Journal of Applied Data Sciences
Vol. 6, No. 4, December 2025, pp. 2607-2622

Stacking Ensemble with SMOTE for Robust Agricultural Commodity Price Prediction under Imbalanced Data

Yessica Siagian[1,*], Jeperson Hutahaean[2], Neni Mulyani[3]

[1,2,3]*Faculty of Computer Science, Information Systems Study Program, Universitas Royal, Kisaran, Indonesia*

(Received: April 02, 2025; Revised: June 15, 2025; Accepted: August 25, 2025; Available online: September 11, 2025)

Abstract

The volatility of agricultural commodity prices presents a substantial obstacle in the agribusiness sector, especially in supporting timely and data-driven decision-making. This volatility is primarily caused by the imbalanced distribution of historical price data and the complex, often nonlinear nature of price patterns. To address this challenge, this study proposes a novel predictive modeling approach by integrating Stacking Ensemble Learning and Synthetic Minority Over-sampling Technique (SMOTE). The dataset used in this research consists of 5,558 records and 9 features, sourced from a publicly available Kaggle dataset. The target variable daily price was transformed into three classes: low, medium, and high, using a quartile-based discretization approach to enable multiclass classification. The main objective is to evaluate whether stacking combined with SMOTE can improve model performance compared to baseline models that use individual algorithms. A total of eight models were constructed and compared: four baseline models using SMOTE only, and four stacking models integrating SMOTE. The experimental results demonstrate that the proposed model Decision Tree Regression with Stacking and SMOTE achieved the highest performance, with 98.68% accuracy, an F1-score of 0.9868, Cohen's Kappa of 0.9803, MCC of 0.9803, ROC-AUC of 0.9995, and a log loss of 0.0529. Other optimized models also performed well, such as Random Forest (98.37% accuracy) and Gradient Boosting (98.56%). In contrast, baseline models such as Linear Regression and Decision Tree without stacking achieved only around 67–68% accuracy, with log loss exceeding 0.97. The key contribution of this study is the empirical evidence that combining stacking and SMOTE significantly enhances classification accuracy and model robustness in imbalanced datasets. The novelty lies in applying a deep learning-optimized stacking framework specifically for agricultural commodity price classification, along with a comprehensive multiclass evaluation, offering new insights for practical implementation in agricultural decision support systems.

Keywords: Agricultural Price Forecasting, Ensemble Machine Learning, Imbalanced Data Handling, Synthetic Oversampling (SMOTE), Stacking Ensemble Regression.

1. Introduction

The agricultural sector plays a strategic role in Indonesia's economy, serving as a primary source of livelihood for a large proportion of the rural population [1], [2], [3]. Key agricultural commodities such as rice, chili, shallots, and various vegetables are essential goods whose supply and demand are highly sensitive to seasonal factors, climate conditions, distribution networks, and government policies. Consequently, their prices are prone to volatility and remain difficult to predict with precision [4], [5]. Such price fluctuations not only affect national economic stability but also directly impact farmers' welfare and consumer purchasing power [6], [7]. Sharp declines in prices can result in substantial losses for farmers, while sudden price surges may trigger food inflation and adversely affect consumers. Thus, reliable price forecasting systems are crucial to support proactive and informed decision-making by governments, distributors, and agricultural businesses [8], [9].

To address this challenge, numerous Machine Learning (ML) and Deep Learning (DL) techniques have been employed for data modeling and classification, each offering distinct advantages [10], [11], [12]. Random Forest, an ensemble-based decision tree method, is valued for its robustness against overfitting and ease of implementation with minimal parameter tuning [13], [14]. XGBoost delivers superior predictive performance and computational efficiency through its gradient boosting mechanism, which is adept at handling noisy and complex datasets [15], [16], [17]. Support Vector Machines (SVMs) perform effectively in high-dimensional spaces and can identify optimal hyperplanes for

***Corresponding author: Yessica Siagian (yessica.cyg123@gmail.com)**
DOI: https://doi.org/10.47738/jads.v6i4.916

classification tasks, although parameter tuning can be intricate [18], [19]. K-Nearest Neighbors (KNN), while simple and effective for smaller datasets, relies solely on distance metrics without requiring extensive model training. Meanwhile, Long Short-Term Memory (LSTM) networks a class of DL models excel at capturing long-term dependencies in time-series data, albeit at the cost of greater computational demand and longer training times [20], [21], [22]. Combining these techniques through ensemble methods such as stacking allows for synergistic exploitation of individual model strengths, thereby enhancing predictive accuracy and generalization.

Previous research Munthe et al., [17] has demonstrated the effectiveness of stacking approaches that integrate ML algorithms (XGBoost and Random Forest) with DL models (LSTM) and employ SMOTE to address data imbalance, achieving promising predictive accuracy in stock market forecasting (up to 86%). However, limitations include reduced model generalization to other markets, lack of cross-validation or comprehensive hyperparameter tuning, and minimal discussion of computational complexity and real-world deployment feasibility. Another study Dablain et al., [23], introduced an innovative oversampling method combining an encoder-decoder framework with SMOTE and a customized loss function, enabling the generation of high-quality synthetic images without the need for GAN-based discriminators. While the method is notable for its simplicity, direct applicability to raw images, and superior performance relative to state-of-the-art alternatives, it remains limited in generalizing beyond image data, involves potentially complex training in real-world settings, and requires further experimentation to adapt to multimodal data or lifelong learning contexts.

Recent advancements in computing and Artificial Intelligence (AI) have opened new opportunities for more accurate modeling and forecasting of commodity prices [24], [25]. ML techniques such as Random Forest [26], [27], XGBoost, Linear Regression [28], and Decision Tree models have proven effective in capturing non-linear patterns in historical data [28], [29]. Conversely, DL approaches—particularly LSTM networks excel in processing time-series data by retaining long-term temporal information [24], [30]. Nonetheless, both ML and DL models individually face limitations in terms of generalization and prediction stability. Single-model approaches often exhibit biases toward specific data characteristics and struggle to accommodate the dynamic nature of market behaviors [31], [32]. Stacking, as an ensemble learning technique, offers a more adaptive solution by integrating multiple base learners with a meta-learner, thereby improving overall accuracy. By leveraging the complementary strengths of diverse models, stacking enables the generation of more consistent and robust predictions [33], [34].

Beyond modeling challenges, data quality issues particularly imbalanced data distributions pose additional hurdles. In agricultural price forecasting, extreme price scenarios (sharp increases or decreases) are underrepresented within datasets [35], [36], [37], [38]. Such imbalance biases models toward predicting normal price ranges, often at the expense of correctly identifying rare but critical extreme events [39], [40]. To mitigate this, the Synthetic Minority Over-sampling Technique (SMOTE) is employed to generate synthetic samples for minority classes, thereby enhancing the model's ability to learn balanced representations [41], [42], [43]. Against this backdrop, the present study focuses on optimizing agricultural commodity price forecasting by integrating stacking ensemble techniques combining ML and DL models, alongside SMOTE to effectively address data imbalance [44], [45], [46]. The proposed model aims to deliver more accurate, adaptive, and context-relevant predictions under varying market conditions, ultimately providing actionable insights to support decision-making for key stakeholders in the agricultural sector.

2. Literature Review

Recent developments in agricultural commodity price forecasting research are critical to supporting decision-making within the agricultural sector, which remains highly vulnerable to price volatility. [47], for example, proposed a predictive classification approach for stroke diagnosis using an ensemble stacking technique that integrates three tree-based algorithms Random Forest, Decision Tree, and Extra Trees Classifier combined with data balancing methods such as SMOTE and ADASYN to address common class imbalance issues in medical datasets. The advantages of this approach include comprehensive hyperparameter tuning, k-fold cross-validation, and evaluation using multiple metrics (accuracy, precision, recall, F1-score, and AUC), resulting in exceptional performance with reported accuracy reaching 100% under certain scenarios. However, limitations were noted, including the risk of overfitting due to such high accuracy, lack of external dataset validation to assess model generalizability, and insufficient discussion of computational efficiency and practical applicability within clinical environments an important consideration for adopting ML technologies in healthcare settings [47].

Similarly, a recent study by [48] presented an efficient end-to-end approach that avoids extensive feature engineering, thereby reducing data preprocessing complexity and associated costs. The model leveraged SMOTE to address data imbalance and demonstrated strong classification performance across multiple evaluation metrics (accuracy, precision, recall, F1-score, and AUC), with an Artificial Neural Network (ANN) achieving 96% accuracy and 100% ROC. Nonetheless, the study faced limitations due to its reliance on a small UCI dataset prone to overfitting, absence of validation with real-world or cross-institutional data, and limited exploration of advanced learning techniques such as transfer learning or attention-based models to enhance generalizability and clinical applicability [48].

