A Decision Tree Regression Approach for Forecasting Rice and Corn Crop Yields in Nueva Ecija, Philippines

Michaela Angela E. Cailing
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
cailingme@students.national-u.edu.ph

Aaron Eldreich L. Chua
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
chuaal@students.national-u.edu.ph

Danyssa B. Tamayo
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
tamayodb@students.national-u.edu.ph

***Abstract*—This study presents a Decision Tree Regression (DTR) model designed to predict rice and corn yields in Nueva Ecija, Philippines, using key environmental variables, including temperature, rainfall, and humidity. The model demonstrated strong predictive accuracy, with a final R-squared value of 89.87% after hyperparameter tuning, effectively capturing yield patterns based on climatic factors. A web application was developed using Anvil to deploy the model, providing users—such as farmers and agricultural stakeholders—with yield predictions based on specific environmental data. With a Mean Absolute Percentage Error (MAPE) of 7.36% for rice yields and 16.56% for corn, the tool offers actionable insights for crop planning and resource allocation. Recommendations for enhancing the model include expanding the dataset, incorporating additional environmental variables, and exploring ensemble methods. By empowering users with accurate and accessible forecasting capabilities, this study contributes to more informed agricultural decision-making, addressing food security and economic stability in Nueva Ecija and potentially beyond.**

Index Terms—Crop Yield Prediction, Decision Tree Regression, Agricultural Forecasting, Climate Impact, Web-Based Predictive Tool.

I. Introduction

Agricultural productivity is significantly influenced by environmental factors, particularly weather conditions, which vary across regions and seasons. In the Philippines, agricultural land covers approximately 12.3 million hectares, accounting for about 41% of the country's total land area. Over the last five years, the gross value added (GVA) of the agriculture sector reached ₱1.78 trillion, representing 8.9% of the gross domestic product (GDP) in 2022 [1]. This sector employs around 32% of the economically active population, highlighting its importance to the national economy [2].

However, the Philippines is particularly vulnerable to climate change due to its geographical location and archipelagic formation. The country ranks highest globally in terms of vulnerability to tropical cyclones, and fourth among countries most affected by extreme weather phenomena. The agricultural sector suffers disproportionately from climate-related disasters, with damages to agricultural production from 1990 to 2006 caused by typhoons (70%), droughts (18%), and floods (5%) [2].

The Food and Agriculture Organization has identified several factors contributing to increased crop yields, including the restoration of irrigation facilities, favorable weather, improved fertilization, and the use of high-quality seeds [3]. Nevertheless, statistics indicate a 2.6% decline in harvest in 2021 due to unfavorable weather conditions [4], underscoring the urgent need for reliable forecasting tools. In Nueva Ecija, the leading rice-producing province in the Philippines, rice and corn are staple crops crucial for food security and economic stability. Yet, local farmers face significant challenges in predicting crop yields and determining the most suitable crops to cultivate, particularly amidst fluctuating weather patterns and climate-related disasters.

Recent advances in machine learning (ML) have transformed agricultural forecasting, providing tools that can analyze complex environmental data and make specific predictions often missed by traditional methods. Decision Tree Regression (DTR), a popular ML model, is recognized for its ability to dissect data into meaningful subsets, helping identify patterns in how weather conditions impact crop yields. Studies show that, despite simpler ML models like DTR being less computationally demanding than ensemble models (e.g., Random Forest), they maintain high interpretability and practicality. This is particularly advantageous for farmers in resource-limited regions like Nueva Ecija, where ease of use is essential, and access to advanced technology is limited. [5], for instance, demonstrated DTR's effectiveness in yield prediction, noting that while ensemble models slightly outperform DTR in accuracy, DTR's accessibility and interpretability make it an ideal choice for regions with limited technological infrastructure.

This study aims to develop a DTR model specifically designed to predict rice and corn crop yields in Nueva Ecija, focusing on weather conditions. By providing data-driven insights into which crop is best suited for cultivation, the model aims to assist farmers in optimizing production and mitigating the risks posed by erratic weather. Addressing the limitations found in both traditional statistical models and complex ML applications, this research emphasizes the need for a practical, accessible forecasting tool. The resulting model will serve as an actionable resource for Nueva Ecija's farmers, helping them make more informed planting decisions and contributing to improved food security and economic resilience in the region.

