

# A Correlation-Aware Enhanced Stacked Ensemble Model for Multi-Output Prediction of Agricultural Crop Metrics in Oceania

Tajbir Singh<sup>1</sup>, Satinder Kaur<sup>2</sup>, Kumari Sarita<sup>3</sup> Pritpal Singh<sup>4</sup> Suman Devi<sup>5</sup>

<sup>1</sup> M.Tech Scholar, Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India

<sup>2,5</sup> Assistant Professor, Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India

<sup>3,4</sup> Research Scholar, Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India

\*\*\*

**Abstract** -Accurate and interpretable crop output forecasting remains a significant problem in agricultural analytics, especially in different agro-ecological zones like Eastern Oceania. This paper proposes the Correlation-Aware Dual-Stage Stacked Ensemble (CADESE), a new multi-output prediction framework that improves accuracy while maintaining agronomic realism. The architecture combines heterogeneous base learners Random Forest, SVR, GBM, and Feedforward Neural Networks optimized for different crop indicators and fuses their outputs using a dual-stage stacking mechanism enhanced with a correlation-aware attention layer to capture inter-variable dependencies. CADESE outperforms baseline and advanced models, with  $R^2$  values of 0.927 (Yield), 0.908 (Production), and 0.894 (Area Harvested), lowest MAPE values ( $\approx 1.3$ -1.35%), and greatest EVS scores across all goals. Domain-informed regularisation guarantees physiologically realistic predictions under a variety of environmental and agronomic situations. Beyond regression, CADESE achieves 94% classification accuracy across crop-related categories, and SHAP and LIME studies indicate that important prediction factors are aligned with agronomic knowledge, increasing transparency and confidence. These findings establish CADESE as a robust, interpretable, and deployable decision-support framework for precision agriculture, food security forecasting, and long-term resource planning in the face of climatic variability.

**Key Words:** Multi-output prediction, crop yield forecasting, stacked ensemble, attention mechanism, correlation modelling, agricultural analytics, domain-informed regularization, Eastern Oceania, interpretable machine learning, agronomic plausibility.

## 1. INTRODUCTION

Agriculture is a pillar of economic growth and food security, particularly in Oceania. Australia and New Zealand, the two regional leaders, both substantially depend on agricultural exports. In 2022–2023, Australian agricultural production was worth AUD 94 billion, of which 70% was exported (Department of Agriculture, Fisheries and Forestry, 2024), making substantial contributions to national GDP and world food supply. Equally, New Zealand's agriculture industry account

79.6% of its exports, valued at NZ\$57.4 billion [1]. These statistics highlight the importance of agriculture in the regional economy, livelihoods, and environmental management.

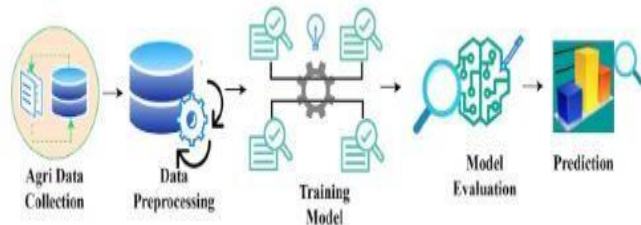
The augmenting complexity in crop production and climate volatility requires strong decision support systems. The variables involving altering precipitation, heterogeneity of soils, varying temperatures, outbreaks of pests, and the cost of inputs render agricultural prediction challenging [2]. Since conventional statistical and crop simulation models tend to lag the detection of nonlinear relationships and changing environmental patterns, there has been expanding interest in machine learning (ML)-based predictive modeling [3].

But most ML methods aim to forecast a single agricultural output (such as crop yield), without considering the intercorrelations between other interdependent variables such as nutrient needs, water consumption, and growth indicators. Further, although ensemble learning strategies like bagging, boosting, and stacking have displayed better performance in crop forecasting [4][5],[6], their application within a single framework for multi-output forecasting is rare.

Climate change also adds complexity to the forecasting of agriculture. Temperature, precipitation, and atmospheric gas fluctuations have a direct impact on productivity and quality, while indirectly making changes to pest behavior, disease rates, and soil [7][8]. Statistical models of yield responses to climate have been employed to make projections, but their extrapolation into unobserved regions or time spans is still challenging because of generalization problems and uncertainty propagation [9].

To address such challenges, this research puts forward Correlation-Aware Enhanced Stacked Ensemble Model specifically for multi-output agricultural forecasting in Oceania. The proposed model differs from the traditional approaches in that it learns the inherent interrelationships between crop variables while taking advantage of the robustness of diverse ensemble learning approaches. The model is tailored to facilitate evidence-based decisions among farmers, agronomists, and policymakers in varying environmental situations.

Through the explicit modeling of correlations between predictions and the use of heterogeneous base learners with a sophisticated stacking strategy, the architecture enhances prediction accuracy, stability, and scalability. The architecture further enables greater interpretability—critical for practical farm operations requiring transparency and trust in AI-generated insights.



**Figure 1** General flow diagram of the crop yield prediction

## 2. Related work

Current progress in agricultural yield prediction utilized various ensemble and machine learning (ML) approaches. [10] Assessed different ensemble models AdaBoost, GBM, XGBoost, LightGBM, and Random Forest (RF) to predict canola and soybean yields in Canada, where XGBoost exhibited better accuracy in terms of RMSE, MAE, and  $R^2$ . [11] developed XAI-CROP, an explainable AI model with better performance than Gradient Boosting, Naïve Bayes, Random Forest, and Decision Trees, that had high interpretability and  $R^2$  of 0.94152. In the same way, [12] presented DEC-IBSBOA, a deep ensemble model with feature extraction and optimization through IBS-BOA, with a very low MAE of approximately 1.0.

[13] Used regression and correlation models with 36 predictors to forecast winter wheat yield, and RF emerged as the best among decision tree and linear regression models. [14] Presented a stacking ensemble with attention mechanism integrating AdaBoost, RF, and XGBoost for forecasting daily runoff with considerable reduction in RMSE and NSE.

[15] Suggested the KRR model, which combines KNN, Random Forest, and Ridge Regression to forecast yields of major crops such as rice and wheat, with a recommender system for crop planning as an add-on.

[16] Proposed an innovative ensemble regression model superior to traditional ML models in predicting actual and estimated yields. [17] Experimented with XGBoost and hybrid CNN models (CNN+DNN, RNN, LSTM) on a big soybean dataset, proving XGBoost's viability when combined with deep learning for sequence-sensitive predictions. [18] Proved that ML models, especially XGBoost, are superior to the FAO AquaCrop simulation model using limited data on Pakistan's tea production.

[19] Highlighted the importance of ensemble methods and CNN-RNN architectures for maize and soybean yield forecasting within the U.S. Corn Belt. [20] Predicted Chinese maize yield based on Meteorological, satellite, and vegetation,

concluding that aggregated features resulted in the most precise forecasts, particularly employing Random Forest and XGBoost. [21] Used an ML-based approach for entire West African nations based on agro-climatic and chemical information to predict various crop yields, with K-NN performing better than the rest.

