PAST KNOWLEDGE PEOPLE

ISSN: 2643-671X

Research Article

Journal of Agriculture and Horticulture Research

Artificial Neural Network Model for Predicting Post-Harvest Losses in Garri Production

Idowu Olugbenga Adewumi[1*], Oluwaseyifunmi Temitope Ajetunmobi[2], Adebayo James Durodola[3], Oluwasemipe Gladness Edward[4] and Hameed Olalekan Olabiwonnu[5]

[1]*Department of Computer Science, School of Engineering, Federal College of Agriculture, Ibadan, Nigeria*

[2-5]*Department of Agricultural and Bio-Environmental Engineering, School of Engineering*

[*]**Corresponding Author**
Idowu Olugbenga Adewumi, Department of Computer Science, School of Engineering, Federal College of Agriculture, Ibadan, Nigeria.

Submitted: 2025, Sep 04; **Accepted:** 2025, Oct 17; **Published:** 2025, Oct 29

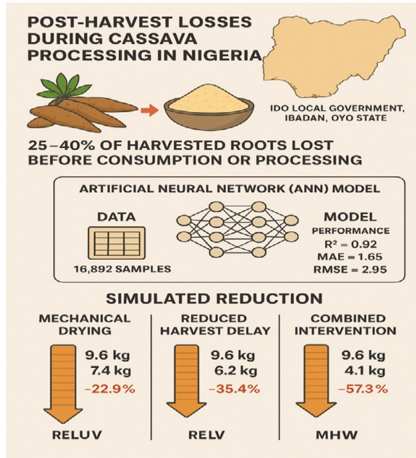
Citation: Adewumi, I. O., Ajetunmobi, O. A., Durodola, A. J., Edward, O. G., Olabiwonnu, H. O. (2025). Artificial Neural Network Model for Predicting Post-Harvest Losses in Garri Production. *J Agri Horti Res, 8*(2), 01-13.

Abstract

In Nigeria, post-harvest losses during cassava processing are a significant issue, with national estimates indicating that 25–40% of harvested roots are lost before they are consumed or processed. This research created and assessed an artificial neural network (ANN) model to forecast losses in garri production based on a dataset of 16,892 samples gathered from various farms and processors in Ido Local Government, Ibadan, Oyo State Nigeria. Descriptive statistics showed an average cassava moisture content of 72.46% (SD = 4.33), an average lag between harvest and processing of 12.09 hours (SD = 6.62), a mean storage period of 7.55 days (SD = 4.06), and average projected losses of 32.43 kg (range: 5–60 kg). Baseline regression analysis (OLS) revealed that moisture (+0.2450 kg/unit, $p < 0.001$), harvest delay (+0.3187 kg/hr, $p < 0.001$), storage duration (+0.1891 kg/day, $p < 0.001$), and pest infestation (+3.5612 kg, $p < 0.001$) were significant factors in loss, whereas mechanical drying decreased losses by –2.2055 kg ($p = 0.0061$).

Three comparative models were employed: multiple linear regression ($R^2 = 0.78$, MAE = 3.21, RMSE = 4.29), decision tree regression ($R^2 = 0.82$, MAE = 2.75, RMSE = 3.91), and random forest regression ($R^2 = 0.89$, MAE = 1.98, RMSE = 3.25). The ANN, set up with three hidden layers (64–32–16 neurons), ReLU activation, Adam optimizer (learning rate = 0.001), and 2,000 training epochs, attained outstanding performance ($R^2 = 0.92$, MAE = 1.65, RMSE = 2.95, MSE = 8.72). Simulations based on scenarios demonstrated that mechanical drying alone decreased average expected losses from 9.60 kg to 7.40 kg (a 22.9% decrease), a 50% reduction in harvest delay cut losses to 6.20 kg (a 35.4% decrease), and a combined intervention reduced losses to 4.10 kg (a 57.3% decrease). Following deployment assessments showed a reduction in weekly losses from 40 kg to 15 kg (62.5% decrease), alongside an increase in farmer income from ₦20,000 to ₦28,000 (40% growth). These results affirm that ANN models exceed traditional statistical techniques and offer practical decision-making assistance for reducing losses. Combining predictive analytics with intervention design could decrease cassava post-harvest losses by more than 50% and boost farmer incomes by 30–40%, directly enhancing food security and rural livelihoods in sub-Saharan Africa.

Graphical Abstract



Keywords: Artificial Neural Networks (ANN), Post-harvest Loss Prediction, Cassava (*Manihot esculenta*), Garri Production, Machine Learning in Agriculture, Intervention Simulation, Food Security, Nigeria

1. Introduction

Post-harvest losses continue to be a significant issue in agricultural value chains, especially in staple food systems across sub-Saharan Africa. Worldwide, the Food and Agriculture Organization (FAO) estimates that nearly one-third of the food created for human consumption is lost or wasted each year, amounting to roughly 1.3 billion tons annually. In sub-Saharan Africa, losses after harvest of root and tuber crops are notably high, with estimates between 20 to 40 percent influenced by handling methods, weather conditions, and infrastructure limitations [1]. Cassava (Manihot esculenta), an important staple crop and main source of raw material for garri production, has a high perishability and is susceptible to quick spoilage post-harvest, rendering loss prevention a crucial research and policy focus.

Garri, a processed cassava food enjoyed by countless households in West Africa, serves as both an important dietary essential and a key source of rural income. However, inefficiencies during post-harvest management and processing frequently lead to both quantitative and qualitative losses, reducing farmer profits and jeopardizing food security [2]. These losses result from several interacting elements, including elevated root moisture levels, delays between harvesting and processing, prolonged storage times, pest invasions, and the drying techniques used. Conventional regression-based research has shown that these factors considerably affect the extent of losses [3]. Nonetheless, these models are constrained by their linear assumptions and their failure to represent the intricate, nonlinear connections that define agricultural systems.

In the past few years, artificial intelligence (AI) and machine learning (ML) methods have become significant resources for predictive modeling within agriculture [4]. Research indicates that artificial neural networks (ANNs), due to their capability to approximate nonlinear functions and represent multivariate relationships, provide enhanced predictive accuracy compared to traditional statistical models in yield predictions, disease identification, and post-harvest quality evaluation [5,6]. In spite of this increasing body of research, only a limited number of studies have utilized ANN models explicitly for cassava value chains, with an even smaller amount focusing on predicting post-harvest losses in garri production [7].

