Journal of Agriculture and Food Research 22 (2025) 102021

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

An integrated framework for multi-commodity agricultural price forecasting and anomaly detection using attention-boosted models

Eko Sediyono [a], Kristoko Dwi Hartomo [a], Christian Arthur [b], Intiyas Utami [c],
Ronny Prabowo [c], Raymond Chiong [d,e,*]

[a] Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia
[b] School of Transdisciplinary, University of Technology Sydney, Ultimo, NSW 2007, Australia
[c] Faculty of Economic and Business, Satya Wacana Christian University, Salatiga, Indonesia
[d] School of Science and Technology, University of New England, Parramatta, NSW 2150, Australia
[e] School of Information and Physical Sciences, University of Newcastle, Callaghan, NSW 2308, Australia

ARTICLE INFO

Keywords:
Agricultural commodity
Price forecasting
Transformer
LSTM-VAE
Anomaly detection
Attention mechanism

ABSTRACT

Global agricultural systems are increasingly exposed to price instability driven by climate extremes, logistic disruptions, and market uncertainty. These conditions complicate efforts to monitor and manage price behaviours in essential commodity markets. Micro, small, and medium enterprises (MSMEs), which operate with constrained resources and limited access to data-driven tools, are particularly susceptible to sudden and irregular price shifts. Their ability to maintain stable operations depends on timely identification of market anomalies and reliable planning information. This underscores the importance of accurate price forecasting, yet deep learning models such as Bidirectional Long Short-Term Memory (LSTM) and the Gated Recurrent Unit often struggle to capture long-term dependencies and detect irregular price behaviors. To bridge the gap, this study proposes a deep learning framework that integrates Transformer models for price prediction and an attention-boosted LSTM Variational Autoencoder (VAE) for anomaly detection. Using daily price data collected from the period of January 2020 to mid-June 2024, this study demonstrated that Transformers outperformed traditional models while accurately capturing market trends and sudden fluctuations. Additionally, the attention-boosted anomaly detection model can outperform standard LSTM and artificial neural network-VAEs in identifying unexpected price changes. The proposed models outperformed baseline methods by achieving lower forecasting and anomaly detection errors. By addressing critical limitations in existing forecasting approaches, specifically their inability to capture abrupt anomalies, this study provides essential support for enhancing MSMEs' resilience and improving decision-making under volatile market conditions.

1. Introduction

Price volatility in agricultural commodities is a growing global concern, driven by climate change, supply chain disruptions, and geopolitical tensions that affect production and trade. These global dynamics have introduced persistent uncertainty in commodity pricing across both developed and developing economies [1,2]. Developing countries are particularly exposed due to weaker infrastructure, limited market access, and low adaptive capacity. In Indonesia, for example, these challenges are increased by geographic fragmentation and uneven logistics, especially in remote and underdeveloped regions [3]. With around 61.5 % of micro, small, and medium enterprises (MSMEs) in the country involved in agriculture, a large part of its economy remains vulnerable to price volatility and distribution risks [4,5]. To illustrate the severity of price volatility, the price of red chili, a strategic commodity, increased by 42.7 % between November and December 2023, reaching 79,793 Indonesian Rupiah (Rp) per kilogram (kg) from its average of 55,952 Rp per kg [6]. During the same period, chili prices in East Nusa Tenggara and Maluku surged 114.4 %, from 69,900 to 149, 900 Rp per kg, adding more financial strain on already struggling MSMEs [7]. The volatility of prices within agricultural markets has a disproportionate impact on MSMEs due to their limited capital reserves and weak market positioning, increasing their vulnerability to dramatic rises in input costs [8–10].