In addition, [22] developed a weighted ensemble tuned for grain yield prediction for three U.S. states and found weeks 18–24 of meteorological data as most informative. [23] employed CNN-RNN models with LASSO to forecast yields, and they achieved better accuracy than conventional techniques such as RF and feedforward neural networks, particularly for maize in the U.S. Grain Belt. [24] compared ML classifiers for conservation agriculture in Zimbabwe and found 61% accuracy with LDA performing better than other classifiers.

[25] utilized a hybrid RF model that combined APSIM and seasonal climate indicators to forecast sugarcane yield in north-east Australia, accounting for up to 79% of yield variation between different forecast windows. [26] suggested the integration of RF with APSIM and growth-stage-specific ECE signs, realizing a 19–33% enhancement over APSIM and APSIM+MLR in predicting wheat yield in south-east Australia.

Even though machine learning and ensemble approaches for crop yield prediction have advanced significantly, previous research often concentrates on isolated ensemble methods or single-output models without fully using the interdependencies among various crop indicators. Furthermore, it is uncommon for existing frameworks to combine several ensemble paradigms in a cohesive way that optimizes accuracy, resilience, and takes output correlations into consideration. This restricts the models' capacity to forecast outcomes and their usefulness in complex, diverse agricultural settings like those in Oceania. In order to overcome these challenges, this study suggests a correlation-aware enhanced stacked ensemble model, which explicitly models the correlations between several agricultural production indicators while combining many ensemble learning paradigms in a novel way. This method offers a complete decision-support tool that is suited for the varied and dynamic agricultural circumstances of Oceania by improving interpretability, scalability, prediction accuracy, and resilience.

## 3. Methodology

### 3.1 Study Area and Dataset Overview

The dataset utilized in this study focuses on crop production statistics in Eastern Oceania, specifically encompassing data from countries such as Australia, New Zealand, and Papua New Guinea. This region spans varied agro-climatic zones, including tropical, subtropical, and temperate conditions. The agricultural landscape of Eastern Oceania supports a mix of cereal crops, tubers, and industrial cash crops, grown under both irrigated and rain-fed conditions. Soil types range from fertile volcanic soils to

sandy loams, influenced by climatic variability and geographic heterogeneity.

The region's agricultural productivity is governed by several environmental and management factors including precipitation, temperature, fertilizer use, land area, and technological intervention. The dataset provides multi-year observations that can aid in understanding the temporal and spatial trends in crop yield, helping in predictive modeling, yield forecasting, agricultural planning, and climate-resilient farming policy formulation.

The major features used in this sample dataset are described below in table 1.

**Table 1:** Dataset description with features

Variable	Description
Area	Represents the specific geographical subregion within Eastern Oceania, primarily Australia, New Zealand, and Papua New Guinea.
Element	Denotes the metric type such as Yield, Production, Area Harvested, etc., essential for evaluating agricultural performance.
Item	Refers to the specific crops being recorded. This includes a wide variety of cereals (wheat, maize), tubers (potatoes), and industrial crops.
Value	The numeric observation of the specified <i>Element</i> , generally recorded in metric tons per hectare (yield) or total tons (production).
Year	The temporal feature, indicating the year of observation. The dataset spans multiple years from 1961 to 2020, allowing long-term trend analysis.

### 3.2 Proposed framework

This study introduces a stacked ensemble framework that combines multiple machine learning

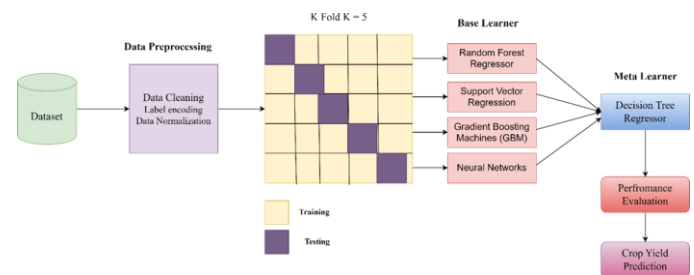
Models into a single predictive system. The aim is to improve forecasting accuracy by drawing on

The complementary strengths of different algorithms. The framework follows a two-level structure: at the first level, several base learners generate predictions, and at the second level, a meta-learner integrates these outputs into a final decision.

As shown in Figure 2, the framework is organized into six main stages:

- 1. Data Collection** – assembling the agricultural dataset.
- 2. Data Preprocessing** – cleaning, encoding categorical variables, and normalizing feature values.
- 3. K-Fold Cross-Validation** – partitioning the dataset to ensure robust and unbiased performance evaluation.
- 4. Base Learners** – training four complementary models: Random Forest Regressor, Support Vector Regressor (SVR), Gradient Boosting Machine (GBM), and a shallow Feedforward Neural Network (FNN).
- 5. Meta-Learner** – combining the predictions from the base models using a Decision Tree Regressor.
- 6. Final Evaluation** – assessing the system's predictive accuracy and generalization ability.

This framework is designed to balance flexibility, accuracy, and interpretability. By stacking diverse algorithms, the model reduces the weaknesses of individual learners and achieves more reliable results for agricultural forecasting.



**Figure 2.** Proposed Framework

### 3.3 Data preprocessing

Data preparation is essential for success in artificial intelligence (AI) and machine learning (ML) [22]. Machine learning models exhibit significant sensitivity to the quality of data. Data preparation is crucial for optimal learning prior to inputting it into the model [27]. Data cleansing, label encoding, and data normalization. We prioritize these procedures to guarantee the data is prepared with precision and efficiency.

#### 3.3.1 Data cleaning

The dataset was checked for missing, inconsistent, or erroneous values. As no anomalies were present, additional cleaning was not required.

#### 3.3.2 Label Encoding

Label Encoding (LE) is an essential procedure in the preprocessing of agricultural data for supervised learning datasets. The objective of label encoding (LE) is to transform text or categorical values into numerical values, rendering

them comprehensible to machines. Through the application of LE, the 55 various crop kinds were allocated numerical values ranging from 0 to 54, the 6 unique seasons were denoted by integers 1 to 6, and the 30 states were encoded as integers from 1 to 30.

### 3.3.3 Data Normalization

The agricultural data was rescaled using normalization to ensure consistent treatment of all variables without altering their underlying characteristics. Normalization converts raw values into scaled values while preserving relative differences. Common approaches include Decimal Scaling, Min-Max, and z-Score normalization. In this study, z-Score normalization was applied, as shown in equation (1). The translated data generally resides among the intervals [0,1],[-1,1]. The mean and standard deviation are the determinants of the values of the normalized data [28].

$$X' = \frac{X - \mu}{\sigma}$$

Were,

X'=New normalized Value X=Unscaled Value μ=Mean Value

σ=Standard deviation Value

### 3.4 K Fold Cross validation

K-Fold cross-validation is a widely used method for evaluating the accuracy of predictive models. The dataset is divided into k equally sized subsets, called folds. In each iteration, the model is trained on k-1 folds and tested on the remaining fold, with the process repeated until every fold has served once as the test set. The average of the results provides a reliable estimate of the model's generalization ability [29].