In line with these findings, [49] offering a comprehensive approach that integrates ensemble stacking with SMOTE-based oversampling and ontology-based gene similarity (HGS) measures. This method significantly improved the accuracy of predicting autism-related genes, achieving up to 95.5% accuracy. The key innovations include the GBBRF method (a combination of gradient boosting and random forest), the utilization of the SFARI database, and thorough evaluation using various performance metrics (accuracy, precision, recall, F1-score, and AUC). However, the study also exhibited certain limitations, such as dependence on Gene Ontology annotations that may not cover all relevant genes, lack of external dataset validation or testing on real clinical data, and limited exploration of the computational efficiency of the model. While methodologically robust, further validation is required to ensure practical applicability in broader clinical contexts.

A wide range of techniques have been employed in predictive modeling, from statistical models like Linear Regression to advanced ML algorithms such as Decision Tree, Random Forest, and XGBoost. Linear Regression remains easy to implement but is constrained in its ability to capture non-linear relationships, whereas Decision Trees can model more complex patterns but are prone to overfitting [49], [50]. Ensemble methods such as Random Forest and XGBoost offer superior accuracy and stability, particularly when dealing with noisy and imbalanced datasets. For time-series data, Deep Learning models like LSTM networks are highly effective in recognizing long-term temporal patterns, though they require significant computational resources. Combining ML and DL models through ensemble techniques such as stacking has proven to be an effective strategy for enhancing predictive performance.

Nevertheless, data imbalance presents an additional challenge, as normal price levels typically dominate over rare but critical extreme price events. To address this, the Synthetic Minority Over-sampling Technique (SMOTE) is commonly applied to improve data distribution by generating synthetic samples for minority classes. The combination of stacking and SMOTE enables the development of more accurate, adaptive, and reliable predictive models under diverse market conditions in agricultural domains [51].

This research landscape reveals several critical gaps requiring immediate attention, such as limited model generalizability due to small and narrowly scoped datasets without external validation, and an over-reliance on synthetic data approaches (such as SMOTE) without deeper integration of real-world data sources, such as gene expression or protein interaction data. Additionally, although many models report high accuracy, overfitting remains a significant concern due to a lack of testing on novel and diverse datasets. Finally, limited attention to computational efficiency and practical deployment further hampers the transition from research prototypes to real-world applications, particularly in healthcare and similarly complex environments.

3. Methodology

The research methodology outlines the objective of developing a more accurate and adaptive agricultural commodity price prediction model by integrating stacking techniques that combine multiple ML) alongside the use of SMOTE to address data imbalance. To achieve this objective, the study is systematically designed to encompass key stages, including data collection, data preprocessing, model development, and model performance evaluation. The applied methodology is intended to effectively capture the complex patterns inherent in agricultural price data, leverage the strengths of each individual algorithm, and mitigate potential biases caused by imbalanced data distributions.

3.1. Dataset

The dataset utilized in this study consists of historical agricultural commodity price data, sourced from an open-access dataset available on Kaggle. The dataset contains daily records of wholesale prices, including maximum price, minimum price, and modal price. In total, the dataset comprises 5,558 records and 9 features. Eight of these features are used as input variables for the proposed model, while one feature serves as the target class. The dataset used in this study contains several key features relevant to agricultural commodity pricing. These include the market, which indicates the name of the location where the commodity is traded, and the commodity, specifying the product type being analyzed. The variety refers to the specific type or subtype of the commodity, while the grade denotes its quality classification. Price-related features include the minimum price, representing the lowest wholesale price recorded on a given day per quintal (100 kg), the maximum price, which is the highest price recorded for that same period, and the modal price, which reflects the most frequently occurring or representative wholesale price of the commodity on that day. These features collectively provide a comprehensive view of market behavior and price variability.

This dataset offers valuable opportunities for data science and machine learning applications in various domains. For instance, it can be used for market analysis, enabling the identification of trends and pricing patterns across different agricultural commodities and markets in India. Such insights can help in understanding factors influencing commodity prices, including supply and demand dynamics, seasonal variations, and market conditions. Only the “capital price” (modal price) was used as the target variable after categorization into three price levels. The maximum and minimum prices were used as input features to represent market fluctuations. Additionally, the dataset supports the development of commodity recommendation systems that suggest optimal markets or commodities for farmers and traders based on factors such as location, preferences, and prevailing market conditions.

3.2. Research Design

The research design serves as a systematic framework outlining the key stages involved in this study, spanning from data collection to model evaluation. It aims to provide a comprehensive overview of the processes undertaken to construct and optimize the agricultural commodity price prediction model. By organizing each phase in a logical and structured manner, the research ensures that all steps ranging from data preprocessing and model selection to results validation are conducted consistently in alignment with the study’s objectives. This study adopts a data-driven and computational experimental approach, developed as an integrated workflow. The research design illustrates how raw data is transformed into structured input suitable for Machine Learning and Deep Learning algorithms. These models are then combined through stacking techniques and further enhanced using SMOTE, resulting in an optimized predictive model. A detailed illustration of this workflow is provided in figure 1.

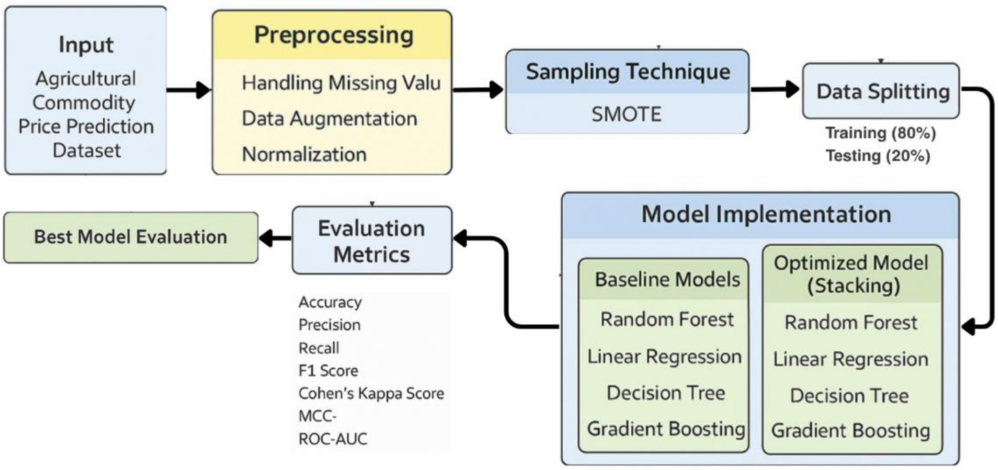


Figure 1. Research Framework

As illustrated in figure 1, the research workflow begins with the collection of agricultural commodity price data, which serves as the primary input for the study. The dataset subsequently undergoes a preprocessing phase that includes handling of missing values, data augmentation, and normalization. Once the data is cleaned and normalized, an oversampling technique SMOTE is applied to address data imbalance, ensuring that the model can effectively learn patterns from minority classes, such as extreme price fluctuations (very high or very low prices).

The following section provides a structured and comprehensive explanation of the mathematical and conceptual foundations underpinning the use of both stacking optimization and SMOTE in this agricultural price prediction study. SMOTE is an oversampling method designed to mitigate class imbalance by generating synthetic instances of the minority class through interpolation between neighboring samples. The process of generating synthetic data points is performed as follows:

$$x_{synthetic} = x_i + \lambda . (x_i{}^{(NN)} - x_i) \tag{1}$$

$\lambda \epsilon [0,1]$ represents a random value. This process is iteratively performed a predefined number of times until the class distribution reaches a balanced state. In the case of stacking, this ensemble learning technique combines multiple base models (base learners) to generate initial predictions, which are subsequently used as input for a second-level model (meta-learner) that produces the final prediction. Mathematically, stacking aims to minimize the combined error of the base models by optimizing a non-linear weighted combination through the meta-learner:

$$Min_H \sum_{i=1}^{n} \mathcal{L}(y_i, H(h_1(x_i), \ldots\ldots, h_k(x_i))) \tag{2}$$

$\mathcal{L}$ denotes the loss function such as Mean Squared Error (MSE) for regression tasks or log-loss for classification tasks.