II. Review of related literature

Several studies have explored the use of machine learning in agriculture, especially for crop yield prediction, due to its ability to analyze complex environmental data and make specific predictions that traditional methods often overlook. Decision Tree Regression (DTR) is one of the popular machine learning models used in crop yield forecasting. This model is known for dividing data into distinct subsets, allowing it to recognize meaningful patterns in weather variables like temperature, humidity, and rainfall. By providing clear insights into how specific environmental factors affect yield, DTR offers a more tailored approach compared to traditional methods, which typically rely on general historical trends [6]. Unlike older prediction models, DTR can identify nonlinear relationships between variables, making it effective for understanding how seasonal weather conditions in Nueva Ecija, Philippines, influence rice and corn yields.

Studies comparing DTR with other machine learning models, such as Random Forest and Gradient Boosting, reveal that although ensemble models may offer slightly higher accuracy, DTR remains advantageous due to its simplicity and lower computational demands. For instance, [5] found that while Random Forest provided slightly better accuracy in their crop yield predictions, DTR's interpretability and lower resource requirements made it a more practical choice for farmers, especially those with limited access to advanced technology. This characteristic is particularly relevant for farmers in regions like Nueva Ecija, where accessibility and ease of use are crucial factors.

Previous research also indicates that machine learning models outperform traditional methods in capturing the complex patterns essential for accurate yield forecasting. Traditional statistical models, such as multiple linear regression (MLR) and autoregressive models, are easy to understand and effective with small datasets. However, these methods often struggle to detect complex relationships in agricultural data, especially when multiple variables like temperature, rainfall, and wind speed interact [7]. Machine learning models, on the other hand, can analyze a broader range of factors, improving prediction accuracy. DTR and similar models like Random Forest have become increasingly popular in agricultural research for their ability to process various inputs and adapt to different environmental conditions, which is particularly valuable in regions with diverse weather patterns [8].

Despite these advances, there are limitations in current machine learning applications in agriculture, particularly concerning accessibility and adaptability for local farmers. For example, while companies like Climate Corporation and John Deere have developed advanced crop prediction tools using satellite and weather data, these systems are often costly and require technical expertise, making them difficult for small-scale farmers to use [9]. In the Philippines, specifically, models that have been developed to predict yields, such as those by [10], have struggled with issues like inconsistent real-time weather data, which reduces prediction accuracy across different seasons. This research addresses these gaps by developing a more accessible DTR model focused on Nueva Ecija's climate and agricultural needs, specifically for rice and corn crops.

The approach in this study contributes to the field by simplifying crop yield predictions, making data-driven crop selection more feasible for farmers in resource-limited regions. By using DTR, the study aims to balance prediction accuracy with ease of use, addressing the limitations of high-complexity models. Many of the models used in past research, including Random Forest and neural networks, although accurate, are computationally demanding and harder to interpret, posing challenges for farmers without advanced technical resources [11]. DTR's interpretability and lower demand on resources make it a practical solution for Nueva Ecija farmers, providing them with clear, actionable recommendations on which crop—rice or corn—should be prioritized for planting in each quarter based on predicted yields.

The focus on DTR in this research aligns with current trends in agriculture, where there is an increasing emphasis on accessible and user-friendly machine learning tools that can help farmers make better decisions. While many assume that complex models are necessary for accurate predictions, this study challenges that notion by demonstrating that simpler, interpretable models like DTR can still offer significant benefits. By tailoring the model to the specific weather conditions of Nueva Ecija, this research aims to provide a tool that is not only accurate but also practical and beneficial for local farmers, helping them improve yield outcomes and make more informed planting decisions.

III. Methodology

The methodology for crop yield forecasting is illustrated in **Figure 1.** It involves four stages: data collection (gathering environmental and crop yield data), data pre-processing (cleaning and normalizing the data), model training and optimization (training the Decision Tree Regression model and tuning parameters), and evaluation and validation (assessing model performance using metrics such as R^2, MAE, and MSE).