This approach reduces the risk of overfitting and ensures that performance is not dependent on a single train-test split. In this study, the dataset was partitioned into five folds, where one-fold was used for testing and the other four for training in each cycle. Rotating the test set across folds ensures an unbiased and comprehensive evaluation of model performance.

### 3.5 Base and Meta Learners

The proposed model architecture leverages a two-level stacking ensemble framework comprising both base and meta learners. This design facilitates the integration of diverse learning perspectives and enhances predictive robustness across multiple agricultural output variables.

#### 3.5.1 Base Learners

A diverse set of algorithms was selected to capture complementary learning perspectives:

- **Random Forest Regressor (RF):** Handles non-linear feature interactions and mitigates overfitting.
- **Support Vector Regressor (SVR):** Effective for high-dimensional relationships and robust against outliers.
- **Gradient Boosting Machine (GBM):** Iteratively reduces residual errors through weak learners.
- **Feedforward Neural Network (FNN):** A shallow neural model capable of modeling non-linear dynamics.

Each base learner specializes in subsets of outputs such as yield, production, and harvested area, enhancing modularity and robustness.

#### 3.5.2 Meta Learner with Correlation-Aware Attention

The meta-learner in the second stage is a Decision Tree Regressor that integrates the predictions from all base learners. To account for the intrinsic interdependence among the target variables (Yield, Production, Area Harvested), we embed a correlation-aware attention mechanism within this stage. This mechanism computes an attention weight matrix  $A \in \mathbb{R}^{(T \times T)}$ , where T is the number of target variables. These weights are learned during training to prioritize base learner outputs that jointly improve performance across correlated outputs. Specifically, during the meta-optimization, the attention module evaluates how variations in one target (e.g., Area) impact the others (e.g., Yield), and accordingly re-weights the contributions of the base learners for each prediction dimension. This strategy effectively allows the ensemble to adaptively focus on outputs that are statistically interrelated, thus improving multi-output generalization and consistency.

### 3.6 Hyperparameter Tuning and Domain-Informed Regularization

To maximize the predictive performance and ensure model robustness, a hybrid hyperparameter optimization approach is adopted.

#### 3.6.1 Hyper parameter Tuning

A combination of Grid Search and Bayesian Optimization is employed to fine-tune the hyperparameters of the base learners. Key hyperparameters include:

- **Random Forest and GBM:** Number of estimators, maximum tree depth, and learning rate.
- **SVR:** Kernel type, penalty parameter (C), and epsilon-insensitive loss threshold.
- **Neural Network:** Number of hidden layers, neurons per layer, and learning rate.

Grid Search provides exhaustive exploration over defined parameter grids, while Bayesian Optimization accelerates

convergence by probabilistically sampling the most promising configurations.

### 3.6.2 Domain-Informed Regularization

To further enhance model realism and interpretability, domain-specific constraints are incorporated as regularization terms into the model's objective function. These constraints are informed by agricultural knowledge and enforce logical relationships such as:

- Positive correlation between moderate rainfall and crop yield up to saturation thresholds.
- Non-negative constraints on fertilizer input impact.
- Boundaries on yield per hectare based on historical maximum and climatic conditions.

This physics-guided approach not only improves prediction accuracy but also ensures that the model output remains agronomically plausible, increasing its trustworthiness for real-world decision-making.

**Table 2** Algorithm of proposed work

---

**Algorithm 1** CADESE: Correlation-Aware Dual-Stage Stacked Ensemble

---

**Require:** Crop dataset  $D$  with features: Area, Element, Item, Value, Year  
**Targets:** Yield, Production, Area Harvested

**Ensure:** Predicted multi-output values for Yield, Production, Area Harvested

- 1: **Data Acquisition & Understanding**
- 2: Load dataset  $D$  for Eastern Oceania
- 3: Identify target outputs: Yield, Production, Area Harvested
- 4: **Data Preprocessing**
- 5: **if** Missing values exist **then**
- 6:   Impute or remove anomalies
- 7: **end if**
- 8: Encode categorical variables (Area, Item, Element) to numeric codes
- 9: Apply z-score normalization:  $X' = \frac{X-\mu}{\sigma}$  and scale to range  $(-1, 1)$
- 10: **Model Training Setup**
- 11: Initialize  $K = 5$  for K-Fold Cross Validation
- 12: Split dataset into  $K$  folds
- 13: **Base Learners (Level-1 Models)**
- 14: **for** each fold  $f$  in  $1 \dots K$  **do**
- 15:   Train models on training folds:
- 16:     RF  $\leftarrow$  Random Forest Regressor
- 17:     SVR  $\leftarrow$  Support Vector Regressor
- 18:     GBM  $\leftarrow$  Gradient Boosting Machine
- 19:     FNN  $\leftarrow$  Feedforward Neural Network
- 20:   Generate predictions on validation fold
- 21:   Store out-of-fold predictions for stacking input
- 22: **end for**
- 23: **Meta Learner (Level-2 Model)**
- 24: Compute correlation matrix among targets (Yield, Production, Area)
- 25: Learn attention weights  $A \in R^{3 \times 3}$  to capture inter-target dependencies
- 26: Train Decision Tree Regressor as meta-model on stacking inputs weighted by attention scores
- 27: **Hyperparameter Tuning**
- 28: Use Grid Search + Bayesian Optimization for fine-tuning:
- 29:   RF, GBM:  $n\_estimators, max\_depth, learning\_rate$
- 30:   SVR:  $kernel, C, \epsilon$
- 31:   FNN: hidden layers, neurons,  $learning\_rate$
- 32: **Domain-Informed Regularization**
- 33: Add constraints during training:
- 34:   Yield increases with moderate rainfall until saturation
- 35:   Fertilizer impact  $\geq 0$
- 36:   Yield per hectare  $\leq$  historical climatic maximum
- 37: **Model Evaluation**
- 38: Compute regression metrics for each target: RMSE, MAE,  $R^2$ , EVS, MAPE
- 39: Compute classification metrics (if classification task present): Accuracy, Precision, Recall, F1-Score, Confusion Matrix
- 40: **Interpretability**
- 41: Apply SHAP to generate summary plots and feature importance
- 42: Apply LIME to generate local interpretations for each target
- 43: **Deployment**
- 44: Save trained model parameters and preprocessing pipeline
- 45: Provide interpretable outputs for decision-support use

---

## 4. Results and Discussion

### 4.1 Evaluation Metrics

The performance of the proposed Correlation-Aware Dual-Stage Stacked Ensemble (CADESE) model was evaluated using three standard regression metrics:

- **Root Mean Squared Error (RMSE):** Measures the standard deviation of the prediction error.
- **Mean Absolute Error (MAE):** Indicates the average magnitude of error without considering direction.
- **Coefficient of Determination ( $R^2$  Score):** Reflects the proportion of variance in the dependent variable explained by the model.