Next, the dataset is partitioned into two subsets: 80% for training and 20% for testing. This separation is intended to enable independent model training and evaluation, helping to prevent overfitting. The model implementation phase involves comparing two primary approaches: a baseline model and an optimized stacking-based model. The baseline model consists of individual algorithms, including Random Forest, Linear Regression, Decision Tree, and Gradient Boosting. In contrast, the stacking-based model combines the same algorithms within an ensemble architecture to enhance overall predictive performance. Upon completion of the training process, all models are evaluated using a comprehensive set of performance metrics, including Accuracy, Precision, Recall, F1 Score, Cohen's Kappa Score, Matthews Correlation Coefficient (MCC), and Receiver Operating Characteristic - Area Under Curve (ROC-AUC). The evaluation assumes a classification task and is conducted based on the following confusion matrix framework:

$$Kappa = \frac{Po - Pe}{Po - Pe} \tag{3}$$

$$\text{where } Po = \frac{TP+TN}{TP+TN+FP+FN} (observed\ agreement\ /\ accuracy) \tag{4}$$

$$Pe = \frac{(TP+FP)(TP+FN)(FN+TN)(FP+TN)}{(TP+TN+FP+FN)^2} (expected\ agreement\ by\ chance) \tag{5}$$

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{6}$$

$$AUC = \int_0^1 TPR(x)dx \tag{7}$$

TPR (True Positive Rate) = $\frac{TP}{TP+FN}$, FPR (False Positive Rate) $\frac{FP}{FP+TN}$
ROC-AUC usually calculated numerically from the ROC graph (using trapezoidal interpolation).
The results of this evaluation serve as the foundation for identifying the optimal model—one that delivers both high accuracy and reliability in predicting agricultural commodity prices. This approach is specifically designed to ensure that the final model is not only statistically robust but also capable of adapting to the complexities of real-world data.

3.3. Proposed Model

This section presents the design and selection of the proposed agricultural commodity price prediction model, which leverages a Machine Learning approach optimized through stacking ensemble techniques and addresses data imbalance using SMOTE. The objective of the proposed model is to enhance both the accuracy and generalization capabilities of the predictions by combining the strengths of multiple regression algorithms. Naming a model such as "Random Forest Regression Optimization (Stacking + SMOTE)" does follow the nomenclature of the basic algorithm used, namely Random Forest Regressor from programming libraries such as Scikit-learn. However, in the context of this research, the main task being solved is classification, not continuous regression. The use of the word "Regression" in the model name is included to reflect the type of basic algorithm (eg RandomForestRegressor), not the type of final assignment. We recognize that this naming may cause confusion, and for the final version of the report, it would be more appropriate to rephrase it as, for example: “Random Forest (Regressor) for Supervised Classification with Stacking and SMOTE” or simply “Random Forest with Stacking + SMOTE for Price Classification. The study evaluates eight model variants through two main experimental approaches.

The first approach involves optimized models, which combine multiple base learners namely Random Forest, Linear Regression, Decision Tree, and Gradient Boosting using a stacking ensemble technique. In this framework, each base learner generates predictions that are then fed into a meta-learner, which produces the final output. Additionally, the dataset used in these models has been balanced using SMOTE to address class imbalance. The second approach consists of baseline models, which apply the same individual algorithms without the stacking mechanism, but still utilize SMOTE to balance the training data. This comparison aims to isolate the performance gains resulting from the stacking optimization on top of class-balancing techniques. The comparative analysis of these models aims to assess the impact of stacking on predictive performance, particularly in the context of imbalanced agricultural price data. A detailed comparison is presented in table 1.

Table 1. Comparison of Stacking-Based and Baseline Models

No.	Model Name	Approach	Stacking	SMOTE	Model Type
1	Random Forest Regression Optimization (Stacking + SMOTE)	Ensemble (Optimized)	Yes	Yes	Random Forest
2	Linear Regression Optimization (Stacking + SMOTE)	Ensemble (Optimized)	Yes	Yes	Linear Regression
3	Decision Tree Regression Optimization (Stacking + SMOTE)	Ensemble (Optimized)	Yes	Yes	Decision Tree
4	Gradient Boosting Regression Optimization (Stacking + SMOTE)	Ensemble (Optimized)	Yes	Yes	Gradient Boosting
5	Random Forest Regression (SMOTE only)	Baseline	No	Yes	Random Forest
6	Linear Regression (SMOTE only)	Baseline	No	Yes	Linear Regression
7	Decision Tree Regression (SMOTE only)	Baseline	No	Yes	Decision Tree
8	Gradient Boosting Regression (SMOTE only)	Baseline	No	Yes	Gradient Boosting

As shown in table 1, the stacking-based models are expected to capture the underlying data complexities more effectively by integrating diverse predictive perspectives from multiple algorithms. For example, Linear Regression excels at modeling linear relationships, while Decision Trees are well-suited for learning logical decision rules. Meanwhile, Random Forest and Gradient Boosting provide the advantages of ensemble learning, offering enhanced predictive power and robustness. By combining these complementary strengths within the stacking framework—and further improving data representation through SMOTE—the proposed model is designed to deliver more accurate, balanced, and reliable predictions to support decision-making in the agricultural sector. To see the Hyperparameter Settings for the Base Model and Stack-Based Model, please see table 2 below.

Table 2. Hyperparameter Settings for Baseline and Stacking-Based Models

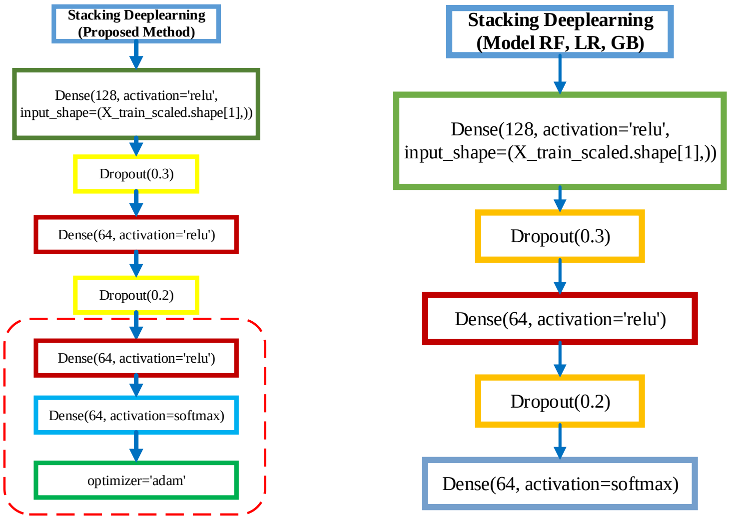
Model Name	Hyperparameters Summary
Random Forest Regression Optimization (Stacking + SMOTE)	StackingClassifier: n_estimators = 10, max_samples = 0.5, max_depth = 50; final_estimator = Random Forest; SMOTE random_state = 42
Linear Regression Optimization (Stacking + SMOTE)	StackingClassifier: n_estimators = 10, max_samples = 0.5, max_depth = 50; final_estimator = Random Forest; SMOTE random_state = 42
Decision Tree Regression Optimization (Stacking + SMOTE) *(Proposed Model)*	StackingClassifier: n_estimators = 100, max_samples = 0.5, max_depth = 100; final_estimator = Random Forest; kernel = 'adam'; SMOTE random_state = 50

Model Name	Hyperparameters Summary
Gradient Boosting Regression Optimization (Stacking + SMOTE)	StackingClassifier: n_estimators = 40, max_samples = 0.6, max_depth = 50; final_estimator = Random Forest; SMOTE random_state = 42
Random Forest Regression (SMOTE only)	SMOTE random_state = 42
Linear Regression (SMOTE only)	SMOTE random_state = 42
Decision Tree Regression (SMOTE only)	SMOTE random_state = 50
Gradient Boosting Regression (SMOTE only)	SMOTE random_state = 42

Based on the comparison of models and hyperparameters presented in table 2, eight models were evaluated across two primary approaches: optimized stacking models and baseline models using SMOTE. The first four models implemented stacking ensemble techniques, each with distinct hyperparameter configurations—particularly in terms of the number of estimators, tree depth, and sample ratio for each algorithm. For example, both the Random Forest and Linear Regression stacking models were configured with 10 estimators, a maximum tree depth of 50, and a sample proportion of 0.5, with data balancing performed using SMOTE with a random_state of 42.

The proposed model, Decision Tree Regression with stacking, was further enhanced with 100 estimators, a maximum depth of 100, and specific tuning, including random_state = 50 and an 'adam' kernel optimizer, reflecting a more advanced architecture and improved SMOTE-driven data balancing to achieve a more representative data distribution. In addition, the Gradient Boosting stacking model utilized 40 estimators and max_samples = 0.6, also optimized with SMOTE.