---

* Corresponding author. School of Science and Technology, University of New England, Parramatta, NSW 2150, Australia.
E-mail address: rchiong@une.edu.au (R. Chiong).

https://doi.org/10.1016/j.jafr.2025.102021
Received 15 February 2025; Received in revised form 30 April 2025; Accepted 14 May 2025

The economic consequences of this are significant. Approximately 75 % of MSMEs reported an inability to maintain operations in the face of sudden surges in commodity prices, and some businesses have witnessed a 30−50 % decline in revenue while attempting to adjust their pricing models in response to unpredictable market conditions [11–14]. This price volatility poses a critical threat to MSMEs, which are often financially constrained and lack adaptive mechanisms. Despite the availability of forecasting techniques, existing models predominantly focus on trend prediction and fail to capture sudden, irregular fluctuations that are common in volatile markets [15]. This represents a significant forecasting gap that leaves MSMEs quite vulnerable and in urgent need of advanced predictive tools [16]. Failure to predict such anomalies results in suboptimal inventory management, further disrupting business operations [17]. Anomaly detection systems, which employ machine learning techniques [18–22], including Variational Autoencoders (VAEs), show promise for mitigating these risks by identifying any deviation from patterns and issuing early warnings about price fluctuations [23].

The application of deep learning models has also been demonstrated in multiple studies to markedly enhance forecasting accuracy [24–28]. It is plausible that similar models could be adapted in Indonesia, where the agricultural sector confronts significant challenges in managing price volatility [29]. This capability becomes even more critical when forecasting not only predicts future prices but also detects anomalies early, enabling businesses to respond proactively to market disruptions. Deep learning-based methods have demonstrated superiority over traditional machine learning and statistical models in capturing complex market behaviors [30,31]. Their advanced pattern recognition capabilities make them particularly effective in identifying unusual price fluctuations in volatile environments. For instance, in the event of an unexpected drought affecting crop yields [32] or the imposition of sudden trade restrictions that alter supply chains [33–35], anomaly detection systems promptly alert businesses to adjust pricing strategies and manage inventory more effectively [36]. Such proactive measures improve MSMEs resilience and reduce revenue losses associated with unforeseen market disruptions [37].

Moreover, the integration of attention mechanisms has recently emerged as a breakthrough in predictive forecasting, offering a robust solution for capturing long-term dependencies in time series data [38]. These mechanisms enable models to selectively focus on the most relevant temporal patterns, enhancing accuracy in highly volatile markets. Notably, Transformer-based models such as TSLANet enhance the prediction process by enabling the algorithm to capture the most relevant data points, thereby enhancing forecast accuracy, particularly in complex and highly volatile agricultural commodity markets [39]. While deep learning-based forecasting and anomaly detection models have proven effective, further enhancements are needed to improve adaptability in volatile markets. The integration of early price anomaly detection with attention mechanisms offers a more powerful forecasting tool for MSMEs, enabling proactive responses to market volatility and greater operational resilience [40]. This study is motivated by the goal of enhancing the resilience of agricultural MSMEs to price fluctuations by delivering timely insights into potential market changes. It focuses on developing a framework capable of capturing long-term contextual price patterns and providing reliable anomaly detection. The novelty of this study lies in the development of an integrated deep learning framework that not only provides accurate multi-commodity price forecasting using Transformer-based models, but also incorporates anomaly detection through an attention-boosted Long Short-Term Memory VAE (LSTM-VAE). This dual functionality allows the system to capture both long-term price trends and sudden irregularities, offering predictive insight and early warnings simultaneously. The proposed framework introduces the following key contributions.

- **A unified architecture combining forecasting and anomaly detection:** designed for highly volatile agricultural markets.
- **Anomaly detection using the attention-boosted LSTM-VAE:** Identifies irregular price patterns, providing early warnings to mitigate risks and adapt strategies proactively.
- **Supporting agriculture-dependent MSMEs in volatile markets:** Strengthens resilience for businesses heavily reliant on agricultural commodities by improving price insights and reducing uncertainty. This enables better inventory management, cost control, and strategic decision-making, helping businesses sustain profitability despite market fluctuations.

2. Related work

This section provides a structured review of the existing literature related to agricultural price forecasting and anomaly detection. It begins with a discussion of models used for price forecasting, particularly in the context of agricultural price prediction, to determine the most suitable models for handling volatile prices. The review examines approaches ranging from statistical methods and machine learning to the latest deep learning models, including Transformer-based architectures. Following this, anomaly detection techniques are compared based on reported research findings. Finally, a comprehensive comparison is provided between our proposed model and existing frameworks, highlighting the contributions of our study and the specific research gap it aims to address.