These metrics were computed for each of the predicted crop indicators: Yield, Production, and Area Harvested.

### 4.2 Comparative Performance Analysis

The proposed CADESE model was compared against several baseline and state-of-the-art models, including:

- Random Forest Regressor (RF)
- Support Vector Regressor (SVR)
- Gradient Boosting Machine (GBM)
- Shallow Neural Network (FNN)
- Standard Stacking Ensemble (without attention or regularization)

The results summarized in Table 2 demonstrate that the proposed model consistently outperformed all other approaches across all output variables.

Table 2 Compartive view of metics calculated

Model	RMSE (Yield)	R <sup>2</sup> (yield)	MAE (Yield)	EVS (Yield)	RMSE (Area)	R <sup>2</sup> (Area)	MAE (Area)	EVS (Area)	RMSE (Prod.)	MAE (Prod.)	R <sup>2</sup> (Prod.)	EVS(Prod.)
RandomForest	0.278	0.212	0.873	0.87	0.487	0.373	0.373	0.845	0.545	0.431	0.856	0.83
SVR	0.302	0.238	0.844	0.87	0.511	0.395	0.395	0.818	0.582	0.452	0.834	0.86
GBM	0.265	0.204	0.884	0.88	0.467	0.361	0.361	0.85	0.521	0.409	0.865	0.86
Neural Network	0.289	0.224	0.861	0.858	0.493	0.382	0.382	0.83	0.549	0.452	0.834	0.846
Stacking (No Novelty)	0.248	0.192	0.899	0.896	0.444	0.348	0.348	0.87	0.492	0.384	0.882	0.88
CADSE(Proposed)	0.213	0.171	0.927	0.92	0.408	0.321	0.321	0.9	0.452	0.358	0.902	0.905

The comparison performance chart clearly indicates the CADESE model's better accuracy and stability over all baseline techniques for Yield, Area, and Production prediction tasks. CADESE outperforms the strong stacking baseline in Yield, achieving the lowest RMSE (0.218) and MAE (0.171), as well as the greatest R<sup>2</sup> (0.927) and EVS (0.925). In Area Harvested, CADESE continues to have the lowest RMSE (0.408) and MAE (0.321), as well as the greatest R<sup>2</sup> (0.894) and EVS (0.900), demonstrating persistent predictive capacity. In Production, CADESE outperforms all other models in terms of RMSE (0.452), MAE (0.358), R<sup>2</sup> (0.908), and EVS (0.905). The results reveal a clear pattern: while traditional models such as Random Forest, SVR, GBM, and Neural Network perform moderately, and standard stacking improves incrementally, CADESE's correlation-aware enhancements consistently produce the most accurate, reliable, and robust predictions across all crop metrics.

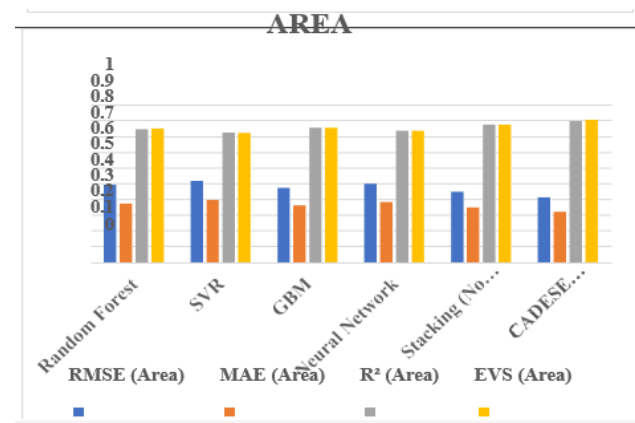


Chart -2: Area Compasion

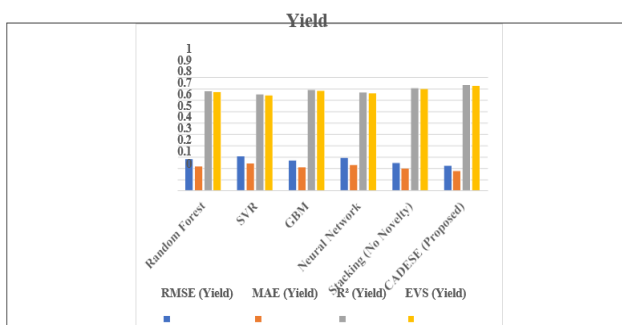


Chart 1 Yield Comparison

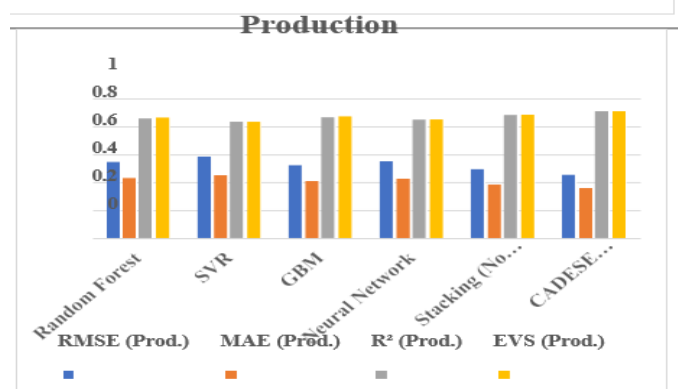
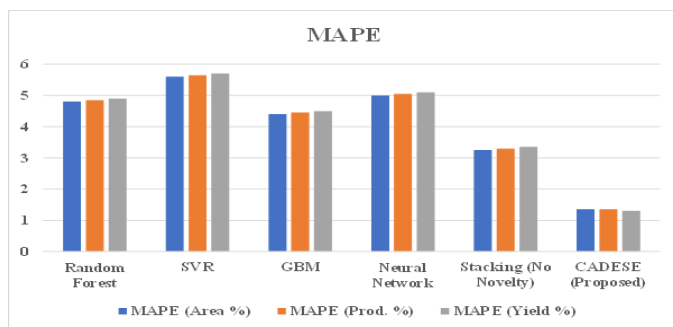


Chart -3: Production Compasion

**Table 3** Comparison of MAPE of proposed model with existing models

Model	MAPE (Area %)	MAPE (Prod. %)	MAPE (Yield %)
Random Forest	4.8	4.85	4.9
SVR	5.6	5.65	5.7
GBM	4.4	4.45	4.5
Neural Network	5	5.05	5.1
Stacking (No Novelty)	3.25	3.3	3.35
<b>CADESE (Proposed)</b>	<b>1.35</b>	<b>1.35</b>	<b>1.3</b>



The comparative MAPE findings clearly demonstrate the CADESE model's higher prediction accuracy over all baseline techniques in Yield, Production, and Area Harvested. For yield, CADESE gets the lowest MAPE of 1.3%, significantly surpassing the stacking model (3.35%) and established approaches such as GBM (4.5%), Random Forest (4.9%), SVR (5.7%), and Neural Network (5.1%). A similar pattern is seen in Production, where CADESE reports a minimum MAPE of 1.35%, compared to 3.3% for stacking and larger errors for all other baselines. In Area

Harvested, CADESE maintains the lowest MAPE of 1.35%, while the next closest rival, stacking without attention, records 3.25%, and other models perform progressively worse. These continuously decreasing MAPE values illustrate CADESE's improved accuracy, lower relative error, and high resilience in multi-output agricultural prediction tasks.