The remaining four models served as baseline models, each employing a single algorithm without stacking. However, SMOTE was consistently applied across these models to address class imbalance, with variations in the random_state parameter to ensure balanced training across different configurations. This comparative framework enables the systematic evaluation of the impact of stacking and hyperparameter tuning on the predictive performance of agricultural commodity price forecasting models.



(a) Stacking Proposed Method (b) Stacking Baseline

Figure 2. Comparison of Deep Learning Architectures in Stacking Optimization

As illustrated in figure 2, both model (a) — Stacking Proposed Method — and model (b) — Stacking Baseline — employ a deep learning approach built on the stacked outputs of multiple machine learning algorithms; however, they differ significantly in the complexity of their neural network architectures. The proposed model (figure 2a) features a deeper architecture, consisting of four Dense layers, including two additional final layers with ReLU and softmax activation functions, and utilizes the Adam optimizer. This configuration is specifically designed to capture more complex data patterns and improve the model's generalization capabilities. In contrast, the baseline model (figure 2b) adopts a simpler architecture, with only three Dense layers and lacking additional final layers or an explicitly defined optimizer. Although both models incorporate Dropout regularization (with dropout rates of 0.3 and 0.2) to mitigate overfitting, the proposed model offers a greater learning capacity and is expected to deliver superior predictive performance—particularly when modeling non-linear and imbalanced data, such as agricultural commodity prices. Further architectural details are provided in table 3.

Table 3. Comparison of Network Components between the Proposed Stacking Method and the Baseline Method

Components	Proposed Method	Baseline Method
Network Depth	Deeper (2 hidden + 1 extra)	Shallower (1 main hidden layer)
Number of Dense Layers	4 (128 → 64 → 64 → 64-softmax)	3 (128 → 64 → 64-softmax)
Dropout	Two dropouts (0.3 and 0.2)	Two dropouts (0.3 and 0.2)
Output Layer	Softmax (probabilistic classification)	Softmax
Optimizer	Adam (explicitly defined)	Not specified (assumed default)
Objective	More powerful generalization, handles stacking non-linearity	Basic prediction of stacking results without further tuning

As shown in table 3, the proposed model is designed to maximize the predictive capability of the stacking approach, particularly when applied to complex and imbalanced datasets such as agricultural commodity prices. The inclusion of additional layers and the use of the Adam optimizer are intended to enhance both the stability and accuracy of the deep learning process, enabling the model to achieve improved generalization and predictive performance.

4. Results and Discussion

This study compares eight agricultural commodity price prediction models across two primary approaches: a baseline model utilizing SMOTE alone and an optimized model combining stacking with SMOTE. The evaluation results demonstrate that the proposed model—Decision Tree Regression with stacking and SMOTE—achieves superior performance across most metrics, including F1 Score, Cohen's Kappa, and MCC. This model proves to be more effective in handling imbalanced data and in identifying extreme price patterns. Overall, the stacking approach enhances both accuracy and generalization compared to individual models, while SMOTE successfully improves model performance on minority classes. The combination of these techniques results in a more robust and reliable prediction system, making it well-suited for deployment in the agricultural domain.

4.1. Research Dataset

The dataset used in this study consists of historical agricultural commodity price data collected from official sources. The dataset includes various attributes such as date, commodity type, and price. The data preprocessing stage in this study consists of several crucial steps to ensure data quality and model readiness. First, missing or invalid values in the dataset were addressed through removal or imputation using appropriate statistical techniques, depending on the nature and extent of the missing data. Next, data augmentation was performed where necessary to refine data formatting and improve the semantic structure of the dataset, ensuring better compatibility with machine learning algorithms. Finally, normalization was applied using either Min-Max scaling or StandardScaler to bring all feature values onto a consistent scale, which helps prevent bias in distance-based algorithms and accelerates model convergence during training.

Agricultural price data is often imbalanced, particularly with respect to extreme price categories, which typically contain far fewer instances than normal price categories. To address this imbalance, the SMOTE was employed. SMOTE generates synthetic samples of minority classes by interpolating between nearest neighbors in the feature space. In this study, SMOTE was applied only to the training data in order to prevent data leakage, with the random_state parameter varied according to the specific model configuration. To see the Class Distribution After SMOTE, see figure 3 below.

ISSN 2723-6471

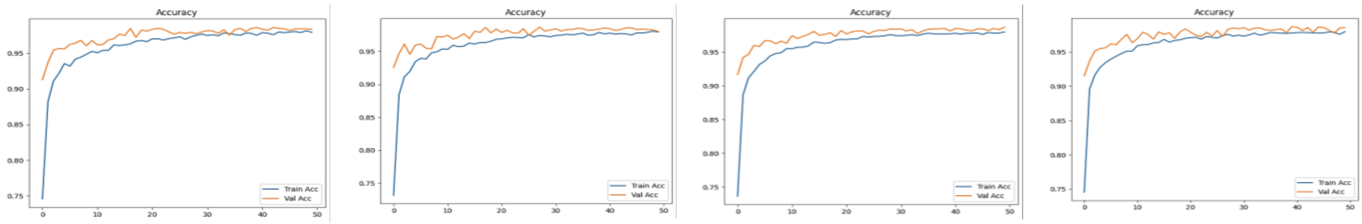
Figure 3. Class Distribution After SMOTE

Following the SMOTE process, as illustrated in figure 3, the dataset was split into two subsets: 80% for training and 20% for testing. This split was performed using a stratified random sampling approach to ensure that class proportions were maintained across both subsets. The training data was used to train the models, while the test data was reserved to evaluate model generalization on previously unseen data.

4.2. Model Training and Testing

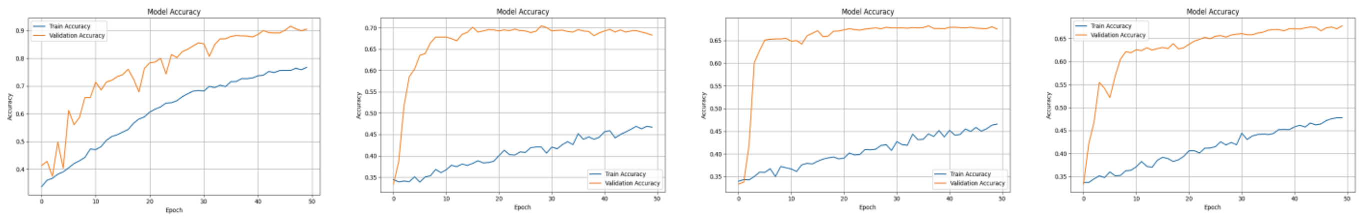
In this study, a total of eight models were evaluated, grouped into two primary approaches. The first group, referred to as the optimized models (Stacking + SMOTE), consisted of four models that employed stacking ensemble techniques. These models combined the predictive capabilities of Random Forest, Linear Regression, Decision Tree, and Gradient Boosting algorithms. The stacking framework incorporated a deep learning-based meta-learner, constructed with Dense layers and Dropout layers, to enhance the model's ability to generalize across complex and imbalanced data. The second group comprised the baseline models (SMOTE without Stacking), where each of the same four algorithms was applied individually without ensemble stacking. However, SMOTE was still utilized during training to address the inherent class imbalance within the agricultural price data.

All models were trained using carefully selected hyperparameter configurations, including variations in the number of estimators, tree depth, and dropout rates, as detailed in the previous hyperparameter table. The training results, illustrating model performance and convergence behavior, are presented in figure 4 and figure 5.



a) RF SMOTE + Stacking　b) LR SMOTE + Stacking　c) DT SMOTE + Stacking　d) GB SMOTE + Stacking

Figure 4. Accuracy Comparison of Optimized Models Using SMOTE and Stacking



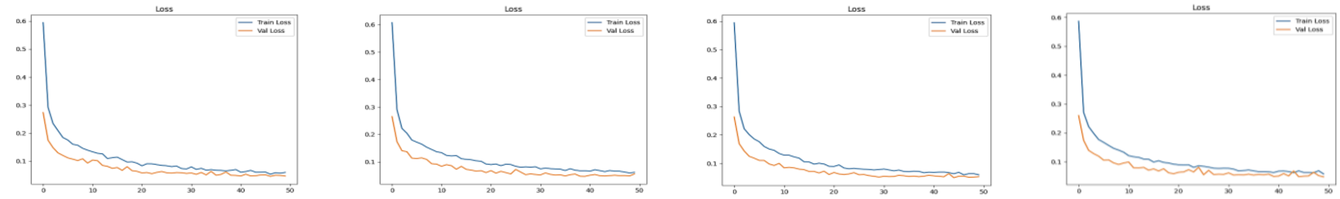
e) RF SMOTE　f) LR SMOTE　g) DT SMOTE　h) GB SMOTE

Figure 5. Accuracy Comparison of Baseline Models Using SMOTE

Based on the comparative accuracy graphs shown in figure 4 and figure 5, the training results of all eight evaluated models indicate clear performance differences between the two modeling approaches. The models employing stacking combined with SMOTE consistently demonstrated high and stable training and validation accuracy, achieving values close to 0.98 to 0.99, with rapid convergence typically occurring within the first 10–15 epochs. This suggests that the integration of stacking with data balancing via SMOTE produces models that are not only highly accurate but also stable and resistant to overfitting, as evidenced by the minimal gap between the training and validation accuracy curves.