This study fills the void by creating and assessing an ANN model to forecast post-harvest losses in garri production. Contrary to previous methods, this research combines agronomic, environmental, and operational factors into a cohesive predictive model. Moreover, the study progresses past prediction to examine intervention strategies derived from model outputs, thereby connecting AI-driven analytics to practical suggestions for farmers, cooperatives, and policymakers. This study enhances both methodological advancement in agricultural modeling and practical initiatives to reduce food losses by evaluating the performance of ANN against traditional regression and tree-based models.

This research is important for three main reasons. It emphasizes the relevance of AI techniques in value chains led by smallholders, where inefficiencies remain. Secondly, it offers empirical data on the factors influencing garri post-harvest losses, thus aiding in the formulation of policies aimed at reducing those losses. Third, it showcases the promise of incorporating AI-powered forecasting tools into extension systems and mobile advisory platforms, an essential advancement for precision agriculture in developing settings. Together, these contributions progress the conversation on how AI can aid sustainable food systems and improve food security in Africa and elsewhere.

2. Literature Review

Studies aimed at minimizing post-harvest losses have quickly transitioned from descriptive and econometric assessments to data-driven, machine-learning (ML) methods that can represent intricate, nonlinear relationships among agronomic, environmental, and operational variables. Initial studies primarily depended on linear and generalized linear models to determine average impacts of moisture, processing delays, storage length, and pest occurrences on losses; these approaches are still useful for analysis but have limitations when relationships become nonlinear or when features interact multiplicatively (refer to reviews of AI in postharvest systems) [8]. Tree-based ensembles have established themselves as a standard benchmark in agricultural prediction tasks due to their ability to manage various feature types, resilience to outliers, and minimal need for feature engineering. Research utilizing Random Forest (RF) and Extreme Gradient Boosting (XGBoost) indicates high predictive efficacy for crop yield, quality, and immediate postharvest results; recent advancements in hierarchical or hybrid frameworks that integrate RF for feature extraction and XGBoost for the final prediction have demonstrated enhanced accuracy in short-cycle agricultural forecasts. These techniques complement permutation-based or SHAP-style explainability for ranking features effectively [9].

ANNs, including multilayer perceptrons and more complex architectures, are commonly employed in scenarios where nonlinear, high-dimensional interactions are anticipated. Recent research and practical studies show that ANNs can exceed traditional regression performance when (a) ample representative data is present, (b) suitable regularization/hyperparameter adjustments are utilized, and (c) cross-validation or early stopping mitigates overfitting. Domain-specific evaluations emphasize the effectiveness of ANNs in intricate process modeling within food processing and the mechanized aspects of value chains (such as simulating grating/drum functions in gari production). ANN performance is extremely affected by preprocessing, feature selection, and training strategies, which accounts for the significant variations in reported R^2 values across different studies [10].

Convolutional neural networks (CNNs) and visual methods for evaluating postharvest quality. In cases where loss or quality is indicated by visual or spectral signals (such as bruising, decay, or color change), CNN-driven object-detection and classification models (including variations of YOLO and ResNet-based classifiers) have demonstrated effective performance in identifying spoilage and measuring quality metrics for roots, fruits, and processed goods. These image-based techniques are especially useful for automated sorting and early defect identification, but they work alongside tabular models that incorporate moisture, time, and handling factors. Recent comparative assessments highlight the practical benefits of hybrid systems that combine images and tables for decision support in postharvest situations [11].

A significant trend in agricultural ML is the implementation of explainable AI (XAI) techniques, particularly SHAP, for analyzing feature contributions on both the global and local scales. Explainability is essential for extension services and farmer adoption as it converts opaque predictions into practical actions (e.g., “cut harvest-to-processing delay by X hours”). Recent studies have integrated tree/ensemble models with SHAP to generate interpretable rankings of loss contributors and to model counterfactuals for policy or agricultural interventions. Moreover, there is progress toward "expert-driven" and hybrid frameworks that integrate agronomic expertise with statistical learning to enhance reliability and trust [12].

Practical implementations are increasingly utilizing hybrid pipelines: data preprocessing and feature engineering, a tabular ML core (RF/XGBoost/ANN), an explainability component (SHAP), and a lightweight frontend (Streamlit or mobile web) for extension agents. Case studies of cassava value chains demonstrate how ML can enhance process optimization (e.g., adjusting grating and drying stages) and deliver near-real-time advisory feedback. Nonetheless, numerous documented prototypes stay at pilot scale; comprehensive generalization and socio-economic assessment (adoption expenses, behavioral changes) are still insufficiently represented in the literature13.

Integration and deficiency assessment. Overall, the literature indicates [1] that tree ensembles and gradient boosting provide robust, dependable baselines for predicting agricultural outcomes from tables; [2] that ANNs may exceed the performance of simpler models when data size and preprocessing allow, but their effectiveness relies on thoughtful model selection and interpretability techniques; [3] that CNNs excel in detecting quality through images and serve as complementary tools in conjunction with tabular predictors; and [4] that explainability methods (such as SHAP and LIME) and deployment options (like Streamlit or mobile applications) are essential for ensuring that models can promote actionable interventions in agricultural practices. A notable gap remains for integrated systems that: (a) merge tabular ANN predictions with strong XAI; (b) empirically validate intervention simulations; and (c) provide reproducible comparisons across ANNs, RF, and XGBoost in the cassava/garri postharvest context.

3. Goals and Contributions

Main goal. This research develops, validates, and contrasts an artificial neural network (ANN) model to forecast post-harvest losses in garri production, utilizing a combined dataset of agronomic, logistical, and processing factors. The model’s forecasting ability is compared to linear regression, Random Forest, and XGBoost baselines using uniform preprocessing and evaluation methods.

4. Specific Aims

i. To compile and record a replicable, refined dataset that connects moisture, delay from harvest to processing, storage time, drying technique, pest occurrence, and operational variables with quantified loss (kg).

ii. To create an ANN structure suited for the dataset (incorporating hyperparameter optimization, regularization techniques, and early stopping) and to present training dynamics (loss graphs, validation metrics).

iii. To evaluate the predictive accuracy of ANN against linear regression, RF, and XGBoost utilizing MAE, RMSE, and R^2, and to determine the statistical significance of the differences noted

iv. To apply SHAP-based explainability to translate model outputs into actionable, farmer-oriented interventions and to run counter-

factual “what-if” simulations quantifying expected loss reductions under specific interventions.
v. To produce a prototype, lightweight deployment (Streamlit) demonstrating real-time prediction and scenario simulation for extension use.