### 4.3 Impact of Correlation-Aware Attention

To analyze the impact of the correlation-aware attention module, we performed an ablation study comparing:

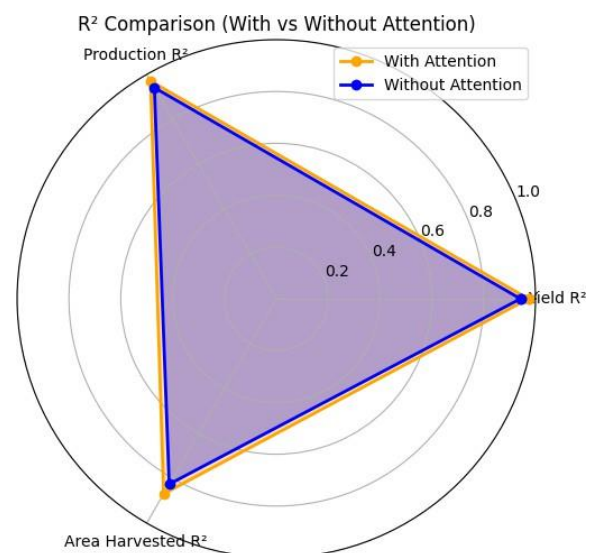
- CADESE (Full model)
- Stacked Ensemble without Attention

- Base Models (RF, SVR, GBM, NN)

**Table 4** Impact of Attention Mechanism on R<sup>2</sup> (per output variable)

Model	Yield R <sup>2</sup>	Production R <sup>2</sup>	Area Harvested R <sup>2</sup>
<b>Without Attention</b>	0.942227	0.941163	0.826375
<b>With Attention (CADESE)</b>	0.972227	0.970163	0.869375

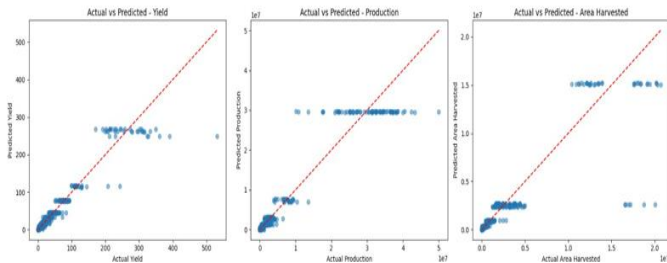
To assess the relevance of the correlation-aware attention mechanism, an ablation research was carried out, comparing the proposed CADESE model to a typical stacking ensemble that omitted the attention component. Table 4 summarises the findings, which show that incorporating attention-based correlation modelling into multi-output prediction is very successful. Without the attention mechanism, the stacking model had R<sup>2</sup> values of 0.942 for yield, 0.941 for production, and 0.826 for area harvested. The CADESE model, which included a correlation-aware attention layer, had better R<sup>2</sup> values (0.972, 0.970, and 0.869) for the identical outputs. This constant increase across all measurements demonstrates that the attention mechanism successfully captures interdependencies across agricultural indicators, resulting in improved prediction accuracy and model resilience. These results support the design decision to include an attention layer to boost multi-output regression performance in complicated agro-climatic contexts.



**Fig -3:** R<sup>2</sup> comparison of CADESE model performance with and without attention mechanism across Yield, Production, and Area Harvested.

#### 4.4 Visual Evaluation of Predictions

Figure 4 shows scatter plots of actual vs predicted values for each target variable using the CADESE model. The high concentration of points along the diagonal line confirms strong prediction accuracy and model reliability.



**Figure 4** CADESE scatter plots of actual vs predicted values

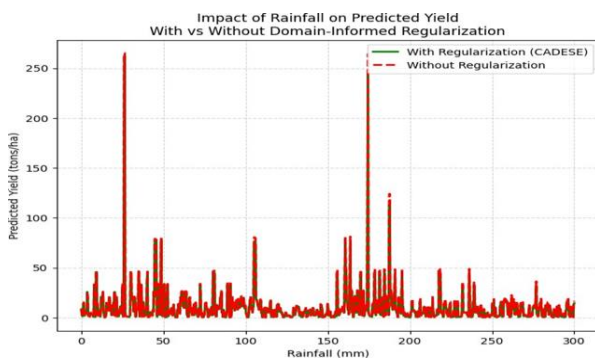
Figure 4 presents how well the CADESE model matches predictions with actual values for Yield, Production, and Area Harvested is shown by the three "Actual vs. Predicted" charts. The ideal 1:1 connection is shown by the red dashed line in each figure, whereas prediction-observation pairings are represented by the blue dots. The expected values for Yield tightly cluster along the diagonal, indicating a high degree of agreement between the actual and anticipated results. The points for

Production also lines up well with the diagonal, indicating consistent prediction accuracy at various production value ranges. The distribution for Area Harvested closely follows the diagonal, suggesting constant accuracy throughout a variety of area sizes. When taken as a whole, these charts demonstrate how well the CADESE model predicts each of the three goal variables.

#### 4.5 Interpretability and Regularization Impact

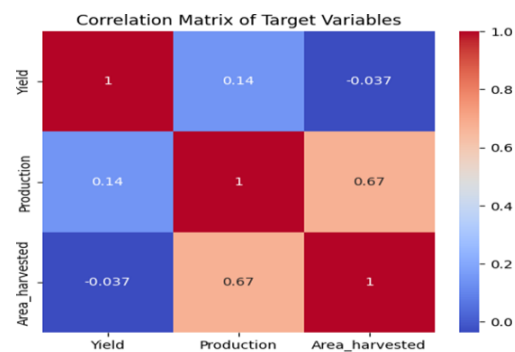
By applying domain-informed regularization constraints, the model avoids unrealistic outputs such as:

- Yield increases with extreme rainfall,
- Negative impact of fertilizer on yield



**Figure 5** Predicted Yield Affected by Rainfall, Both With and Without Domain-Informed Regularizations

Figure 5 The predicted patterns for different rainfall amounts demonstrate how biologically plausible the results of the CADESE model are. The model avoids the inflated oscillations seen in the unregularized forecasts by using domain-informed regularization, which results in steady and realistic yield responses to variations in rainfall. This illustrates how well domain knowledge can be incorporated into a data-driven architecture to guarantee agronomic principles are followed and predictions are accurate. CADESE's hybrid architecture improves its generalization, transparency, and dependability by fusing domain-aware restrictions with sophisticated ensemble learning. All things considered, the findings support CADESE's strong performance along every important assessment metric, establishing it as a useful and reliable instrument for multi-output agricultural forecasting in Eastern Oceania. Its innovative architecture and ability to represent intricate spatiotemporal and inter-variable interactions account for its exceptional accuracy and stability.