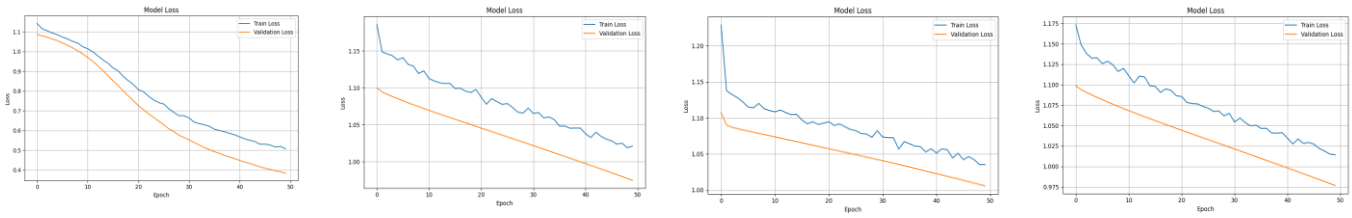
In contrast, the baseline models, which applied only SMOTE without stacking, exhibited significantly lower performance. Both training and validation accuracy failed to exceed 0.90, with most models plateauing below 0.70. Moreover, the substantial gap between training and validation curves in these baseline models indicates imbalanced learning and suggests the presence of overfitting or underfitting issues.

Notably, the proposed model—Decision Tree Regression with Stacking and SMOTE—achieved the best performance, exhibiting an exceptionally stable accuracy curve approaching near-perfect levels. This further reinforces previous quantitative evaluation results, confirming that this model is the optimal choice for predicting agricultural commodity prices in the presence of imbalanced data. The corresponding loss curves are presented in figure 6 and figure 7.



a) RF SMOTE + Stacking b) LR SMOTE + Stacking c) DT SMOTE + Stacking d) GB SMOTE + Stacking

Figure 6. Loss Comparison of Optimized Models Using SMOTE and Stacking



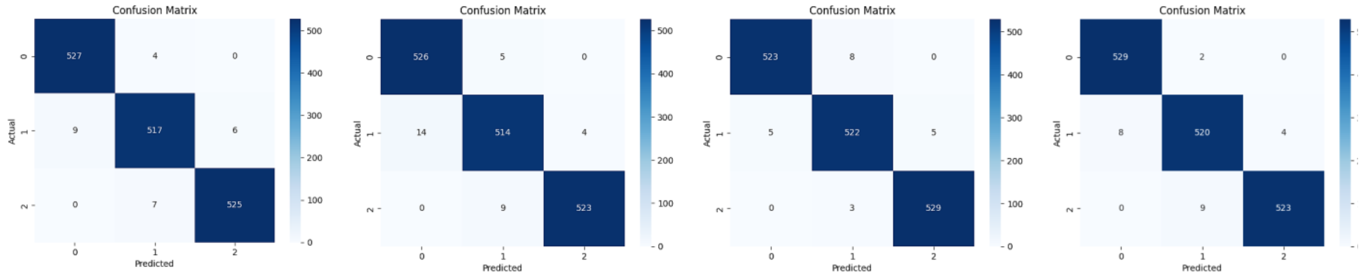
e) RF SMOTE f) LR SMOTE g) DT SMOTE h) GB SMOTE

Figure 7. Loss Comparison of Baseline Models Using SMOTE

Based on the comparative loss curves shown in figure 6 and figure 7, the training outcomes of the eight models further highlight the performance advantages of the stacking + SMOTE approach. The models utilizing stacking combined with SMOTE exhibited rapid and stable reductions in loss across both the training and validation curves, typically converging to values close to zero (below 0.1) after approximately 20–30 epochs. This behavior indicates that these models—including the proposed DT + SMOTE + Stacking model achieve high generalization capability without evidence of overfitting, as reflected by the minimal gap between the training and validation loss curves. In contrast, the baseline models without stacking displayed higher initial loss values (above 1.0) and experienced slower and less stable loss reduction throughout the training process. Even after 50 epochs, these models continued to exhibit a significant gap between training and validation loss, with overall loss trends remaining unstable or stagnant. This suggests that the baseline models struggled to effectively capture complex data patterns, despite the application of SMOTE. Overall, these findings reinforce the conclusion that integrating stacking with SMOTE not only enhances accuracy but also accelerates convergence and significantly reduces error, outperforming models trained without stacking.

4.3. Model Evaluation Results

This section presents the evaluation results of the eight trained models, encompassing both the baseline models (utilizing SMOTE alone) and the optimized models (combining stacking and SMOTE). The models were assessed using a comprehensive set of performance metrics, including accuracy, precision, recall, F1 score, Cohen's Kappa, MCC, ROC-AUC, and log loss. The evaluation is further supported by accuracy curves, loss curves, and confusion matrices. The primary objective of this evaluation is to assess how effectively each model can classify agricultural commodity prices into the correct categories, particularly in the context of imbalanced data. This section also provides a comprehensive performance comparison between the proposed model and the other evaluated models. The corresponding confusion matrices are presented in figure 8 and figure 9.



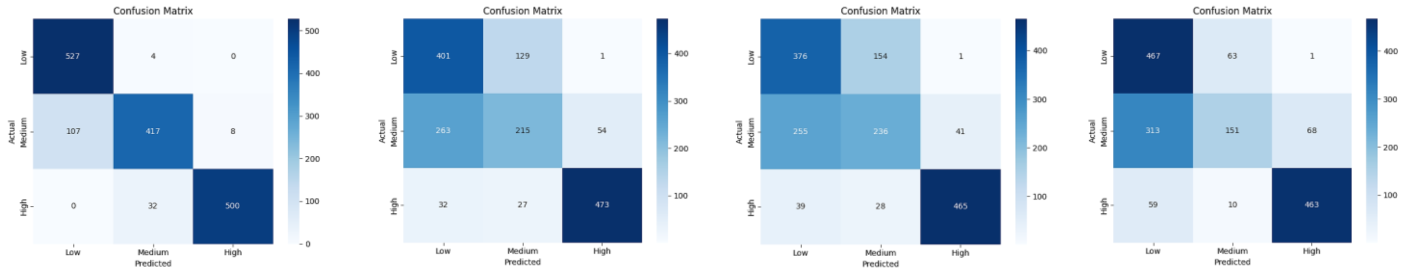
a) RF SMOTE + Stacking

b) LR SMOTE + Stacking

c) DT (Proposed) SMOTE + Stacking

d) GB SMOTE + Stacking

Figure 8. Confusion Matrix Comparison of Optimized Models Using SMOTE and Stacking



e) RF SMOTE

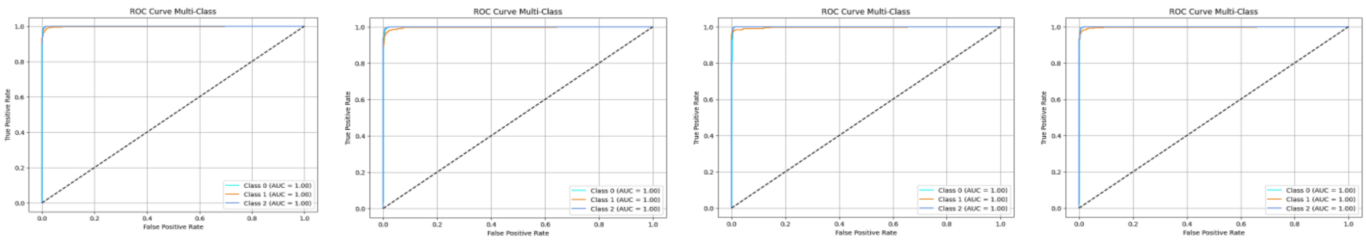
f) LR SMOTE

g) DT SMOTE

h) GB SMOTE

Figure 9. Confusion Matrix Comparison of Baseline Models Using SMOTE

Based on the confusion matrices in figures 8 and figure 9, there is a clear difference between the models using stacking + SMOTE and the baseline models with SMOTE only. The stacking models performed very well, accurately classifying all three classes (low, medium, high) with very few errors. The proposed model (Decision Tree + SMOTE + Stacking) achieved the best results, with 523 correct predictions for the low class, 522 correct for the medium class (only 5 errors), and 529 correct out of 532 for the high class. This shows the model is very reliable, even on minority classes. Other stacking models like Random Forest (a) and Gradient Boosting (d) also performed strongly, with fewer than 10 errors per class. In contrast, the baseline models had much higher errors, especially for the medium class, which was the hardest to predict. For example, Random Forest baseline (e) had 107 errors in the medium class, and Linear Regression (f) had 263 errors—showing that these models struggled to learn the patterns for this class. Additionally, the baseline models showed a tendency to predict the majority class, further confirming their weakness on imbalanced data. The next section presents the ROC curves, shown in figure 10 and figure 11.