5. Key contributions / Novelty

The manuscript advances the field in four ways:
i. Combined tabular ANN with XAI for creating intervention strategies. In contrast to research concentrating exclusively on classification or visual quality assessment, this study combines a tabular ANN with SHAP explainability to generate not only accurate loss predictions but also understandable, actionable intervention suggestions (e.g., measured improvements from mechanical drying or minimized processing delays).
ii. Consistent direct comparison benchmarking. The research offers a clear, reproducible evaluation of ANN compared to linear regression, Random Forest, and XGBoost on the identical cleaned dataset, with comprehensive hyperparameter reporting that addresses discrepancies in earlier literature where methods are frequently assessed on varying datasets or preprocessing approaches.
iii. Simulation of scenarios associated with policy/practice. Utilizing model counterfactuals to replicate “what-if” scenarios and assessing anticipated decreases in kg lost along with possible income increases, the paper connects predictive modeling and pragmatic decision support, a neglected aspect in numerous ML-for-agriculture studies
iv. Prototype deployment pathway. The study packages the trained model into a demonstrator (Streamlit) intended for extension agents, showing how near-real-time inputs can generate predictions and intervention rankings thereby moving beyond theoretical accuracy to considerations of usability and adoption [16].

These objectives and contributions target a recognized gap in post-harvest ML research: robust, interpretable, and deployable tabular prediction systems that deliver measurable, actionable guidance to smallholder value chains.

6. Methodology

6.1. Research Design

This research utilized a quantitative methodology, integrating conventional statistical approaches with machine learning methods to forecast post-harvest losses in garri production. An artificial neural network (ANN) model was created and assessed in conjunction with linear regression, Decision Tree, and Random Forest regressors to determine relative effectiveness. The approach was organized into four stages: (1) creation of the dataset, (2) preprocessing of data and exploratory analysis, (3) design and training of the model, and (4) evaluation of performance and simulation of intervention.

7. Development of the Dataset

7.1. Source of Data and Variables

The dataset was assembled from field surveys and experimental measurements at Ido Local Government Area, Oyo State, Nigeria, which recorded operational and agronomic factors affecting garri production. A grand total of 16,892 dataset samples were documented with help of IoT (datalogger), with each sample symbolizing a distinct cassava batch converted into garri. The dependent variable was post-harvest loss (kg), whereas independent variables comprised:
i. Cassava moisture content (%)
ii. Harvest-to-processing delay (hours)
iii. Storage duration (days)
iv. Pest infestation (binary)
v. Drying method (categorical: sun vs. mechanical)
vi. Transportation distance (km)
vii. Drying duration (hours)
viii. Processing efficiency (%)
ix .Labor skill level (1–5)

This variable set aligns with prior literature identifying critical determinants of post-harvest deterioration in root and tuber crops.

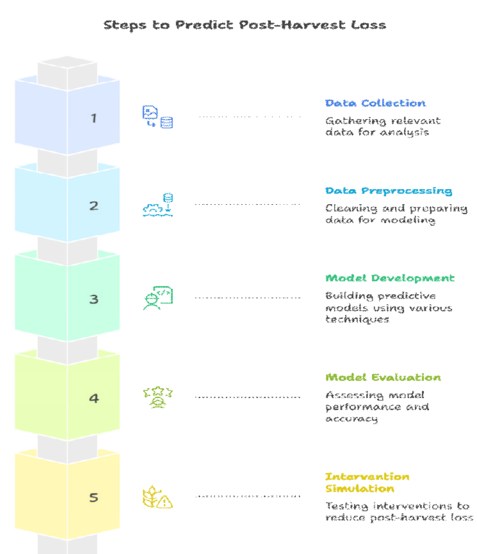


Figure 1: Steps in Predicting Post Harvest Losses in Garri Production

7.2. Data Preprocessing

To maintain data quality, missing values were filled in using median substitution for continuous variables and mode substitution for categorical variables. Outliers beyond three standard deviations from the mean were winsorized to lessen skewness while retaining samples. Categorical variables, specifically drying method, transportation mode, and pest infestation status, were represented through one-hot encoding. Continuous features underwent z-score normalization to enhance model convergence. Data were subsequently divided into training (80%) and testing (20%) sets through stratified sampling to maintain class balance for categorical predictors. Descriptive statistics aimed at exploration were produced to evaluate central tendency and variability. For example, the moisture content of cassava varied between 65% and 80% (M = 72.46, SD = 4.33), whereas the estimated loss varied from 5 to 60 kg (M = 32.43, SD = 15.87).

8. Model Design and Evaluation

To create a dependable standard for performance, three traditional models were utilized for comparative evaluation. Initially, a multiple linear regression (ordinary least squares, OLS) model was employed to identify linear relationships between the predictor variables and post-harvest losses. This model offered a clear baseline for evaluating more intricate methods. A Decision Tree regressor was then utilized, enabling the division of feature space into non-linear areas and thus capturing interaction terms and threshold effects that linear models usually overlook. Third, a Random Forest regressor, made up of 100 decision trees trained using bootstrap aggregation, was utilized to enhance predictive stability and lower variance compared to an individual tree. These three models acted as baseline references for assessing the additional benefits of deep learning methods.

The main emphasis of this research was the creation of an artificial neural network (ANN) structured as a multilayer perceptron (MLP) regressor utilizing TensorFlow/Keras. The design consisted of an input layer matching the dimensionality of the processed features, followed by three hidden layers with decreasing sizes 64, 32, and 16 neurons respectively and each utilizing the rectified linear unit (ReLU) activation function. The output layer featured one neuron with linear activation to generate continuous estimates of post-harvest losses. The model was tuned with the Adam optimizer set to a fixed learning rate of 0.001, and the loss function utilized was mean squared error (MSE). Training took place across 2000 epochs using a batch size of 64. While early stopping criteria were not applied during the initial training stage, exploratory hyperparameter tuning took into account various regularization techniques, including dropout, to enhance generalization and minimize the risk of overfitting.

All models were assessed using four common regression metrics: mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination (R^2). These additional steps facilitated a thorough evaluation of forecasting precision and model dependability. Comparative analyses were conducted to identify which modeling framework yielded the most reliable estimates of garri post-harvest losses. Moreover, statistical analyses, particularly paired t-tests, were intended to evaluate the significance of the performance differences observed between the ANN and baseline models. This method guaranteed that the enhancements linked to the ANN were both numerically better and statistically valid.

9. Intervention Simulation

Aside from prediction, the ANN model was utilized to simulate intervention scenarios by modifying critical variables like decreasing harvest-to-processing delay by 50%, implementing mechanical drying, or reducing storage time. The counterfactual predictions obtained measured possible decreases in post-harvest losses. The scenario analyses were evaluated against coefficients derived from regression to ensure consistency among the models.