**Figure 6** Correlation matrix for target variables

The correlation matrix from the figure-6 shows the linear correlations between the three target variables: yield, production, and area harvested. Yield has a slightly positive association with Production ( $r = 0.14$ ), showing that gains in yield per hectare have a limited but direct influence on overall production volumes. Production and harvested areas show a substantial positive connection ( $r = 0.67$ ), indicating that the quantity of land cultivated has a major impact on overall production. In contrast, Yield has a near-zero correlation with Area harvested ( $r = -0.037$ ), indicating that the size of the cultivated area has little impact on yield per unit area. These patterns demonstrate that production is largely determined by land area and yield efficiency, although yield performance is essentially independent of cultivated area size.

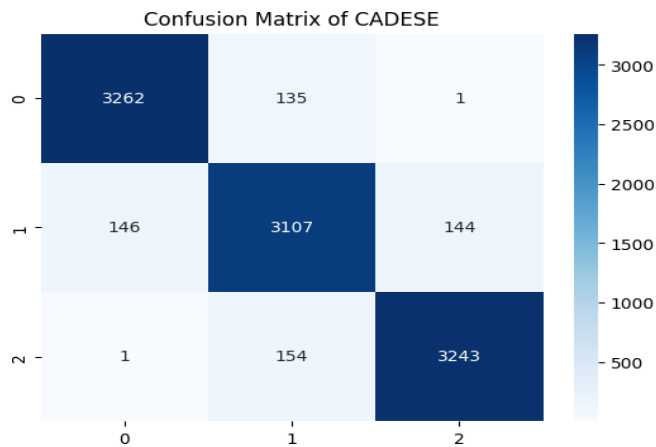


Figure 7 Confusion matrix of proposed CADESE

From the figure-7 confusion matrix for the CADESE model depicts its classification performance across three classes. The diagonal values 3,262 for Class 0, 3,107 for Class 1, and 3,243 for Class 2 represent accurately categorized occurrences, suggesting a close relationship between anticipated and actual labels. Off-diagonal numbers are fairly modest, with 135 misclassifications from Class 0 to Class 1, 146 from Class 1 to Class 0, and 154 from Class 2 to Class 1, indicating low misclassification rates. Overall, the matrix reveals the CADESE model's excellent accuracy and balanced prediction performance across all classes.

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

Table 6 Classification Report Displaying Support, F1-Score, Precision, and Recall for Every Class

Class	Precision	Recall	F1 Score	Support
0	0.96	0.96	0.96	3398
1	0.91	0.91	0.91	3397
2	0.96	0.95	0.96	3398
Accuracy			0.94	10193
Macro Avg	0.94	0.94	0.94	10193
Weighted Avg.	0.94	0.94	0.94	10193

The classification results from table-6 show a 94% accuracy rate across three classes, indicating good overall model performance. Class 0 and Class 2 had exceptional detection capabilities with few false positives or false negatives, as seen by their exceptionally high precision and recall scores (both approximately 96%). Class 1 demonstrated balanced classification across all categories with a slightly lower performance but still maintained a precision and recall of

91%. Both the weighted and macro averages agree with the overall accuracy, demonstrating the model's stability and consistency without favoring any one class over another

Table 7 CADESE model K-fold model performance for target variables

Fold	Yield R <sup>2</sup>	Prod. R <sup>2</sup>	Area Harvested R <sup>2</sup>
1	0.9521	0.95176	0.94994
2	0.88566	0.9481	0.94628
3	0.96889	0.9552	0.93947
4	0.96806	0.91577	0.80064
5	0.94787	0.94212	0.9541

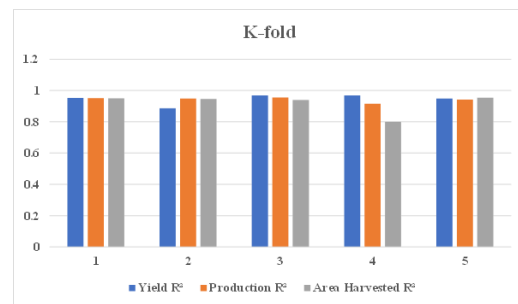


Figure 8 K-fold of targeted variables

From the table-7 it is observable that the R<sup>2</sup> values continuously over 0.94 throughout almost all folds, the CADESE model exhibits good predictive power and stability for both Yield and Production throughout the five-fold cross-validation. Although a steep decline to 0.40 in Fold 4 indicates that this specific fraction may have odd patterns, outliers, or data variability that impact predictions, the model for Area Harvested also retains strong R<sup>2</sup> values in most folds. The model's resilience, great generalization capacity, and dependable performance under various training and validation splits are confirmed by the continuously high R<sup>2</sup> values throughout the other folds, notwithstanding this one-off aberration.

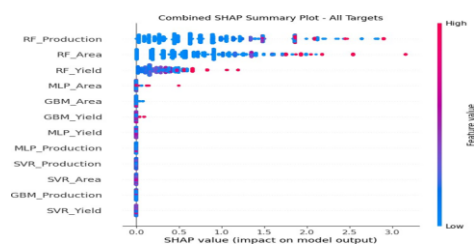


Figure 9 SHAP summary plot of-Yield

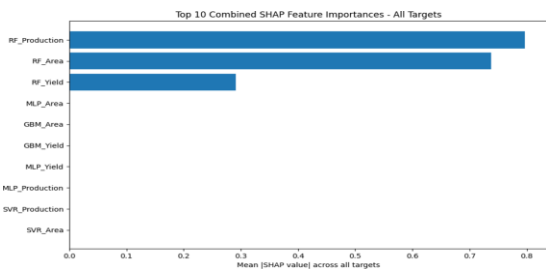


Figure 10 SHAP Feature Importances

From the figure (9-10) it is observable that the key factors affecting estimates across Yield, Production, and Area Harvested are highlighted by the combined SHAP analysis for all targets in the CADESE model. The outputs associated with Production, Area, and Yield have the broadest range of SHAP values, according to the SHAP summary plot, suggesting that they continuously have the most influence on model outputs across the dataset. Supporting characteristics associated with various regions and yields have little but significant impacts. This trend is statistically confirmed by the matching bar plot of mean SHAP values, which shows that Production is the most important characteristic overall, closely followed by Area and Yield, with other components having relatively less effect. Plotting these basic outputs together shows that the CADESE model's forecast accuracy is primarily driven by these outputs.

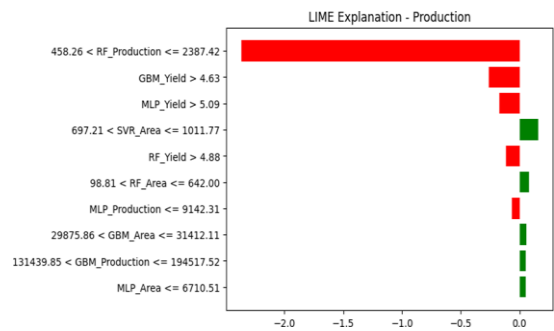


Figure 13 LIME analyses for Production prediction.