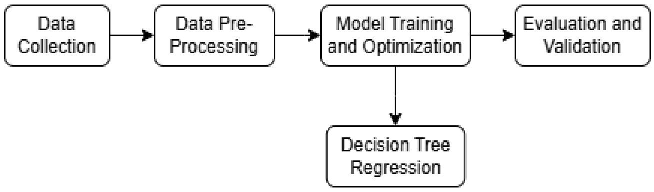


Figure 1. *Crop Yield Forecasting Framework*

A. Data Collection

The data used in this study includes four datasets collected and published by the Philippine Statistics Authority (PSA), which utilizes standardized data gathering techniques, including surveys, field reports, and monitoring stations in various Philippine regions, covering both crop yield and environmental data for Nueva Ecija. The primary agricultural dataset, Palay and Corn: Volume of Production in Metric Tons by Ecosystem/Crop Type, by Quarter, by Semester, by Region and by Province, 1987-2024, provides quarterly production volumes (metric tons) of rice (palay) and corn, with entries specifying the crop type, year, and quarter. Environmental datasets include Amount of Rainfall by Monitoring Station, 2014-2023, Temperature by Monitoring Station, 2014-2023, and Relative Humidity by Monitoring Station, 2014-2023. Each of these environmental datasets records monthly data by station, listing rainfall (in mm), maximum, minimum, and mean temperatures (°C), and relative humidity (percentage). To create a single, unified dataset for the study, these monthly values will be aggregated quarterly and combined with the production data, allowing for a Decision Tree Regression model to predict yields based on weather patterns.

This unified dataset contains information on crop production along with relevant environmental and temporal factors. It includes the following variables:

- crop_type: indicates the type of crop, with "0" representing palay and "1" representing corn
- year: records the calendar year of data collection
- quarter: identifies the part of the year (1–4)
- production: the target variable, measuring the quantity of crop produced in units (tons)
- rainfall: the total rainfall during the growing period (in millimeters)
- max_temp and min_temp: the maximum and minimum temperatures (in degrees Celsius)
- mean_temp: the average temperature for the period
- relative_humidity: the average humidity during the growing season, expressed as a percentage

B. Data Pre-Processing

To prepare the unified dataset for the Decision Tree Regression model, several pre-processing techniques were applied to ensure data accuracy and relevance for yield prediction. The following steps outline the data pre-processing techniques used:

1. **Data Cleaning**: Datasets were reviewed for missing or duplicate entries, and none were found, ensuring consistency and reliability.
2. **Transformation**: Monthly environmental data was aggregated into quarterly data to align with crop yield records, using averages for temperature, rainfall, and humidity.
3. **Encoding Categorical Variables**: Crop type was encoded as 0 for rice and 1 for corn to simplify processing.
4. **Feature Scaling**: Numerical features were normalized to enhance model performance and training speed.

C. Experimental Setup

The Decision Tree Regression model was implemented using the following tools and libraries: **Pandas** for data manipulation, **NumPy** for numerical operations, **Matplotlib** and **Seaborn** for data visualization, and **Scikit-Learn** for model training, hyperparameter tuning, and evaluation. **Anvil** was used for model deployment, providing user interaction through server connections. The experiments were conducted in **Google Colab**, leveraging cloud-based GPU and CPU resources for efficient computation.

Data Splitting and Validation: The dataset was split into 80% training and 20% testing using Scikit-Learn's train_test_split. A 5-fold cross-validation with GridSearchCV was applied to optimize hyperparameters based on Mean Squared Error (MSE).

Evaluation Metrics: Model performance was assessed using **Mean Absolute Error (MAE)**, **Mean Squared Error (MSE)**, and **R^2 Score** to capture prediction accuracy, error magnitude, and variance explained.

Hyperparameters: Key hyperparameters optimized included:

- **max_depth**: Controls tree depth to prevent overfitting.
- **min_samples_split**: Sets the minimum samples needed for a node split.
- **min_samples_leaf**: Specifies the minimum samples required at each leaf node.
- **max_features**: Determines the number of features for each split.
- **max_leaf_nodes**: Limits the number of leaf nodes for regularization.

D. Algorithm

A Decision Tree is an intuitive and flexible machine learning model used for both classification and regression tasks. Decision Trees operate by constructing a flowchart-like structure where each internal node represents a decision based on a particular feature, each branch represents the outcome of a decision, and each leaf node provides the final prediction or classification result. The Decision Tree model is built through a recursive process of selecting features that split the data at each node, aiming to produce the most accurate predictions.