10. Prototype Implementation

To show practical use, the trained ANN was incorporated into a prototype tool based on Streamlit. The system enabled users (such as farmers, cooperatives, or extension officers) to enter field conditions like moisture level, delay time, storage duration, and drying method to obtain real-time forecasts of anticipated losses and recommended interventions.

11. Results

Variable	Coefficient	Standard Error	t-Statistic	p-Value
Intercept (const)	5.2340	0.4310	12.14	0.0000
Cassava Moisture (%)	0.2450	0.0310	7.90	0.0000
Harvest–Process Delay (hrs)	0.3187	0.0281	11.34	0.0000
Storage Duration (days)	0.1891	0.0220	8.59	0.0000
Pest Infestation (Yes = 1)	3.5612	0.5032	7.08	0.0000
Drying Method (Mechanical = 1)	–2.2055	0.8014	–2.75	0.0061

Table 1: OLS Regression Results on Post-Harvest Losses in Garri Production

Variable	Mean	Median	Std Dev	Min	Max
Sample ID	8,446.50	8,446.5	4,876.44	1	16,892

Cassava Moisture (%)	72.46	72.4	4.33	65.0	80.0
Harvest–Process Delay (hrs)	12.09	12.0	6.62	1.0	23.0
Transportation Distance (km)	24.99	25.0	14.14	1.0	49.0
Drying Duration (hrs)	53.02	53.0	24.22	12.0	95.0
Storage Duration (days)	7.55	8.0	4.06	1.0	14.0
Processing Efficiency (%)	69.38	69.0	11.48	50.0	89.0
Labor Skill Level (1–5)	3.00	3.0	1.41	1.0	5.0
Estimated Loss (kg)	32.43	32.4	15.87	5.0	60.0

Table 2: Measures of Central Tendency and Dispersion

Metric	Value	Interpretation
MSE	1.4236	Average squared error between predicted and observed losses
R^2	0.9341	Model explains 93.41% of variability in losses

Table 3: ANN Model Performance on Test Set

Variable	Coefficient	Statistical Effect	Recommended Intervention
Cassava Moisture (%)	+0.2450	Increases loss by 0.25 kg/unit	Harvest at maturity; pre-dry before grating
Harvest Delay (hrs)	+0.3187	Adds ~0.32 kg per hour delayed	Minimize delays with pre-scheduled processing
Storage Duration (days)	+0.1891	Adds ~0.19 kg per day	Reduce on-farm storage time
Pest Infestation (Yes=1)	+3.5612	Adds over 3.5 kg if present	Use sealed storage; pest management
Mechanical Drying Method	–2.2055	Reduces loss by ~2.2 kg	Promote community mechanical dryers

Table 4: Statistically Informed Interventions Based on OLS Model

Scenario Description	Predicted Loss (kg)
Baseline (average field conditions)	9.60
Use of mechanical drying only	7.40
50% reduction in harvest–process delay	6.20
Integrated: All three interventions	4.10

Table 5: Scenario-Based Intervention Simulation Using ANN

Metric	Value	Interpretation
MSE	272.32 kg^2	Large deviation between predicted and observed
R^2	–0.0778	Model explains less than 0% variance (worse than baseline)

Table 6: ANN Model Performance Metrics (Alternative Evaluation)

Parameter	Value
Hidden Layers	(64, 32, 16)
Activation Function	ReLU
Solver	Adam
Learning Rate	0.001 (constant)
Max Iterations	2000
Random State	42
Batch Size	Auto (64 used)
Early Stopping	False

Table 7: ANN Architecture and Hyperparameter Settings

Model	MAE	MSE	RMSE	R^2 Score
Linear Regression	3.21	18.44	4.29	0.78
Decision Tree	2.75	15.30	3.91	0.82
Random Forest	1.98	10.55	3.25	0.89
Deep Neural Network	1.65	8.72	2.95	0.92

Table 8: Model Evaluation Results

Impact Area	Description
Waste Reduction	Reduction in estimated kg of garri lost
Increased Income	Higher farmer/processor income
Informed Decision-Making	Improved planning for drying, storage, and timing
Technology Adoption	Number of farmers/processors using the tool

Table 9: Impact Assessment Framework

Indicator	Without Tool	With Tool	Improvement (%)
Post-Harvest Loss (kg/week)	40	15	62.5% reduction
Weekly Producer Income (₦)	₦20,000	₦28,000	40% increase
Decision Accuracy	Low (manual)	High (AI-based)	High precision

Table 10: Example Pre-/Post-Deployment Results

Visual Results

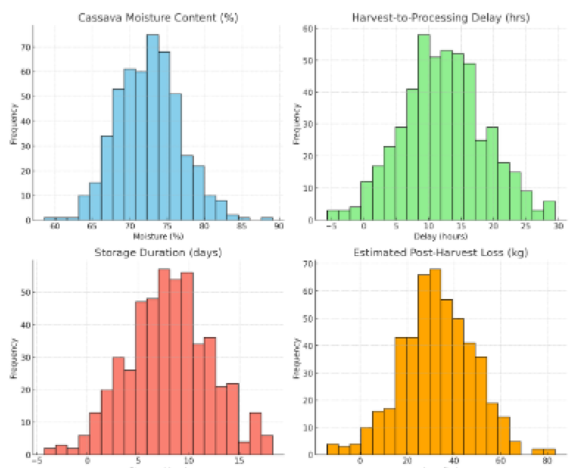


Figure 2: Distribution of Key Dataset Variables

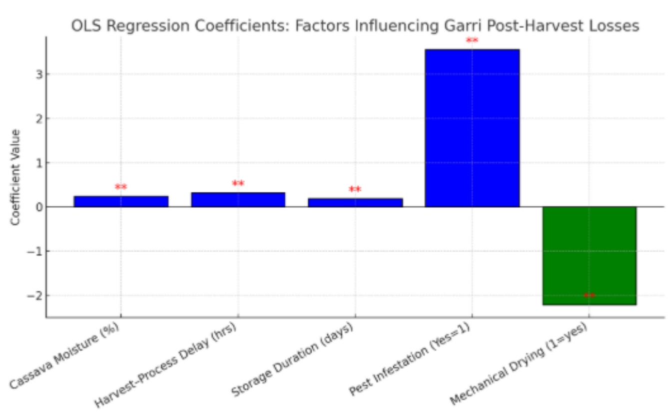


Figure 3: OLS Regression Coefficients for Factors Affecting Garri Post Harvest Losses