2.1. Attention-based forecasting models

In developing economies like Indonesia, forecasting agricultural prices is particularly difficult due to an interplay of dynamic elements such as inconsistent climate patterns [41], inadequate transportation logistics, and fragmented market infrastructures. These elements often interact in unpredictable ways, making traditional forecasting methods insufficient in responding to abrupt market changes [42,43]. This challenge highlights the importance of adopting advanced models, particularly those based on attention-driven deep learning that are capable of capturing such complexities, which are more suited for handling volatile and irregular data patterns [44].

The incorporation of attention mechanisms has improved the performance of time series forecasting models by enabling them to focus on the most relevant parts of an input sequence when generating predictions. This capability is particularly beneficial in agricultural markets, where forecasting is complicated by a range of dynamic and interdependent factors such as seasonal variability, climate anomalies, and supply chain disruptions [45,46]. Recurrent Neural Network (RNN)-based deep learning models, including LSTM, have outperformed traditional statistical approaches such as Autoregressive Integrated Moving Average (ARIMA), particularly in capturing complex temporal dependencies and handling price volatility [47]. Effectively representing complex long-range dependencies and managing noisy data are critical concerns in the field of time series forecasting for agricultural prices [48]. Earlier studies [49,50] have highlighted the limitations of Bidirectional LSTM (BiLSTM) and Gated Recurrent Unit (GRU) models in capturing long-range dependencies in time series data, particularly in volatile agricultural markets.

Vaswani et al. [51] proposed a significant advancement in this field with the introduction of a self-attention mechanism, which enables the model to capture long-range sequence dependencies with greater efficacy. Although originally designed for natural language processing, the attention mechanism has been successfully adapted for time series forecasting due to its capacity to handle complex and noisy datasets [52]. In a recent study [53], Transformer models were employed in agricultural production prediction, where they demonstrated a 79.54 % improvement over traditional models. This serves as an illustration of the importance of temporal features, such as cyclical price patterns, seasonal variations, and exogenous market shocks, which can be more effectively captured by a Transformer model to improve the accuracy of agricultural commodity price forecasts.

Liu et al. [54] additionally demonstrated that Transformer-based models could also outperform RNN-based models and statistical approaches for rice yield prediction, reducing the mean absolute percentage error (MAPE) by 16.56 %. The present study thus sought to examine the efficacy of the Transformer model in processing noisy data, especially those prevalent in agricultural markets, where data may occasionally be incomplete or inconsistent. Furthermore, Transformer-based deep learning models, enhanced by attention mechanisms, are adept at processing multi-dimensional datasets, rendering them well-suited for predicting agricultural commodities that are influenced by numerous factors including climatic variables, soil conditions, and sensor-based measurements [55]. Additionally, Transformer models have demonstrated efficacy in domains related to agriculture, such as energy and finance, where volatility and exogenous factors play a significant role in determining outcomes, outperforming models based solely on Artificial Neural Network (ANN) or other RNN-based architectures that often struggle to capture long-range dependencies and sudden shifts in sequential data. Transformer-based deep learning models were selected in this study for their superior ability to process complex and sequential data influenced by multiple external variables. Unlike traditional statistical approaches, which often assume linearity and limited dependencies, Transformer models can capture long-term relationships and adapt to sudden fluctuations [56]. For example, in electricity demand forecasting, the use of a Transformer led to a 1.78 % reduction in the MAPE compared to statistical methods, highlighting its effectiveness in volatile environment [55]. This capability aligns well with the challenges found in Indonesia's remote regions, where agricultural price volatility is driven by unpredictable supply chains, weather variability, and infrastructure limitations. While deep learning requires larger datasets and higher computational resources, its advantages in modeling non-linear patterns and capturing temporal dependencies make it well-suited for addressing the research gaps identified [57].

2.2. Attention-boosted LSTM-VAE for anomaly detection

VAEs are effective for anomaly detection in volatile time-series data, particularly in agricultural markets where price fluctuations from climatic events and supply chain disruptions impact MSMEs [58]. Attention-boosted LSTM-VAEs enhance anomaly detection by identifying subtle deviations more effectively, providing early warnings to mitigate market instability [59]. VAEs are capable of detecting this type of anomaly because of their latent space encoding, which allows them to learn the normal patterns of the data and identify deviations from those patterns as possible anomalies [60].