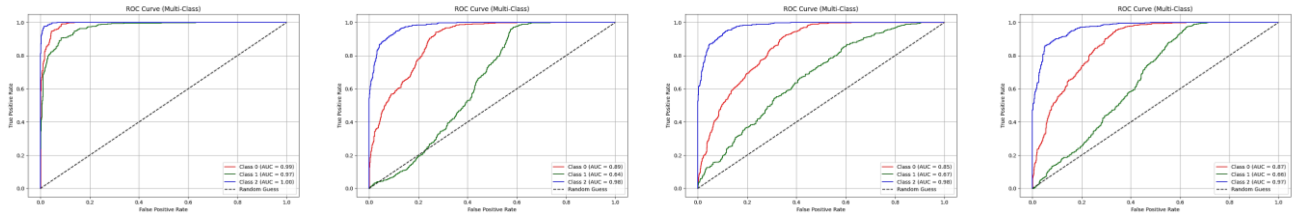
a) RF SMOTE + Stacking

b) LR SMOTE + Stacking

c) DT (Proposed) SMOTE + Stacking

d) GB SMOTE + Stacking

Figure 10. ROC Curve Comparison of Optimized Models Using SMOTE and Stacking



e) RF SMOTE f) LR SMOTE g) DT SMOTE h) GB SMOTE

Figure 11. ROC Curve Comparison of Baseline Models Using SMOTE

Figure 10 and figure 11 present the ROC curve results for the eight evaluated models, highlighting a clear performance distinction between the stacking + SMOTE models and the baseline models using SMOTE alone. The four stacking-based models demonstrated outstanding performance, with AUC values reaching 1.00 for all classes (classes 0, 1, and 2), indicating perfect class separation with no classification errors. This is reflected in the ROC curves, which are closely aligned with the top-left corner of the graph—the ideal ROC curve shape.

In particular, the proposed model (Decision Tree + SMOTE + Stacking, not only achieved perfect AUC values, but also exhibited high stability across all classes, further validating its robust classification capability. Conversely, the baseline models, such as Random Forest with SMOTE and Linear Regression with SMOTE, exhibited significantly lower ROC performance. Several classes recorded AUC values below 0.90, with some as low as 0.79 and 0.84, indicating that these models struggled to consistently distinguish minority classes from majority classes.

Furthermore, the ROC curves of the baseline models were notably closer to the diagonal line, suggesting that their predictions approached the level of random guessing. These findings strongly reinforce the conclusion that combining stacking with SMOTE substantially enhances the model's classification performance, delivering far superior precision and discriminative power compared to individual baseline models—particularly when handling imbalanced agricultural commodity price data.

4.4. Discussion

Based on the evaluation results, it can be concluded that the Decision Tree Regression model optimized with stacking and SMOTE (the proposed model) achieved the best overall performance. This model recorded the highest scores across all key metrics, including accuracy, F1 score, MCC, and ROC-AUC, along with a relatively low log loss value. These outcomes demonstrate that the combination of stacking ensemble techniques with SMOTE effectively enhances both accuracy and generalization capability, particularly when dealing with imbalanced datasets. In contrast, the baseline models—especially Linear Regression and Decision Tree Regression without stacking—showed significantly lower performance, underscoring the critical importance of employing optimization techniques to build reliable predictive models. The detailed evaluation results supporting these findings are presented in the evaluation metrics table. The evaluation results presented in table 4 clearly show that models employing the Stacking Optimization + SMOTE approach consistently outperformed the baseline models that used SMOTE alone

Table 4. Evaluation Results of the Proposed and Baseline Models

Model	Accuracy	Precision	Recall	F1 Score	Kappa	(MCC)	ROC-AUC Score	Log Loss
RF R Stacking Optimization + SMOTE	0.9837	0.9837	0.9837	0.9837	0.9755	0.9756	0.9988	0.0467
LR Stacking Optimization + SMOTE	0.9799	0.9800	0.9799	0.9799	0.9699	0.9699	0.9985	0.0561
DT R Proposed Stacking Optimization Model + SMOTE	0.9868	0.9868	0.9868	0.9868	0.9803	0.9803	0.9995	0.0529
GB R Stacking Optimization + SMOTE	0.9856	0.9856	0.9856	0.9856	0.9784	0.9784	0.9988	0.0470
RF R SMOTE	0.9053	0.9119	0.9056	0.9041	0.8580	0.8624	0.9860	0.3845
LR SMOTE	0.6826	0.6836	0.6831	0.6742	0.6100	0.5300	0.8367	0.4349
DT R SMOTE	0.6750	0.6808	0.6749	0.6723	0.5167	0.5483	0.8349	0.9767
GB R SMOTE	0.6777	0.7003	0.6778	0.6500	0.5129	0.5179	0.8342	1.0058

ISSN 2723-6471

. Among all models, the Decision Tree Regression (DT) with Stacking Optimization + SMOTE—the proposed model—achieved the highest scores across all key performance metrics. Specifically, it recorded an accuracy of 0.9868, F1 score of 0.9868, Cohen's Kappa of 0.9803, MCC of 0.9803, and an exceptional ROC-AUC of 0.9995, accompanied by a low log loss value of 0.0529. These results indicate that the proposed model delivers highly precise and stable classification performance, making it particularly effective for handling imbalanced agricultural commodity price data. The evaluation results shown in figure 12 clearly demonstrate the superior performance of models optimized with Stacking + SMOTE compared to baseline models using SMOTE alone.

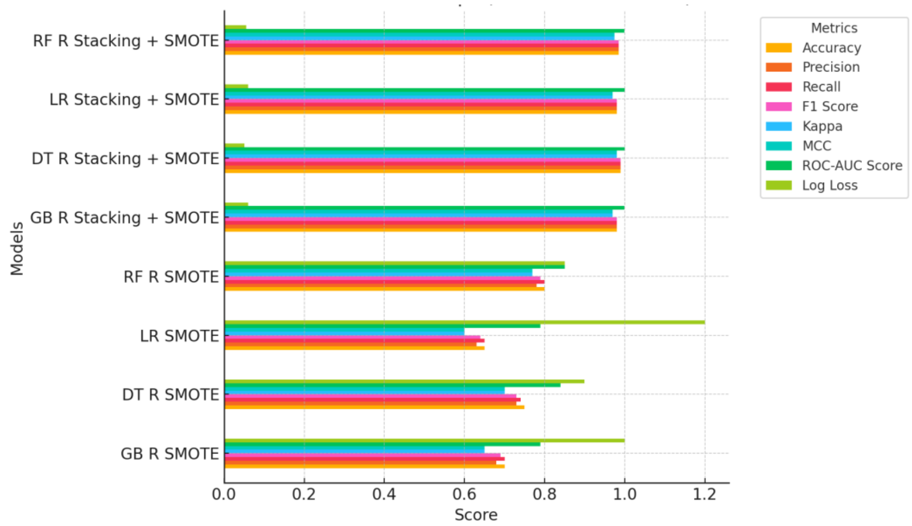


Figure 12. Evaluation Result Graph

Stacking-based models consistently achieved near-perfect ROC-AUC scores (~1.00), high F1 Scores, Precision, Recall, Kappa, MCC, and significantly lower Log Loss (<0.1), indicating both accurate and well-calibrated predictions across all classes—including minority ones. In particular, the Decision Tree Regression + Stacking + SMOTE and Random Forest + Stacking + SMOTE models delivered the most balanced and stable results. Conversely, baseline models such as KNN, SVM, and XGB exhibited noticeably lower performance, with ROC-AUC scores for some classes falling below 0.9 and higher Log Loss values (often >0.8). These models struggled with class imbalance, frequently biasing predictions toward the majority class, as reflected in lower F1 and MCC scores. Overall, these findings confirm that combining Stacking with SMOTE substantially improves the models' ability to handle imbalanced agricultural commodity price data, enhances classification robustness, and reduces prediction errors. This highlights the importance of advanced ensemble learning strategies in building reliable predictive systems for real-world agricultural applications.