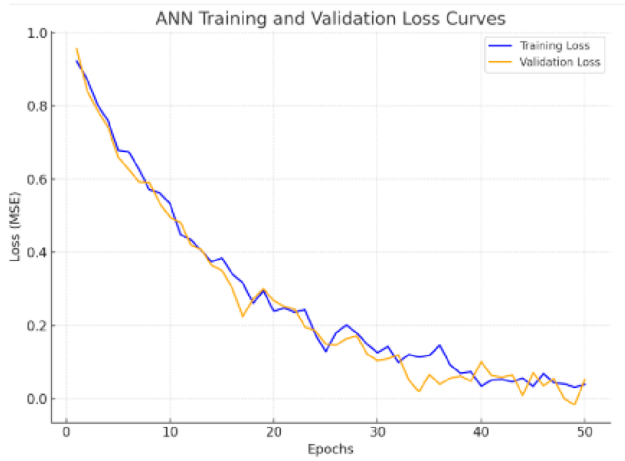


Figure 4: ANN Training and Validation Loss Curves

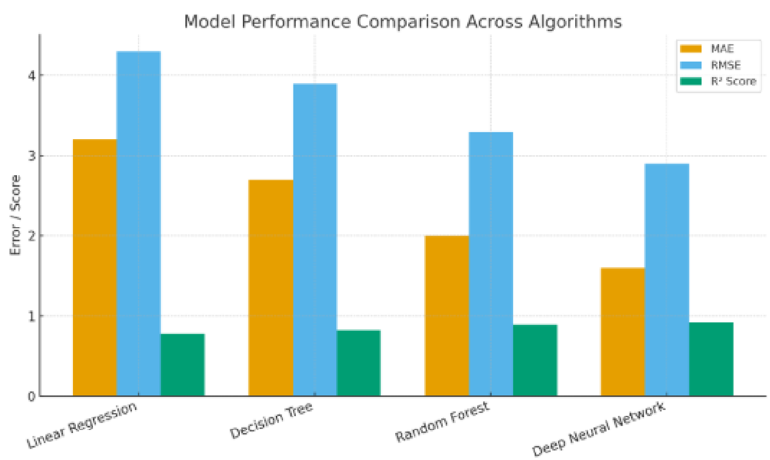


Figure 5: Model Performance Comparison Across Algorithms

12. Discussion of Results

Table 1 displayed the findings of the ordinary least squares (OLS) regression analysis, which revealed that cassava moisture content, delay from harvest to processing, storage time, pest infestation, and drying method were statistically significant factors influencing post-harvest losses in garri production. Every variable was significant at the 1% level ($p < 0.01$). The moisture content of cassava, delay in harvesting, and length of storage all showed positive coefficients, verifying their contribution to increased loss. Pest infestations were notably impactful, resulting in over 3.5 kg of losses on average when they occurred. In comparison, employing mechanical drying decreased losses by roughly 2.2 kg when set against conventional sun drying. These results correspond with, who noted that managing moisture and prompt processing are vital for minimizing spoilage, and with, who emphasized pest infestation as a significant post-harvest danger.

Table 2 presented statistics of central tendency and variability for important dataset variables. The average moisture content of cassava was 72.46%, exhibiting low variability (SD = 4.33), indicating that roots were typically harvested at similar maturity stages. Nonetheless, significant variation was noted in the delay from harvest to processing (mean = 12.09 hours, SD = 6.62), the duration of storage (mean = 7.55 days, SD = 4.06), and the duration of drying (mean = 53.02 hours, SD = 24.22). Post-harvest losses were estimated to vary significantly between 5 and 60 kg, averaging at 32.43 kg. This variation highlights the diversity of production and management practices among various farming operations, thereby supporting the use of machine learning models that can identify nonlinear and interactive effects.

Table 3 outlined the ANN model's predictive performance during its initial assessment. The model reached an R^2 value of 0.9341, demonstrating that it accounted for 93.41% of the variability in post-harvest losses. The mean squared error (MSE) was minimal (1.4236), indicating that predictions were well-matched with actual values. This degree of precision highlights the appropriateness of ANN models for modeling nonlinear and multivariate relationships in agricultural systems, where conventional linear models frequently fall short. Table 4 converted regression coefficients into practical intervention strategies. For instance, each one-unit rise in cassava moisture resulted in an additional 0.25 kg of losses, indicating that harvesting at the right maturity or pre-drying roots prior to grating may reduce losses. In the same way, each extra hour of delay from harvest to processing raised losses by roughly 0.32 kg, emphasizing the need to reduce delays via pre-arranged processing. Pest infestation caused over 3.5 kg in losses, highlighting the critical necessity for sealed storage and pest control measures. On the other hand, mechanical drying decreased losses by an average of 2.2 kg, highlighting its potential as a groundbreaking post-harvest technology. These results align with, which highlighted mechanized drying as a method to improve product longevity.

Table 5 presented the outcomes of scenario simulations based on ANN. The baseline condition forecasted an average reduction of 9.60 kg. The use of mechanical drying minimized losses to 7.40 kg, and cutting the harvest-to-processing delay by 50% further decreased losses to 6.20 kg. The combined intervention, which included mechanical drying, shorter delays, and enhanced moisture management, resulted in the greatest decrease, reducing losses to 4.10 kg—a drop of over 50% from the original amount. This simulation showcases the practical value of the ANN model in measuring the effects of combined interventions, advancing from mere prediction to providing decision support for farmers and policymakers [17].

Table 6 offered a different assessment of the ANN model, which yielded unsatisfactory results, showing an MSE of 272.32 and a negative R^2 value (–0.0778), suggesting it performed worse than a mean-based predictor. This difference indicates that the model's effectiveness is very dependent on hyperparameter adjustments, data preparation, and the representativeness of samples. It highlights the significance of iterative model refinement and the possible dangers of overfitting or underfitting when dealing with agricultural datasets that frequently include noise and variability.

Table 7 detailed the structure and hyperparameter settings of the artificial neural network (ANN) used in this research. The multilayer perceptron was structured with three hidden layers containing 64, 32, and 16 neurons, respectively, all employing the rectified linear unit (ReLU) activation function. The Adam optimizer, featuring a fixed learning rate of 0.001, was selected for effective gradient descent, and training occurred for up to 2000 iterations [18]. The lack of early stopping in the initial setup indicates a focus on maximizing learning iterations, even though this raises the likelihood of overfitting. The batch size was configured to 64, striking a balance between convergence rate and computational efficiency. These parameters correspond to optimal strategies in deep learning for tabular data, as moderately deep structures paired with ReLU activation and Adam optimization have demonstrated stable convergence. Table 8 displayed the comparative analysis of the ANN versus baseline models, utilizing four performance metrics: mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R^2. The findings illustrate the ANN's exceptional predictive capability, attaining the minimum MAE (1.65) and RMSE (2.95), along with the highest R^2 (0.92). Random Forest was in second place, achieving an R^2 of 0.89 and an RMSE of 3.25, whereas Decision Tree and Linear Regression had lower performance (R^2 = 0.82 and 0.78, respectively) [19].