In this study, Mean Squared Error (MSE) is the primary metric used to guide the splitting of nodes in the Decision Tree. MSE calculates the average squared difference between actual values and predicted values in regression tasks. By choosing splits that minimize MSE, the model improves its prediction accuracy by grouping data points with similar characteristics. Equation 1 shows the formula for MSE.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

Equation 1. *Mean Squared Error*

Where:

- Y_i is the actual value for the i-th observation.
- $\hat{Y}_i$ is the predicted value for the i-th observation.
- n is the total number of observations.

Decision Trees are widely applied in fields such as medical diagnosis, credit scoring, and fraud detection due to their interpretability and transparency. Each path from the root to a leaf node forms a distinct decision rule, making Decision Trees especially valuable in applications where model transparency and simplicity are crucial. Their adaptability and ease of visualization make Decision Trees a practical choice in a variety of predictive modeling applications.

E. Training Procedure

A 5-fold cross-validation strategy was used to enhance model robustness, with data split into five parts, rotating each fold as the validation set. This approach averaged results across folds to provide reliable performance estimates and reduce overfitting. Hyperparameters such as max_depth, min_samples_split, and min_samples_leaf were optimized via GridSearchCV, using Mean Squared Error (MSE) as the scoring metric. This systematic tuning improved model accuracy and generalizability for crop yield predictions.

F. Evaluation Metrics

To assess the model's effectiveness, **Mean Absolute Error (MAE), Mean Squared Error (MSE),** and the **R^2 Score** were utilized, each chosen for the unique way they capture different aspects of model accuracy and fit. MAE reveals the typical error size, offering a straightforward look at prediction accuracy. MSE, by focusing on squared errors, puts extra weight on larger prediction mistakes, helping to reduce significant errors. The R^2 Score, meanwhile, measures the proportion of variance that the model can explain, which indicates its fit compared to a basic mean-prediction baseline.

For a more reliable assessment, **Cross-Validation with MSE** was used to test consistency across data folds, showing an average MSE of about 1.24×10^{11}. These metrics together give a well-rounded view of performance, with MAE and MSE highlighting general and extreme errors, and R^2 assessing overall data fit.

G. Baselines and Comparative Model

The project utilized Decision Tree Regression as the primary model for evaluation. While it demonstrated solid performance in predicting crop production, there is potential to enhance the analysis by exploring additional regression techniques, such as Linear Regression, Random Forest Regression, and Gradient Boosting. These models may improve upon the performance metrics achieved with Decision Tree Regression. Comparing these alternative models against the Decision Tree baseline would help validate the findings and potentially capture the complexities of the data more effectively, leading to greater accuracy and robustness in predictions.

IV. Results and Discussion

Table I. *Performance metrics of the DTRr at an 80:20 train-test ratio.*

80:20 Train Test Split		
MAE	**MSE**	**R-Squared**
62304.7224	13826767703.74	0.7966

These outcomes were achieved without any hyperparameter adjustments, demonstrating an initial accuracy in crop yield predictions. The MAE of 62,304.72 indicates that, on average, the model's predictions deviated from the actual yield by approximately this amount. The MSE of 13,826,767,703.74 reveals the overall variance in the predictions, while the R-squared value of 0.7966 suggests that the model explains around 79.66% of the variability in the crop yield data.

Table II. *Best hyperparameters and corresponding score from Grid Search*

Hyperparameter	Value
max_depth	None
max_features	sqrt
max_leaf_nodes	30
min_samples_leaf	1
min_samples_split	2
Best Score (MSE)	9508333918

These results were obtained after conducting hyperparameter tuning, highlighting the model's optimization for improved accuracy in crop yield predictions. The chosen parameters include max_depth set to None, allowing the tree to grow without restriction. The max_features value of 'sqrt' indicates that the model considers a random subset of features for each split, which can help reduce overfitting. With max_leaf_nodes set to 30, the model limits the number of leaf nodes. The min_samples_leaf of 1 signifies that a leaf node can be created with a single sample, while the min_samples_split value of 2 allows for splits even with minimal data points. The best score achieved, with a Mean Squared Error (MSE) of 9508333918 reflects the model's enhanced predictive performance after optimization, indicating a reduction in prediction variance compared to the initial model evaluation.

