by-agricultural-sector-in-nigeria/>
- Sumit, S. S., Awang Rambli, D. R., Mirjalili, S., Ejaz, M. M., and Restinet, M. S. U. (2022). On improving the performance of tiny-yolo-based CNN architecture for applications in human detection. *Applied Sciences*, 12(18), 9331.
- Tengfei Z., Rhea, N., Wenting, Z., Xitian, P., Yu, S., Youxiang, Z., Qingping, L., Guoyuan, W., and Zhenyu, C. (2024).

The impacts of animal agriculture on One Health—Bacterial zoonosis, antimicrobial resistance, and beyond, *One Health*, Volume 18, ISSN 2352- 7714, <https://doi.org/10.1016/j.onehlt.2024.100748>.
(<https://www.sciencedirect.com/science/article/pii/S2352771424000740>)

Uzonwanne, M. C., Francis, O. C., and Nwokoye, M. O. (2023). Impact of livestock production on gross domestic product in Nigeria. *International Journal of Advanced Economics*, 5(5).

Valdes-Donoso, P., VanderWaal, K., Jarvis, L. S., Wayne, S. R., and Perez, A. M. (2017). Using machine learning to predict swine movements within a regional program to improve control of infectious diseases in the US. *Frontiers in Veterinary Science*, 4(2). <https://doi.org/10.3389/fvets.2017.00002>

World Bank. (2021). GDP growth (annual %) Nigeria. Retrieved 22nd, December, 2023 <https://data.worldbank.org/indicator>.

Yoade, J. A., Ibrahim, H. I., and Ibrahim, H. Y. (2009). Analysis of women involvement in livestock production in Lafia area of Nasarawa State, Nigeria. *Livestock Research for Rural Development*, 21, Article #220. Retrieved December 18, 2023, from <http://www.lrrd.org/lrrd21/12/ayoa21220.htm>

Zhang, S., Su, Q., and Chen, Q. (2021). Application of machine learning in animal disease analysis and prediction. *Current Bioinformatics*, 16(7), 972–982.

Zhang, S., Su, Q., and Chen, Q. (2020). Application of machine learning in animal disease analysis and prediction. *Current Bioinformatics*, 15(7), 561–573. <https://doi.org/10.2174/1574893615999200728195613>