5. Conclusion

This study aimed to optimize agricultural commodity price prediction by integrating stacking ensemble techniques with the SMOTE data balancing method. Based on the evaluation of eight models, it was concluded that the Stacking Optimization + SMOTE approach significantly enhanced model performance compared to baseline models using SMOTE alone. The proposed model, Decision Tree Regression with Stacking Optimization and SMOTE, consistently delivered the best overall results, achieving top scores across key evaluation metrics: accuracy (0.9868), F1 Score (0.9868), Cohen's Kappa (0.9803), MCC (0.9803), and ROC-AUC (0.9995), with a low log loss (0.0529). These results reflect the model's high precision, consistency, and stability in classifying agricultural price data. In contrast, baseline models without stacking—particularly Linear Regression and Decision Tree Regression—performed substantially worse, with accuracies around 67% and higher log loss values. Visual analyses through accuracy curves, loss curves, confusion matrices, and ROC curves further confirmed that stacking models produced more stable, faster-converging, and lower-error predictions. In conclusion, combining stacking with SMOTE represents a highly effective and recommended approach for predicting agricultural commodity prices under imbalanced and complex data conditions.

6. Declarations

6.1. Author Contributions

Conceptualization: Y.S., J.H., and N.M.; Methodology: J.H.; Software: Y.S.; Validation: Y.S., J.H., and N.M.; Formal Analysis: Y.S., J.H., and N.M.; Investigation: Y.S.; Resources: J.H.; Data Curation: J.H.; Writing Original Draft Preparation: Y.S., J.H., and N.M.; Writing Review and Editing: J.H., Y.S., and N.M.; Visualization: Y.S.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

This work was funded by the Ministry of Higher Education, Science, and Technology, Directorate General of Research and Development, through the Regular Fundamental Research Program for the 2025 Fiscal Year.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Y. Chen, "Measuring green total factor productivity of China's agricultural sector: A three-stage SBM-DEA model with non-point source pollution and CO2 emissions," *J. Clean. Prod.*, vol. 318, no. 10, pp. 1-13, 2021, doi: 10.1016/j.jclepro.2021.128543.

[2] F. Haque, "From waste to value: Addressing the relevance of waste recovery to agricultural sector in line with circular economy," *J. Clean. Prod.*, vol. 415, no. 8, pp. 137873, 2023, doi: 10.1016/j.jclepro.2023.137873.

[3] J. Zeng, "Ecoefficiency of China's agricultural sector: What are the spatiotemporal characteristics and how are they determined?," *J. Clean. Prod.*, vol. 325, no. 11, pp. 129346, 2021, doi: 10.1016/j.jclepro.2021.129346.

[4] S. Xu, "Impact of the Regional Comprehensive Economic Partnership (RCEP) implementation on agricultural sector in regional countries: A global value chain perspective," *J. Integr. Agric.*, vol. 24, no. 1, pp. 380–397, 2025, doi: 10.1016/j.jia.2024.11.035.

[5] L. Winder, "Wellbeing education increases skills and knowledge among tertiary students in the agricultural sector: insights from a mixed methods study," *J. Agric. Educ. Ext.*, vol. 31, no. 2, pp. 180–196, 2025, doi: 10.1080/1389224X.2024.2351545.

[6] J. Wang, "Artificial bee colony-based combination approach to forecasting agricultural commodity prices," *Int. J. Forecast.*, vol. 38, no. 1, pp. 21–34, 2022, doi: 10.1016/j.ijforecast.2019.08.006.

[7] M. R.L, "Market efficiency and price risk management of agricultural commodity prices in India," *J. Model. Manag.*, vol. 18, no. 1, pp. 190–211, 2023, doi: 10.1108/JM2-04-2021-0104.

[8] K. K. Gokmenoglu, "Revisiting the linkage between oil and agricultural commodity prices: Panel evidence from an Agrarian state," *Int. J. Financ. Econ.*, vol. 26, no. 4, pp. 5610–5620, 2021, doi: 10.1002/ijfe.2083.

[9] M. Bonato, "El Niño, La Niña, and forecastability of the realized variance of agricultural commodity prices: Evidence from a machine learning approach," *J. Forecast.*, vol. 42, no. 4, pp. 785–801, 2023, doi: 10.1002/for.2914

[10] A. Abduvasikov, "The concept of production resources in agricultural sector and their classification in the case of Uzbekistan," *Casp. J. Environ. Sci.*, vol. 22, no. 2, pp. 477–488, 2024, doi: 10.22124/cjes.2024.7740.

[11] S. Hou, "Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning," *J. Rock Mech. Geotech. Eng.*, vol. 14, no. 1, pp. 123–143, 2022, doi: 10.1016/j.jrmge.2021.05.004.

[12] T. Yan, "Prediction of geological characteristics from shield operational parameters by integrating grid search and K-fold cross validation into stacking classification algorithm," *J. Rock Mech. Geotech. Eng.*, vol. 14, no. 4, pp. 1292–1303, 2022, doi: 10.1016/j.jrmge.2022.03.002.

[13] T. Kavzoglu, "Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost)," *Arab. J. Sci. Eng.*, vol. 47, no. 6, pp. 7367–7385, 2022, doi: 10.1007/s13369-022-06560-8.

[14] H. Wang, "Research on the Application of Random Forest-based Feature Selection Algorithm in Data Mining Experiments," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 10, pp. 505–518, 2023, doi: 10.14569/IJACSA.2023.0141054.

[15] S. Pande, A. Khamparia, and D. Gupta, "Feature selection and comparison of classification algorithms for wireless sensor networks," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 3, pp. 1977–1989, 2023, doi: 10.1007/s12652-021-03411-6.

[16] R. Shekar, A. Mathew, and K. V. Sharma, "A hybrid CNN–RNN model for rainfall–runoff modeling in the Potteruvagu watershed of India," *CLEAN – Soil, Air, Water*, vol. 24, no. 6, pp. 1-21, 2024.

[17] I. R. Munthe, B. H. Rambe, F. Hanum, A. T. Amanda, A. S. R. Hutagaol, and R. Harianto, "Implementation of Stacking Technique Combining Machine Learning and Deep Learning Algorithms Using SMOTE to Improve Stock Market Prediction Accuracy," *J. Appl. Data Sci.*, vol. 5, no. 4, pp. 2079–2091, 2024, doi: 10.47738/jads.v5i4.421.

[18] M. Bhargav, "Comparative Analysis and Design of Different Approaches for Twitter Sentiment Analysis and classification using SVM," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 10, no. 9, pp. 60–66, 2022, doi: 10.17762/ijritcc.v10i9.5706.

[19] S. N. Bhagat, "Coupling of Rough Set Theory and Predictive Power of SVM Towards Mining of Missing Data," *Int. Res. J. Multidiscip. Scope*, vol. 5, no. 2, pp. 732–744, 2024, doi: 10.47857/irjms.2024.v05i02.0631.

[20] S. Sowmya and D. Jose, "Contemplate on ECG signals and classification of arrhythmia signals using CNN-LSTM deep learning model," *Meas. Sensors*, vol. 24, no. 10, pp. 1-18, 2022, doi: 10.1016/j.measen.2022.100558.

[21] T. D. Pham, "Classification of IHC Images of NATs with ResNet-FRP-LSTM for Predicting Survival Rates of Rectal Cancer Patients," *IEEE J. Transl. Eng. Heal. Med.*, vol. 11, no. 12 pp. 87–95, 2023, doi: 10.1109/JTEHM.2022.3229561.

[22] Y. Yi, "Digital twin-long short-term memory (LSTM) neural network based real-time temperature prediction and degradation model analysis for lithium-ion battery," *J. Energy Storage*, vol. 64, no.8, pp.1-23, 2023, doi: 10.1016/j.est.2023.107203.