The incorporation of an attention mechanism into VAEs allows for the detection of anomalies that are both subtle and complex. The incorporation of attention-based VAEs allows for a greater focus on the most critical portions of the input sequence, thereby enhancing the ability to detect anomalies. This quality is particularly beneficial in the context of a volatile market such as agriculture. Correia et al. [61] introduced a multi-head attention-based VAE with a P of 0.98 and an R of 0.67 VAE. This improvement is particularly relevant to the identification of sudden price changes that are triggered by factors such as unexpected weather shocks, which are common in commodity agricultural products.

Ma et al. [62] proposed an attention-based VAE for anomaly detection in energy markets, outperforming non-attention models. Their attention mechanism focused on time-series segments prone to anomalies, such as supply or demand shifts, enhancing performance in high-volatility environments. This approach could similarly improve anomaly detection in agricultural markets, where weather changes and policy shifts cause rapid price fluctuations [63]. An attention-based VAE could serve as an early warning system for anomalies, thus assisting MSMEs in operating within the Indonesian agricultural sector. The earlier anomalies are detected, the more effective decisions can be made regarding prices, inventories, and supply chain adjustments. The combination of attention-based VAEs with predictive models has the potential to empower MSMEs to navigate highly volatile markets. This is particularly crucial for Indonesian MSMEs, which operate on narrow profit margins and are often unable to withstand sudden price changes without significant financial repercussions.

2.3. Price anomaly detection agriculture frameworks

Recent studies have proposed various frameworks for forecasting and detecting anomalies in agricultural price data, employing a range of statistical and machine learning techniques. Table 1 provides a comparative overview of selected works, outlining the forecasting methods, anomaly detection techniques, evaluation metrics, detection time horizons, and key findings.

Table 1 presents a comparative analysis of recent studies focused on agricultural commodity price forecasting and anomaly detection. Although these studies share the objective of anticipating future price fluctuations, their methodologies and scopes differ substantially. Jiang et al. [66] applied a Convolutional Neural Network to classify visual representations of time series data, effectively identifying anomalies but without incorporating a forecasting mechanism. Khadka et al. [64] used a Bayesian classifier to detect trade-related anomalies influenced by economic and policy factors, yet their study also lacked a predictive modeling element. Madaan et al. [65] combined multivariate regression and LSTM networks for price forecasting and utilized Random Forest classifiers to detect anomalies within a fixed 43-day window. However, their approach relied on manual feature engineering and demonstrated limited generalizability. Mohanty et al. [12] developed a modular framework for predicting annual corn prices using various statistical and machine learning models, such as decision tree regressors and ARIMA, but did not address anomaly detection or short-term price dynamics.

Collectively, these studies reflect a separation between forecasting and anomaly detection tasks, with few attempts to integrate both within a single framework. Additionally, most of the reviewed works operate on aggregated data at monthly or yearly intervals, limiting their ability to detect short-term irregularities. The present study addresses these limitations by proposing an attention-based deep learning framework that performs long-term price forecasting and concurrently detects both point anomalies and contextual deviations. By operating on daily data and applying sequence-aware learning techniques, the proposed model achieves consistent detection performance across multiple agricultural commodities, providing a more adaptive and comprehensive approach to price monitoring and risk assessment in agricultural markets.

3. Proposed methods

Building upon the previous sections, this section presents the proposed methods for predicting and detecting anomalies in agricultural commodity prices, as informed by the literature review. The methodology includes several stages: data collection, preprocessing, price prediction, and anomaly detection, employing machine learning models such as Transformers and attention-based autoencoders. The overall workflow of the proposed approach is illustrated in Fig. 1, which outlines four main components, each described in detail within the figure.

The designed algorithm, shown in Fig. 2, combines a Transformer-based prediction model and an attention-boosted autoencoder for accurate price forecasting and anomaly detection in agricultural commodities. The input for both components is historical price data, including features such as commodity type, historical price, and time. The Transformer model uses its encoder to capture temporal patterns and its decoder to generate future price predictions ($\hat{y}_t$) for specified time steps ($T+1$, $T+2$, …). For anomaly detection, the attention-boosted autoencoder compresses the historical data into latent representations and reconstructs it to calculate reconstruction errors ($E_t =$

Table 1
Comparative summaries of recent studies on agricultural price forecasting and anomaly detection.