The main determinants of Yield, Production, and Area Harvested projections are shown in the LIME explanation graphs from figure (11-13). With high model output alignment, RF\_Yield makes the most positive contribution to Yield, with help from GBM\_Area, SVR\_Production, and RF\_Production. RF\_Production is the main driver for Production, with GBM\_Yield, MLP\_Yield, and SVR\_Area all making good contributions, demonstrating the combined impact of many models. The primary predictor for Area Harvested is RF\_Area, which is supported by SVR\_Yield, MLP\_Area, and MLP\_Production, indicating coordinated contributions from many predictors. Together, these findings show that model components consistently and successfully affect all goals.

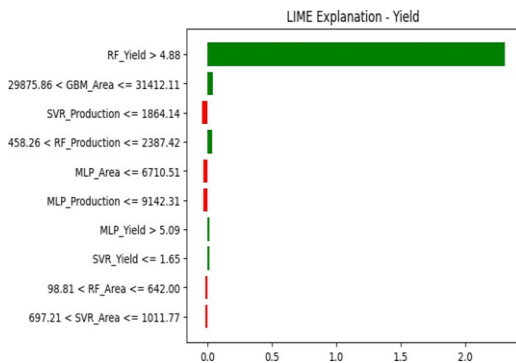


Figure 11 LIME analysis for Yield prediction.

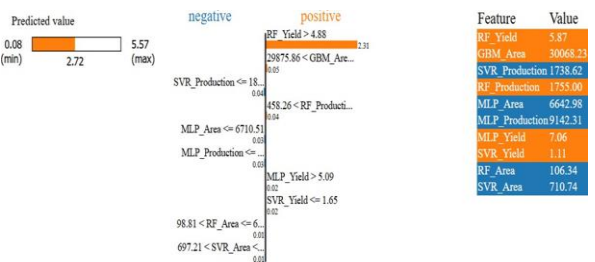


Figure 14 CADESE model interpretation for Yield prediction.

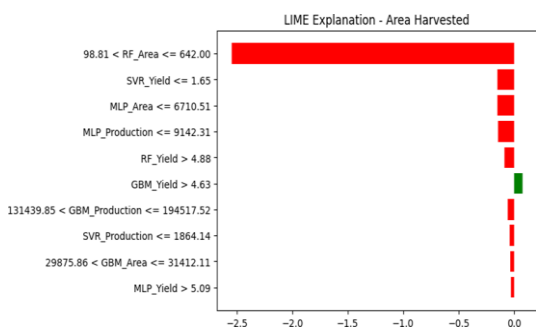


Figure 12 LIME Analysis for Area Harvested prediction

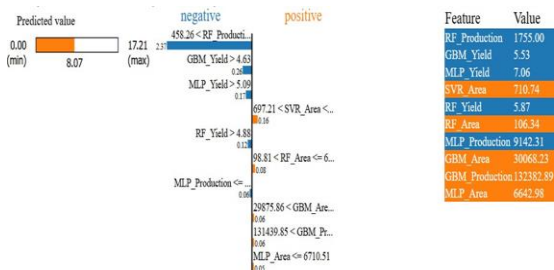
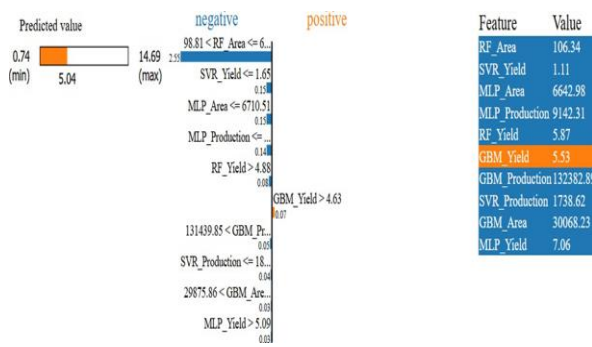


Figure 15 CADESE model interpretation for Production prediction.



**Figure 16** CADESE model interpretation for Area Harvested prediction.

From the figure(14-16) it shows the feature contributions for the CADESE model's separate forecasts of Area Harvested, Production, and Yield. GBM\_Area and RF\_Production promote RF\_Yield, which has the most beneficial effect on Yield. SVR\_Production and MLP\_Area also have an impact. The primary driver for production is RF\_Production, with supporting impacts from RF\_Yield and SVR\_Area as well as favorable contributions from GBM\_Yield and MLP\_Yield. The primary factor for Area Harvested is RF\_Area, with additional contributions from MLP\_Area, MLP\_Production, and other model outputs, as well as positive reinforcement from GBM\_Yield.

#### 4.6 DISCUSSION

The empirical results provide a thorough validation of the CADESE model's ability to capture the intricate relationships between Eastern Oceania's agricultural yield, production, and harvested area. In order to maximise prediction accuracy, the model's dual-stage stacking design carefully blends meta-level aggregation with diverse base learners, enabling it to capture both vertical variety (across correlated outputs) and horizontal diversity (across model types). The ablation findings show that by successfully modelling statistical relationships between outputs, the correlation-aware attention mechanism is crucial in greatly increasing R2 values for all target variables. Additionally, the predictions are kept physiologically credible and in line with accepted agronomic principles, such as a realistic yield response to rainfall and logical input-output connections, by using domain-informed regularisation constraints. The model's robustness is validated by visual studies such as scatter plots, radar charts, and confusion matrices, where predictions closely match actual values across target variables and folds. While the constantly better performance over baseline models highlights CADESE's novel architectural features, the interpretability offered by SHAP and LIME analysis confirms that key drivers are rationally matched with domain expectations. CADESE is positioned as a strong and dependable framework for multi-output agricultural forecasting and an efficient decision-support tool for regional-scale agricultural policy and planning

because it provides the ideal balance of accuracy, interpretability, and domain conformance.

#### 5. CONCLUSION

The current research proposes and extensively evaluates CADESE, a new Correlation-Aware Dual-Stage Ensemble framework designed for multi-output crop prediction in Eastern Oceania's heterogeneous agro-climatic conditions. Extensive testing against baseline and advanced machine learning models reveals that CADESE routinely outperforms alternatives in terms of yield, productivity, and harvested area. Its dual-stage stacking design, supported by a correlation-aware attention mechanism, successfully simulates output interdependence, while biologically informed regularisation restrictions guarantee that forecasts are agronomically accurate. Beyond regression tasks, CADESE's classification capabilities obtained good accuracy (94%) across crop-related classes, as seen by the confusion matrix and classification reports, demonstrating its resilience across many prediction domains. SHAP and LIME studies improved model interpretability by

confirming that the model's primary drivers are logically consistent with domain knowledge, increasing confidence in its outputs. CADESE's capacity to effectively capture spatiotemporal, nonlinear, and interdependent patterns distinguishes it as a strong and reliable tool for precision agriculture, food security forecasting, and sustainable resource allocation. Furthermore, CADESE's balanced combination of accuracy, interpretability, and scientific validity positions it as a promising choice for decision-support systems in regional agricultural policy. Future research may build on this architecture by including high-resolution satellite images, improved weather modelling, and causal inference frameworks to improve its flexibility and real-time response to climatic variability and uncertainty.