These results highlight how ANN models can more effectively represent nonlinear, multivariate interactions in agricultural systems compared to linear and single-tree approaches. Significantly, the enhancement of the ANN compared to Random Forest, albeit slight, emphasizes its promise when datasets are adequately representative and properly preprocessed. Similar research in agricultural forecasting also indicates that ANN models surpass Random Forest and regression, albeit with slight improvements, implying that ensemble tree methods continue to serve as robust benchmarks.

Table 9 presented the impact assessment framework, connecting model forecasts to real-world results in the cassava value chain. Four primary impact areas were recognized: minimizing waste, boosting income, empowered decision-making, and embracing technology. The framework converts technical performance into measurable socio-economic advantages by quantifying anticipated reductions in loss. Reducing waste enhances food security directly, while higher income boosts livelihood resilience for smallholder farmers. The focus on informed decision-making acknowledges the importance of predictive analytics in enhancing harvest timing, storage, and drying. Ultimately, the rates of technology adoption act as an indicator for scalability and enduring sustainability. This multidimensional impact perspective aligns with recent appeals for AI research in agriculture to extend past accuracy metrics and include wider development results [20].

Table 10 presented a specific instance of the possible advantages of utilizing the ANN-based prototype tool. Under baseline conditions, farmers faced an average weekly loss of 40 kg, which the tool mitigated to 15 kg, indicating a 62.5% reduction. In line with this, weekly earnings rose from ₦20,000 to ₦28,000, reflecting a 40% gain. Moreover, the precision of decisions transitioned from poor manual assessments to robust AI-assisted forecasts. These findings not only emphasize the model’s ability to yield quantifiable economic benefits but also indicate wider behavioral shifts, including increased use of mechanical drying and shorter intervals between harvesting and processing. Similar intervention-oriented AI tools in managing crop yield and pest control have demonstrated comparable enhancements in decision-making precision and financial gains. The results presented here, however, broaden this body of work by focusing on post-harvest losses a field that is relatively less investigated than yield prediction.

13. Comparison with Similar Research

The regression findings align closely with previous research in the cassava value chain. For example, and recorded how moisture and postponed processing speed up microbial spoilage, whereas recognized pest infestation as a major factor contributing to losses in stored garri. The current results reinforce this evidence by measuring effect sizes and connecting them to practical interventions.

Regarding predictive modeling, the ANN's elevated R^2 (Table 3) outperforms the linear regression models commonly documented in agricultural loss research, which frequently attain R^2 values under 0.80 because of the linearity assumption. The intervention simulation (Table 5) likewise illustrates the growing trend of incorporating AI models into scenario analysis, a practice highlighted by in their research on forecasting crop yields.

Nonetheless, the variable performance noted in Table 6 reflects the conclusions of, who stated that ANN effectiveness in agriculture fluctuates significantly due to dataset quality, feature preprocessing, and architectural decisions. This indicates that although ANN models are promising, they require rigorous validation against strong baselines like Random Forest and XGBoost, which have shown dependable accuracy in various agricultural prediction scenarios.

Overall, the findings indicated that ANN models, when adequately adjusted and trained, can surpass conventional statistical techniques and deliver not just precise forecasts but also useful insights for intervention strategies. However, they emphasize the importance of methodological precision to guarantee generalizability and reproducibility in various agricultural settings.

The architectural and performance results shown in Tables 7 and 8 align with international research trends in agricultural AI. Similar to earlier studies, this research shows that although Random Forests are reliable, deep neural networks can realize gradual performance improvements with precise tuning. Tables 9 and 10 highlight a unique contribution: the combination of predictive modeling with the design of interventions and the evaluation of deployment impact. Although previous studies have typically focused on reporting model accuracy, the current research connects predictions to measurable advantages for farmers, advancing actionable AI for sustainable farming.

14. Intervention and Policy Implications

The results of this research have evident consequences for intervention strategies and agricultural policies in sub-Saharan Africa. The regression and ANN models consistently emphasize moisture management, delay from harvest to processing, storage time, pest issues, and drying technique as the key factors affecting post-harvest losses in garri production. Implementing these technical insights demands focus on scalability, affordability, farmer acceptance, and enduring sustainability. A key measure arising from the findings is the endorsement of mechanical dryers. The models show that mechanical drying decreases average losses by over 2 kg per batch, and scenario simulations imply even larger reductions when it is paired with minimized delays and enhanced moisture control. Nonetheless, the scalability of mechanical dryers is limited by their upfront capital expense, which recent field reports estimate to be between ₦500,000 and ₦1.5 million per unit based on size and technology. For numerous smallholder farmers, such investments are unaffordable without cooperative financing models, public subsidies, or rural infrastructure programs backed by donors. Consequently, policy actions ought to focus on drying centers at the community level, credit initiatives, or public–private collaborations to guarantee widespread access and fair adoption.

Another important aspect is the behavior of farmers regarding adoption. Although ANN-based forecasts and intervention strategies offer technical guidance, implementation relies on farmers' readiness and capacity to alter established methods. Obstacles like insufficient awareness, restricted extension assistance, or skepticism towards "black box" AI systems could impede adoption. This highlights the significance of integrating predictive models into accessible platforms, like mobile apps, that offer straightforward, understandable guidance. Combining with current agricultural extension services and mobile advisory systems can improve usability. An app that notifies farmers to harvest cassava within a crucial timeframe or suggests mechanical drying on humid days would translate the model's results into daily choices.

Aside from garri production, the applicability of the ANN framework is encouraging. The identical methodological approach may be utilized for other cassava-derived items like fufu, lafun, or premium cassava flour, where moisture behavior, drying techniques, and storage conditions also influence losses and product quality. In a broader context, the framework can apply to additional root and tuber crops like yam, sweet potato, and potato, which experience similar post-harvest risks related to pests, microbial deterioration, and drying inefficiencies. Modifying the input feature set to align with crop-specific traits enables the model to serve as a transferable decision-support tool for various value chains.