Author	Forecasting Method	Anomaly Detection Method	Evaluation Metric	Prediction Window	Key Findings
Khadka et al., 2023 [64]	Not provided	Naïve Bayes (NB) compared with other classifiers	Accuracy (91.89 %), residual analysis	Monthly (30 days)	Detected anomalies linked to economic and trade policy uncertainty; NB outperformed other classifiers.
Madaan et al., 2019 [65]	Multivariate regression with exogenous variables; LSTM	Random Forest classifier on engineered features and residuals	Root mean squared error (RMSE), normalized deviation, accuracy	Daily (30-day forecast and 43-day anomaly)	Integrated forecasting and anomaly detection to identify hoarding and weather-related disruptions.
Jiang et al., 2023 [66]	None (focused solely on anomaly detection)	ResNet-34 with Statistical Dependency Structure-Time Series Imaging (SDS-TSI)	Precision, Recall, F1-score (up to 0.93)	Biweekly (15-day window)	Transformed time series data into images using Markov Transition Fields; SDS-TSI-ResNet34 outperformed LSTM and other machine learning models in anomaly detection.
Mohanty et al., 2023 [12]	ARIMA; Theta (for yield); Decision Tree Regressor for price	None explicitly metioned	Mean absolute error (MAE), RMSE, MAPE	1 Year (aggregated)	Developed a modular machine learning pipeline integrating yield, supply, and demand for corn.
This Study	Deep Transformer	Attention-boosted VAE	MAPE, MAE	100 days	This study conducted long-term price forecasting using attention-based deep learning techniques. In addition to forecasting results, ours also detects anomalies that deviate from the general pattern using an attention-boosted VAE.