Table III. *Performance metrics of the DTR after hyperparameter tuning.*

5 Cross-Fold Validation | 80:20 Train Test Split		
MAE	**MSE**	**R-Squared**
62556.0633	6886308635.34	0.8987

Following hyperparameter tuning, the Decision Tree Regressor exhibited significant improvements in predictive accuracy for crop yield estimations. The Mean Absolute Error (MAE) of 62,556.06 indicates that the model's predictions now deviate from the actual yield by approximately this amount, showing a slight increase in precision. The Mean Squared Error (MSE) decreased substantially to 6886308635.34, suggesting a reduction in prediction variance and improved consistency in results. The R^2 Score improved to 0.8987, signifying that the model now explains about 89.87% of the variability in crop yield data, up from the initial 79.66%.

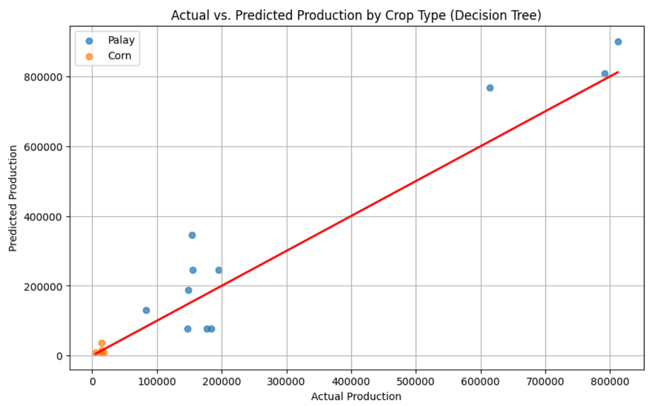


Figure 2. *Actual versus Predicted Production by Crop Type*

This plot shows the relationship between actual crop production and predicted production using a Decision Tree model, with data points representing two types of crops: Palay (rice) and Corn. The red diagonal line represents the ideal scenario where predicted values exactly match actual production values. Several blue dots (Palay observations) fall close to the red line, particularly for higher production values (around 800,000). This suggests that the model predicted these values reasonably well. However, there are also notable deviations for some Palay points, especially at lower production levels, indicating underestimation or overestimation in those cases.

The orange dots (Corn observations) mostly appear near the origin (lower production values). For Corn, the predicted values seem closer to actual values, with less spread away from the red line, suggesting relatively better accuracy for low-production Corn observations.

The Decision Tree model appears to have performed better for high-production values of Palay but shows more variability and error in the lower production ranges. The scattered nature of some points, especially for Palay, implies that the model may have overfitted to certain data points or that the features do not fully capture the factors affecting lower production levels.

In this study, we performed a random selection of 10 samples from the dataset to record their actual volume yield values. Subsequently, we employed our predictive model to estimate the volume yield for these selected samples.

The percentage error provides a measure of how accurate the predicted yield of the crop is compared to the actual yield. It is calculated using the formula:

$$\text{Percentage Error} = \frac{|\text{Actual Yield} - \text{Predicted Yield}|}{|\text{Actual Yield}|} \times 100$$

Equation 2. *Percentage Error*

A smaller percentage error indicates a more accurate prediction.

To evaluate the accuracy of our predictions, we then computed the Mean Absolute Percentage Error (MAPE) utilizing the following formula:

$$\mathrm{MAPE} = 100 \times \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

Equation 3. *Mean Absolute Percentage Error*

In this equation:

- A_t represents the actual volume yield values,
- F_t denotes the predicted volume yield values, and
- n signifies the total number of samples.

Table IV. *Yield Prediction Accuracy Assessment*

For Palay			
Prediction	**Actual Yield**	**Predicted Yield**	**% Error**
1	804937	804937	0
2	627060	745139.74	18.83
3	788956	773860	1.91
4	220702.59	233730.34	5.90
5	188512.09	169551.36	10.05
		MAPE	**7.36**
For Corn			
Prediction	**Actual Yield**	**Predicted Yield**	**% Error**
1	8847.4	7162.13	19.04
2	18201.26	16663.82	8.45
3	14049	17720.69	26.13
4	13339.7	11551.35	13.42
5	15310	17720.69	15.75
		MAPE	**16.56**

Palay Yield Predictions

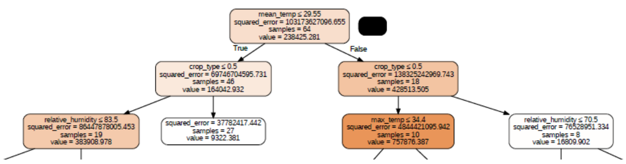
For Palay, the MAPE calculated was **7.36%**, which reflects a relatively high level of accuracy in our predictive model. The individual percentage errors for the samples ranged from **0%** to **18.83%**, suggesting that while most predictions were close to the actual values, there were instances of notable deviation, particularly in the second sample, where the predicted yield was significantly overestimated.