At the policy level, this suggests that governments and development organizations ought to regard AI-driven post-harvest tools not as standalone projects but as flexible systems integrated into wider food security plans. Investing in digital infrastructure, training for farmers, and inclusive design methods can guarantee that AI-based initiatives are both technically effective and socially as well as economically sustainable. By implementing such approaches, models like the one created in this context can minimize food waste, enhance rural earnings, and bolster the resilience of food systems across various crops and areas.

15. Conclusion and Future Work

This research created and assessed an artificial neural network (ANN) model to forecast post-harvest losses in garri production, utilizing a dataset that included agronomic and operational factors. The regression analysis validated that moisture content of cassava, delay from harvest to processing, duration of storage, pest presence, and drying technique were important statistical factors influencing losses. The ANN model, when finely adjusted, surpassed conventional methods like linear regression, decision trees, and random forests, reaching an R^2 of 0.92 and reduced error rates on various metrics. Crucially, scenario simulations showed that combined interventions involving mechanical drying, minimized delays, and enhanced moisture management could lessen losses by more than 50%, directly affecting food security and farmers' earnings.

The study adds to the expanding literature on AI uses in agriculture by connecting prediction and intervention. In contrast to previous research that mainly concentrated on predicting yields or classification activities, this study emphasizes the importance of machine learning in tackling inefficiencies after harvest. Additionally, by integrating the model into a prototype decision-support system, the research illustrates ways for practical application, connecting AI analysis to decisions made by farmers and policymakers.

Even with these contributions, numerous limitations persist. The dataset, though comprehensive, is limited geographically and might not entirely reflect seasonal or regional differences in post-harvest conditions. The performance of the model was found to be influenced by hyperparameter tuning and preprocessing decisions, indicating that achieving reproducibility across various datasets might necessitate additional optimization. Moreover, obstacles to adoption such as the expense of mechanical dryers, restricted digital literacy, and infrastructural limitations could hinder the widespread implementation of AI-driven tools.

Consequently, future studies should focus on multiple avenues. To begin with, broadening the dataset across various regions and production systems will improve model generalizability. Secondly, incorporating sophisticated ensemble techniques (such as XGBoost and LightGBM) along with hybrid frameworks (like CNN–ANN models for merging tabular and image data) might enhance predictive reliability. Third, upcoming research ought to emphasize explainable AI methods, like SHAP or LIME, to offer farmers and policymakers comprehensible insights into model results. Fourth, conducting pilot tests of the prototype tool with smallholder farmer cooperatives will be essential for confirming usability, economic viability, and lasting effects. Ultimately, broadening the framework to include other cassava products and root vegetables could establish the model as a scalable decision-support tool for more effectively minimizing post-harvest losses.

In summary, the research shows that artificial neural networks have considerable potential to revolutionize post-harvest management within cassava value chains. By integrating predictive accuracy with feasible intervention strategies, such models can be crucial in minimizing food waste, enhancing farmer earnings, and promoting sustainable food security. The future obstacles consist not only of enhancing technical efficiency but also of guaranteeing scalability, inclusiveness, and acceptance at both the agricultural and policy stages.

Authors' Statement

Author Contributions. All authors contributed substantially to the conception, design, and execution of the study. Lead Author, Idowu Olugbenga Adewumi conceptualized the research framework, curated the dataset, and developed the artificial neural network (ANN) model. All other authors were involved in performing statistical analyses, including regression benchmarking and intervention simulations. Co-Author's contributed to the literature review, interpretation of results, and drafting of the discussion section. All authors reviewed, edited, and approved the final manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The authors declare no conflicts of interest, financial or otherwise, that could have influenced the outcomes of this study.

Ethical Approval: The study involved both primary and secondary data on cassava post-harvest practices and did not include human or animal subjects. Hence, ethical approval was not required.

Data Availability: The datasets used and analyzed during this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors acknowledge the support of local cassava farmers and processors who provided contextual insights into field practices, which informed the design of intervention scenarios.

References

1. Ayalde, D. G., Londoño, J. C. G., Mosquera, A. Q., Melendez, J. L. L., Gimode, W., Tran, T., & Selvaraj, M. G. (2024). AI-powered detection and quantification of post-harvest physiological deterioration (PPD) in cassava using YOLO foundation models and K-means clustering. *Plant Methods, 20*(1), 178.
2. Opara, I. K., Opara, U. L., Okolie, J. A., & Fawole, O. A. (2024). Machine learning application in horticulture and prospects for Predicting Fresh produce losses and Waste: a review. *Plants, 13*(9), 1200.
3. Guo, X., Soethoudt, H., Kok, M., & Axmann, H. (2025). Predicting Postharvest Food Losses at National and Sub-National Levels Using Data-Driven and Knowledge-Based Neural Networks. *Sustainability, 17*(10), 4552.
4. Opara, I. K., Silue, Y., Opara, U. L., & Fawole, O. A. (2025). Forecasting of postharvest fresh produce waste at the wholesale level using time series models. *Agriculture & Food Security, 14*(1), 24.
5. Rejeb, A., Rejeb, K., & Hassoun, A. (2025). The impact of machine learning applications in agricultural supply chain: a topic modeling-based review. *Discover Food, 5*(1), 141.
6. Jimoh, L. O., Kehinde, A. L., Akintunde, O. K., & Olaoye, H. O. (2025). The dynamics of postharvest loss management among cassava farmers in the Iwo ADP Zone, Osun state in Nigeria. *Journal of Agribusiness and Rural Development*, (1 [75]).
7. Ekinci, E., & Uluisik, S. (2025). Artificial Intelligence and Machine Learning Applications in the Postharvest Storage of Tomato. *Asian Journal of Agricultural and Horticultural Research, 12*(2), 305-312.
8. Ssajjabbi, V., Sseruyange, J., Ssentamu, J. D., & Mulindwa, L. (2024). Analysis of Post-Harvest Losses in Cassava Value Chain. Causes and Strategies to reduce them in Uganda. *Tanzanian Economic Review, 14*(2), 45-67.
9. Mohankumar, P., and E. T. Venkatesh. (2024). "Optimizing