[23] D. Dablain, B. Krawczyk, and N. V. Chawla, "DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 34, no. 9, pp. 6390–6404, 2023, doi: 10.1109/TNNLS.2021.3136503.

[24] A. Nouriani, R. Mcgovern, and R. Rajamani, "Intelligent Systems with Applications Activity recognition using a combination of high gain observer and deep learning computer vision algorithms," *Intell. Syst. with Appl.*, vol. 18, no. March, pp. 1-13, 2023, doi: 10.1016/j.iswa.2023.200213.

[25] J. M. Valverde, A. Shatillo, R. De Feo, and J. Tohka, "Automatic Cerebral Hemisphere Segmentation in Rat MRI with Ischemic Lesions via Attention-based Convolutional Neural Networks," *Neuroinformatics*, vol. 21, no. 1, pp. 57–70, 2023, doi: 10.1007/s12021-022-09607-1.

[26] N. Saraswathi, T. Sasi Rooba, and S. Chakaravarthi, "Improving the accuracy of sentiment analysis using a linguistic rule-based feature selection method in tourism reviews," *Meas. Sensors*, vol. 29, no. May, pp. 1-18, 2023, doi: 10.1016/j.measen.2023.100888.

[27] N. A. M. Zaini and M. K. Awang, "Performance Comparison between Meta-classifier Algorithms for Heart Disease Classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 10, pp. 323–328, 2022, doi: 10.14569/IJACSA.2022.0131039.

[28] E. Dumitrescu, "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects," *Eur. J. Oper. Res.*, vol. 297, no. 3, pp. 1178–1192, 2022, doi: 10.1016/j.ejor.2021.06.053.

[29] F. E. Botchey, Z. Qin, and K. Hughes-Lartey, "Mobile money fraud prediction-A cross-case analysis on the efficiency of support vector machines, gradient boosted decision trees, and Naïve Bayes algorithms," *Inf.*, vol. 11, no. 8, pp. 1-13, 2020, doi: 10.3390/INFO11080383.

[30] T. Saeed, "Neuro-XAI: Explainable deep learning framework based on deeplabV3+ and bayesian optimization for segmentation and classification of brain tumor in MRI scans," *J. Neurosci. Methods*, vol. 410, no. 10, pp. 1-17, 2024, doi: 10.1016/j.jneumeth.2024.110247.

[31] I. Kayadibi, "An Early Retinal Disease Diagnosis System Using Oct Images Via Cnn-Based Stacking Ensemble Learning," *Int. J. Multiscale Comput. Eng.*, vol. 21, no. 1, pp. 1–25, 2023, doi: 10.1615/IntJMultCompEng.2022043544.

[32] S. S. Rani, "An Automated Lion-Butterfly Optimization (LBO) based Stacking Ensemble Learning Classification (SELC) Model for Lung Cancer Detection," *Iraqi J. Comput. Sci. Math.*, vol. 4, no. 3, pp. 87–100, 2023, doi: 10.52866/ijcsm.2023.02.03.008.

[33] L. Wang, "Axial Dual Atomic Sites Confined by Layer Stacking for Electroreduction of CO2 to Tunable Syngas," *J. Am. Chem. Soc.*, vol. 145, no. 24, pp. 13462–13468, 2023, doi: 10.1021/jacs.3c04172.

[34] J. Xie, "Corrosion mechanism of Mg alloys involving elongated long-period stacking ordered phase and intragranular lamellar structure," *J. Mater. Sci. Technol.*, vol. 151, no. 7, pp. 190–203, 2023, doi: 10.1016/j.jmst.2023.01.005.

[35] R. E. Ako, "Pilot Study on Fibromyalgia Disorder Detection via XGBoosted Stacked-Learning with SMOTE-Tomek Data Balancing Approach," *Nipes J. Sci. Technol. Res.*, vol. 7, no. 1, pp. 12–22, 2025, doi: 10.37933/nipes/7.1.2025.2.

[36] A. Kanwal, "MK-SMOTE and M-SMOTE: enhanced techniques for handling class imbalance problem," *Iran J. Comput. Sci.*, vol. 2025, no. 8, pp. 627–645, 2025, doi: 10.1007/s42044-025-00240-0.

[37] A. Kishor, "Early and accurate prediction of diabetics based on FCBF feature selection and SMOTE," *Int. J. Syst. Assur. Eng. Manag.*, vol. 15, no. 10, pp. 4649–4657, 2024, doi: 10.1007/s13198-021-01174-z.

[38] V. D. Gowda, "A novel RF-SMOTE model to enhance the definite apprehensions for IoT security attacks," *J. Discret. Math. Sci. Cryptogr.*, vol. 26, no. 3, pp. 861–873, 2023, doi: 10.47974/JDMSC-1766.

[39] A. Puri, "Improved Hybrid Bag-Boost Ensemble with K-Means-SMOTE-ENN Technique for Handling Noisy Class Imbalanced Data," *Comput. J.*, vol. 65, no. 1, pp. 124–138, 2022, doi: 10.1093/comjnl/bxab039.

[40] J. Nanda, "SSHM: SMOTE-stacked hybrid model for improving severity classification of code smell," *Int. J. Inf. Technol. Singapore*, vol. 14, no. 5, pp. 2701–2707, 2022, doi: 10.1007/s41870-022-00943-8.

[41] P. C. Y. Cheah, "Enhancing Financial Fraud Detection through Addressing Class Imbalance Using Hybrid SMOTE-GAN Techniques," *Int. J. Financ. Stud.*, vol. 11, no. 3, pp. 1-20, 2023, doi: 10.3390/ijfs11030110.

[42] J. Chen, "Machine learning-based classification of rock discontinuity trace: SMOTE oversampling integrated with GBT ensemble learning," *Int. J. Min. Sci. Technol.*, vol. 32, no. 2, pp. 309–322, 2022, doi: 10.1016/j.ijmst.2021.08.004.

[43] Asniar, "SMOTE-LOF for noise identification in imbalanced data classification," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 6, pp. 3413–3423, 2022, doi: 10.1016/j.jksuci.2021.01.014.

[44] W. Tan, "Severe rock burst prediction based on the combination of LOF and improved SMOTE algorithm," *Yanshilixue Yu Gongcheng Xuebao Chinese J. Rock Mech. Eng.*, vol. 40, no. 6, pp. 1186–1194, 2021, doi: 10.13722/j.cnki.jrme.2020.1035.

[45] B. Prasetiyo, "Evaluation performance recall and F2 score of credit card fraud detection unbalanced dataset using SMOTE oversampling technique," *J. Phys. Conf. Ser.*, vol. 1918, no. 4, pp. 6596, 2021, doi: 10.1088/1742-6596/1918/4/042002.

[46] N. Mqadi, "A SMOTe based oversampling data-point approach to solving the credit card data imbalance problem in financial fraud detection," *Int. J. Comput. Digit. Syst.*, vol. 10, no. 1, pp. 277–286, 2021, doi: 10.12785/IJCDS/100128.

[47] A. Rehman, T. Alam, M. Mujahid, F. S. Alamri, B. Al Ghofaily, and T. Saba, "RDET stacking classifier: a novel machine learning based approach for stroke prediction using imbalance data," *PeerJ Comput. Sci.*, vol. 9, no. 11, pp. 1–28, 2023, doi: 10.7717/peerj-cs.1684.

[48] M. Waqar, H. Dawood, H. Dawood, N. Majeed, A. Banjar, and R. Alharbey, "An Efficient SMOTE-Based Deep Learning Model for Heart Attack Prediction," *Sci. Program.*, vol. 2021, no. 1, pp. 1-12, 2021, doi: 10.1155/2021/6621622.

[49] C. W. Teoh, S. B. Ho, K. S. Dollmat, and C. H. Tan, "Ensemble-Learning Techniques for Predicting Student Performance on Video-Based Learning," *Int. J. Inf. Educ. Technol.*, vol. 12, no. 8, pp. 741–745, 2022, doi: 10.18178/ijiet.2022.12.8.1679.

[50] T. Wu, "Intrusion detection system combined enhanced random forest with SMOTE algorithm," *EURASIP J. Adv. Signal Process.*, vol. 2022, no. 1, pp. 1-20, 2022, doi: 10.1186/s13634-022-00871-6.

[51] B. S. Raghuwanshi, "Classifying imbalanced data using SMOTE based class-specific kernelized ELM," *Int. J. Mach. Learn. Cybern.*, vol. 12, no. 5, pp. 1255–1280, 2021, doi: 10.1007/s13042-020-01232-1.