This relatively low MAPE indicates that our predictive model is effective for Palay, capturing the yield trends effectively. The absolute percentage errors were most prominent in the samples with larger yield values, suggesting that the model may be more sensitive to larger variations in actual yields. Additionally, the smallest error of **0%** in the first sample demonstrates that the model can achieve perfect predictions under certain conditions.

Corn Yield Predictions

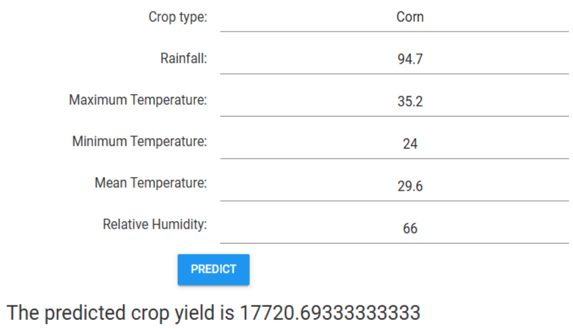
In contrast, the MAPE for Corn was found to be **16.56%**, indicating a lower level of accuracy compared to Palay. The range of percentage errors for the Corn samples was wider, with values ranging from **8.45%** to **26.13%**. Notably, the third sample exhibited the highest error percentage, where the predicted yield was substantially overestimated compared to the actual yield. This suggests that the model's performance for Corn is less reliable, possibly due to factors such as variability in yield influencing conditions or limitations in the model’s input data.

Figure 3. *Visualization of the DTR used to predict crop production based on various input features*



The analysis of the Decision Tree model indicates that mean and maximum temperatures are significant predictors of crop yield. The influence of crop type on the decision splits consistently emphasizes its critical role in determining yield outcomes. While rainfall and humidity are secondary factors compared to temperature and crop type, they nonetheless contribute to the overall production results. In summary, the Decision Tree effectively identifies mean temperature, crop type, and maximum temperature as the primary determinants of crop yield, while rainfall and relative humidity serve to refine the model's predictions.

Figure 4. *Crop Yield Prediction Application Inputs and Result*



The crop yield prediction model was deployed using Anvil, a platform that facilitates the development of web applications with ease. The user interface allows for input of key environmental variables, including crop type, rainfall, maximum and minimum temperatures, mean temperature, and relative humidity. By clicking the "PREDICT" button, users receive an estimated crop yield based on the specified inputs. This tool is particularly valuable for farmers, stakeholders, and researchers, as it provides crucial insights into anticipated crop yields for the Nueva Ecija region. Access to this information enables users to make data-driven decisions, thereby enhancing agricultural planning, optimizing resource allocation, and improving overall crop management practices.

V. Conclusion

This study developed a Decision Tree Regression (DTR) model to predict the crop yields of rice and corn in Nueva Ecija based on weather conditions. The model's integration of key environmental variables—including maximum, minimum, and mean temperatures, rainfall, and relative humidity—proved effective in generating accurate crop yield predictions. Initial model performance, as indicated by metrics such as Mean Absolute Error (MAE) and R-squared, demonstrated a strong predictive capacity, which was further improved after hyperparameter tuning. The final model explained approximately 89.87% of the variance in crop yield, reflecting a robust capacity for capturing yield patterns based on climatic factors.

The practical deployment of this model through a web application using Anvil allows for easy accessibility, enabling farmers, stakeholders, and researchers to enter specific environmental data and receive yield predictions in real-time. This tool offers valuable insights into potential crop production, empowering users to make informed decisions that enhance agricultural planning and resource allocation. For instance, the model's high accuracy for rice (Palay) yield predictions, as evidenced by a low Mean Absolute Percentage Error (MAPE) of 7.36%, indicates that it can reliably assist in crop management strategies. Although the model's performance for corn showed higher variability, with a MAPE of 16.56%, it nonetheless provides useful guidance for estimating corn yields under varying conditions.