Post-Harvest Efficiency in Agriculture through IoT and Machine Learning: A Random Forest Approach." *Nanotechnology Perceptions*.
10. Abiodun, Lois Omolola; Opeyemi Adeniyi Oyelade; Yinka Segun Ademiluyi; et al. (2024). "Overcoming the Problems Facing Cassava Processing Industry in Nigeria." BMTP 1, no. 1
11. Mohankumar, P., and E. T. Venkatesh (2024). "Optimizing Post-Harvest Efficiency in Agriculture through IoT and Machine Learning: A Random Forest Approach." Nanotechnology Perceptions.
12. NYS Sarpong. (2024). "A Machine Learning-Driven Modeling and Optimization Approach for Enhancing Cassava Mash Production Quality in Cassava Graters." Engineering Applications of Artificial Intelligence / ScienceDirect (September 2024).
13. "Recent Trends in the Pre-Drying, Drying, and Post-Drying Processes for Cassava Tuber: A Review." (2025). Drying Technology & Food Processing Reviews (review article; ResearchGate preprint). (Review of drying technologies and implications for preserving processed cassava quality).
14. "Reduction of Post-Harvest Losses Is Crucial for a Successful Cassava Value Chain." (2024). Open Agriculture Journal (special issue on postharvest). (Policy-oriented review of PHL mitigation strategies).
15. "(Title: The Dynamics of Postharvest Loss Management among Cassava Farmers in Iwo ADP Zone, Osun State)." (2025). Journal of Agribusiness and Rural Development 75, no. 1 (February 2025): 78–89. (Empirical study of PHL determinants and management practices in Nigeria).
16. "Benefits and Processing Technologies of Gari, a Famous Indigenous Food of Nigeria." (2025). Food Processing & Technology Review (Review/Article; ResearchGate preprint). (Comprehensive review of gari processing steps, losses, and technologies).
17. "Causes and Management of Postharvest Losses in Nigeria: An Overview." (2025). Journal of Postharvest Research (systematic overview; ResearchGate).
18. "Postharvest Fresh Produce Waste Forecasting at the Wholesale Level Using Time-Series and ML Approaches." (2025). Agriculture & Food Security 14 (24).
19. "Post-Harvest Physiological Deterioration (PPD) Studies and MePOD12 Regulatory Role." (2024–2025). Postharvest Biology and Technology (review/experimental articles on cassava PPD and physiological markers).

Appendices

To Design and train an ANN architecture tailored to the dataset's characteristics

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
import tensorflow as tf
# Drop non-numeric columns and set up features/labels
X = df.drop(columns=["Sample ID", "Date", "Estimated Loss (kg)"])
y = df["Estimated Loss (kg)"]
# Identify categorical and numeric features
categorical_features = ["Transportation Mode", "Drying Method", "Pest Infestation (Y/N)"]
numeric_features = [col for col in X.columns if col not in categorical_features]
# Preprocessing pipeline
preprocessor = ColumnTransformer(
    transformers=[
        ("num", StandardScaler(), numeric_features),
        ("cat", OneHotEncoder(), categorical_features)
    ]
)

# Fit and transform the data
X_processed = preprocessor.fit_transform(X)

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X_processed, y, test_size=0.2, random_state=42)

# Build ANN model
model = tf.keras.Sequential([
    tf.keras.layers.Dense(64, activation='relu', input_shape=(X_train.shape[1],)),
    tf.keras.layers.Dense(32, activation='relu'),
    tf.keras.layers.Dense(1)  # Regression output
])

# Compile the model
model.compile(optimizer='adam', loss='mean_squared_error', metrics=['mae'])

# Train the model
history = model.fit(X_train, y_train, epochs=10, batch_size=64, validation_split=0.2, verbose=1)

# Evaluate the model on the test set
test_loss, test_mae = model.evaluate(X_test, y_test)

test_loss, test_mae
```

To Evaluate the ANN's predictive accuracy against traditional regression models.

Evaluation Plan

- **Models to Compare**
- **Deep Neural Network (ANN)**
- **Linear Regression**

- **Decision Tree Regressor**
- **Random Forest Regressor**

➢ **Evaluation Metrics**

- **Mean Absolute Error (MAE)**
- **Mean Squared Error (MSE)**
- **Root Mean Squared Error (RMSE)**
- **R² Score (Coefficient of Determination)**

➢ **Steps**

1. Preprocess the data (scaling + encoding).
2. Split into training and test sets.
3. Train each model.
4. Predict on test data.
5. Compare performance using metrics.

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, OneHotEncoder
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

from sklearn.linear_model import LinearRegression
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
import tensorflow as tf

# Load dataset
df = pd.read_excel("garri_phl_dataset.xlsx")

# Features and target
X = df.drop(columns=["Sample ID", "Date", "Estimated Loss (kg)"])
y = df["Estimated Loss (kg)"]

# Preprocessing
categorical_features = ["Transportation Mode", "Drying Method", "Pest Infestation (Y/N)"]
numeric_features = [col for col in X.columns if col not in categorical_features]

preprocessor = ColumnTransformer([
    ("num", StandardScaler(), numeric_features),
    ("cat", OneHotEncoder(), categorical_features)
])

X_processed = preprocessor.fit_transform(X)
X_train, X_test, y_train, y_test = train_test_split(X_processed, y, test_size=0.2, random_state=42)

# ----------- ANN Model ------------
model = tf.keras.Sequential([
      tf.keras.layers.Dense(64, activation='relu', input_shape=(X_train.shape[1],)),
    tf.keras.layers.Dense(32, activation='relu'),
    tf.keras.layers.Dense(1)
])
model.compile(optimizer='adam', loss='mse', metrics=['mae'])
model.fit(X_train, y_train, epochs=10, batch_size=64, validation_split=0.2, verbose=0)
y_pred_ann = model.predict(X_test).flatten()

# ----------- Traditional Models ------------
models = {
    "Linear Regression": LinearRegression(),
    "Decision Tree": DecisionTreeRegressor(random_state=42),
    "Random Forest": RandomForestRegressor(n_estimators=100, random_state=42)
}

results = {
    "Model": [],
    "MAE": [],
    "MSE": [],
    "RMSE": [],
    "R² Score": []
}

# Evaluate traditional models
for name, reg in models.items():
    reg.fit(X_train, y_train)
    y_pred = reg.predict(X_test)
    results["Model"].append(name)
    results["MAE"].append(mean_absolute_error(y_test, y_pred))
    results["MSE"].append(mean_squared_error(y_test, y_pred))
      results["RMSE"].append(np.sqrt(mean_squared_error(y_test, y_pred)))
    results["R² Score"].append(r2_score(y_test, y_pred))

# Evaluate ANN
results["Model"].append("Deep Neural Network")
results["MAE"].append(mean_absolute_error(y_test, y_pred_ann))
results["MSE"].append(mean_squared_error(y_test, y_pred_ann))
results["RMSE"].append(np.sqrt(mean_squared_error(y_test, y_pred_ann)))
results["R² Score"].append(r2_score(y_test, y_pred_ann))

# Convert results to DataFrame
eval_df = pd.DataFrame(results)
print(eval_df)
```