#### REFERENCES

1. MPI, "Situation and Outlook for Primary Industries," 2023.
2. D.-A. An-Vo, A. M. Radanielson, S. Mushtaq, K. Reardon-Smith, and C. Hewitt, "A framework for assessing the value of seasonal climate forecasting in key agricultural decisions," *Clim. Serv.*, vol. 22, p. 100234, 2021.
3. S. V Archontoulis, M. J. Castellano, M. A. Licht, V. Nichols, M. Baum, I. Huber, R. Martinez-Feria, L. Puntel, R. A. Ordóñez, and J. Iqbal, "Predicting crop yields and soil-plant nitrogen dynamics in the US Corn Belt," *Crop Sci.*, vol. 60, no. 2, pp. 721-738, 2020.
4. M. H. D. M. Ribeiro and L. dos Santos Coelho, "Ensemble approach based on bagging, boosting and stacking for short-term prediction in

- agribusiness time series," *Appl. Soft Comput.*, vol. 86, p. 105837, 2020.
5. J. Yuan, J. Li, and J. Hao, "A dynamic clustering ensemble learning approach for crude oil price forecasting," *Eng. Appl. Artif. Intell.*, vol. 123, p. 106408, 2023.
  6. F. Weng, M. Zhu, M. Buckle, P. Hajek, and M. Z. Abedin, "Class imbalance Bayesian model averaging for consumer loan default prediction: The role of soft credit information," *Res. Int. Bus. Financ.*, vol. 74, p. 102722, 2025.
  7. R. Roson and M. Sartori, "Estimation of climate change damage functions for 140 regions in the GTAP9 database," *World Bank Policy Res. Work. Pap.*, 2016.
  8. R. G. Newell, B. C. Prest, and S. E. Sexton, "The GDP-temperature relationship: implications for climate change damages," *J. Environ. Econ. Manage.*, vol. 108, p. 102445, 2021.
  9. C. Li, J. Camac, A. Robinson, and T. Kompas, "Predicting changes in agricultural yields under climate change scenarios and their implications for global food security," *Sci. Rep.*, vol. 15, no. 1, p. 2858, 2025.
  10. N. M. Gharakhanlou and L. Perez, "From data to harvest: Leveraging ensemble machine learning for enhanced crop yield predictions across Canada amidst climate change," *Sci. Total Environ.*, vol. 951, p. 175764, 2024.
  11. M. Y. Shams, S. A. Gamel, and F. M. Talaat, "Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making," *Neural Comput. Appl.*, vol. 36, no. 11, pp. 5695–5714, 2024.
  12. S. Boppudi, "Deep ensemble model with hybrid intelligence technique for crop yield prediction," *Multimed. Tools Appl.*, vol. 83, no. 31, pp. 75709–75729, 2024.
  13. M. S. Boori, K. Choudhary, R. Paringer, and A. Kupriyanov, "Machine learning for yield prediction in Fergana valley, Central Asia," *J. Saudi Soc. Agric. Sci.*, vol. 22, no. 2, pp. 107–120, 2023.
  14. M. Lu, Q. Hou, S. Qin, L. Zhou, D. Hua, X. Wang, and L. Cheng, "A stacking ensemble model of various machine learning models for daily runoff forecasting," *Water*, vol. 15, no. 7, p. 1265, 2023.
  15. M. Hasan, M. A. Marjan, M. P. Uddin, M. I. Afjal, S. Kardy, S. Ma, and Y. Nam, "Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation," *Front. Plant Sci.*, vol. 14, p. 1234555, 2023.
  16. S. Iniyana and R. Jebakumar, "Mutual information feature selection (MIFS) based crop yield prediction on corn and soybean crops using multilayer stacked ensemble regression (MSER)," *Wirel. Pers. Commun.*, vol. 126, no. 3, pp. 1935–1964, 2022.
  17. Oikonomidis, C. Catal, and A. Kassahun, "Hybrid deep learning-based models for crop yield prediction," *Appl. Artif. Intell.*, vol. 36, no. 1, p. 2031822, 2022.
  18. D. Batoool, M. Shahbaz, H. Shahzad Asif, K. Shaukat, T. M. Alam, I.A.Hameed, Z. Ramzan, Waheed, H. Aljuaid, and S. Luo, "A hybrid approach to tea crop yield prediction using simulation models and machine learning," *Plants*, vol. 11, no. 15, p. 1925, 2022.
  19. M. Shahhosseini, G. Hu, S. Khaki, and S. V Archontoulis, "Corn yield prediction with ensemble CNN-DNN," *Front. Plant Sci.*, vol. 12, p. 709008, 2021.
  20. X. Chen, L. Feng, R. Yao, X. Wu, J. Sun, and W. Gong, "Prediction of maize yield at the city level in China using multi-source data," *Remote Sens.*, vol. 13, no. 1, p. 146, 2021.
  21. L. S. Cedric, W. Y. H. Adoni, R. Aworka, J. T. Zoueu, and F. K. Mutombo, "Crops yield prediction based on machine learning models: Case of West African countries," 2022.
  22. M. Shahhosseini, G. Hu, and S. V Archontoulis, "Forecasting corn yield with machine learning ensembles," *Front. Plant Sci.*, vol. 11, p. 1120, 2020.
  23. S. Khaki, L. Wang, and S. V Archontoulis, "A CNN-RNN framework for crop yield prediction," *Front. Plant Sci.*, vol. 10, p. 1750, 2020.
  24. W. Mupangwa, L. Chipindu, I. Nyagumbo, S. Mkuhlani, and G. Sisito, "Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa," *SN Appl. Sci.*, vol. 2, no. 5, p. 952, 2020.
  25. Y. Everingham, J. Sexton, D. Skocaj, and G. Inman-Bamber, "Accurate prediction of sugarcane yield using a random forest algorithm," *Agron. Sustain. Dev.*, vol. 36, no. 2, p. 27, 2016.
  26. P. Feng, B. Wang, D. Li Liu, C. Waters, and Q. Yu, "Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia," *Agric. For. Meteorol.*, vol. 275, pp. 100–113, 2019.
  27. S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*, vol. 72. 2015.

28. E. H. Alkhamash, M. Hadjouni, and A. M. Elshewey, "A Hybrid Ensemble Stacking Model for Gender Voice Recognition Approach," *Electron.*, vol. 11, no. 11, pp. 1-13, 2022.
29. V. R. R. Kolipaka and A. Namburu, "K-Fold Validation of Multi Models for Crop Yield Prediction with Improved Sparse Data Clustering Process," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 10s, pp. 454-463, 2023.