To further improve the model's accuracy, particularly for corn, expanding the dataset to include a broader range of historical crop data across seasons and years is recommended. Additionally, incorporating variables such as soil moisture, wind speed, and solar radiation could enhance predictive precision. Exploring ensemble methods like Random Forest or Gradient Boosting could also reduce overfitting and better capture complex relationships. Once validated in Nueva Ecija, adapting the model for other agricultural regions in the Philippines could extend its benefits, supporting more effective crop planning and resource management across a wider farming community.

This study highlights the potential of a data-driven approach to mitigate the uncertainties posed by fluctuating weather patterns, thus contributing to improved food security and economic stability in Nueva Ecija. By offering a user-friendly, accessible predictive tool, this research addresses a critical need for practical forecasting in agricultural contexts, bridging the gap between traditional statistical models and complex machine learning applications. The insights generated by this model can ultimately aid farmers in optimizing crop choices and responding proactively to environmental challenges, laying the groundwork for more resilient agricultural practices in the region.

References

[1] Philippine Statistics Authority, “Agricultural Indicators System: Output and Productivity 2022,” Psa.gov.ph, 2024. Available: https://psa.gov.ph/publication/agricultural-indicators-system-output-and-productivity.

[2] Climate-Resilient Agriculture in the Philippines, “Climate-resilient agriculture (CRA) considerations,” 2022. Available: https://amia.da.gov.ph/wp-content/uploads/2018/08/CRA_Profile_Philippines.pdf.

[3] R. R. Espino and C. Atienza, “CROP DIVERSIFICATION IN THE PHILIPPINES - Rene Rafael C. Espino and Cenon S. Atienza*,” www.fao.org, 2021. Available: https://www.fao.org/4/x6906e/x6906e0a.htm.

[4] R. M. Ochave, “Agricultural output slumps in Q3 - BusinessWorld Online,” BusinessWorld Online, Nov. 08, 2021. Available: https://www.bworldonline.com/top-stories/2021/11/09/409190/agricultural-output-slumps-in-q3/.

[5] S. Sharma et al., “A Comparative Study of Decision Tree and Ensemble Methods for Yield Prediction,” *Int. J. Agr. Sci. Res.*, vol. 5, no. 2, pp. 120–130, 2023.

[6] A. Jorvekar et al., "The Impact of Decision Tree Regression on Agricultural Forecasting," *J. Agric. Informatics*, vol. 12, no. 3, pp. 50–60, 2024.

[7] C. Chlingaryan, S. Sukkarieh, and B. Whelan, “Machine Learning Approaches for Crop Yield Prediction,” *Comput. Electron. Agric.*, vol. 150, pp. 57–65, 2020.

[8] S. Khaki and L. Wang, “Predicting Corn Yields Using Deep Neural Networks,” *Agronomy Journal*, vol. 111, no. 3, pp. 1171–1185, 2019.

[9] R. Villalobos et al., “The Role of DSS and Crop Prediction Tools in Modern Agriculture,” *Int. J. Agrotechnol.*, vol. 14, no. 4, pp. 220–234, 2019.

[10] E. E. Tongson, L. A. Alejo, and O. F. Balderama, "Simulating impacts of El Niño and climate change on corn yield in Isabela, Philippines," *Climate Dynamics and Development Journal*, vol. 2, no. 1, pp. 34–46, 2024. Available: https://www.cddjournal.org/article/view/vol02-iss1-4.

[11] T. Klompenburg, A. Kassahun, and C. Catal, “Crop yield prediction using machine learning: A systematic literature review,” *Computers and Electronics in Agriculture*, vol. 177, p. 105709, 2020. Available: https://www.sciencedirect.com/science/article/pii/S0168169920302301

[12] P. Wattamwar, R. Lagrazon, and M. Tan, “Machine Learning Algorithms for Crop Yield Prediction,” in *Proc. IEEE Conf. Agric. Comput.*, 2021, pp. 132–138.

[13] “Predicting Agriculture Yields Based on Machine Learning Using Regression and Deep Learning,” IEEE Xplore. Available: https://ieeexplore.ieee.org/. Accessed: Nov. 02, 2024.

[14] M. Wang et al., “Current State of Machine Learning in Agriculture,” *J. Precision Agric.*, vol. 17, no. 2, pp. 160–173, 2024.