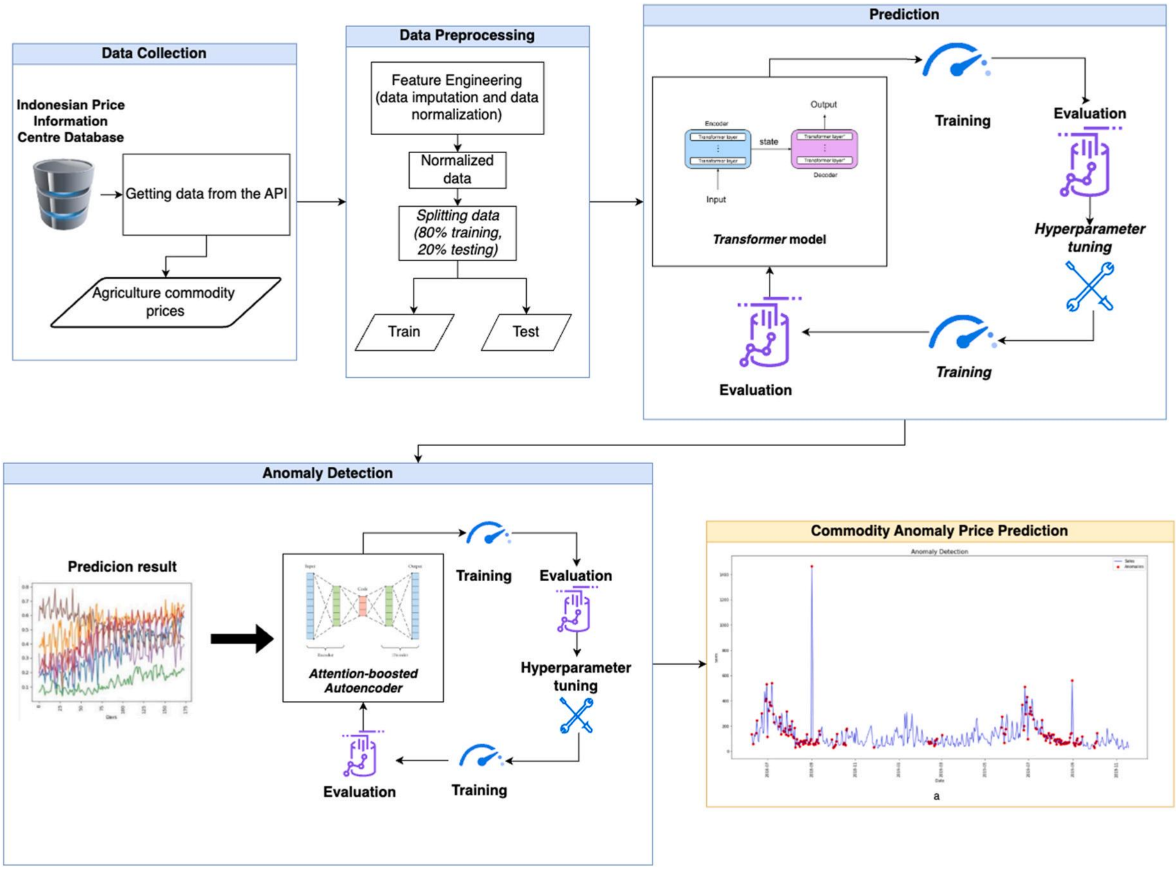
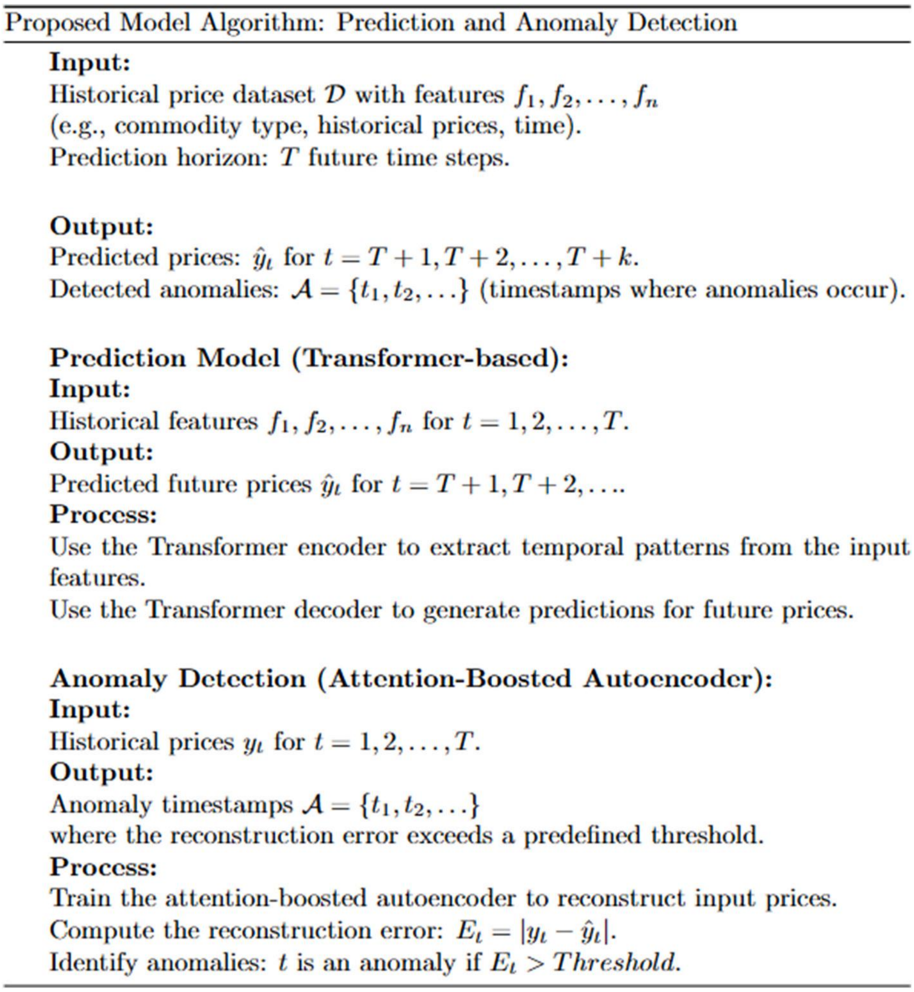


Fig. 1. The proposed framework for commodity price prediction and anomaly detection, integrating Transformer and attention-boosted LSTM autoencoder models.

$|y_t - \hat{y}_t|$). Time points with errors exceeding a defined threshold are flagged as anomalies ($\mathscr{A} = \{t_1, t_2, \ldots\}$), identifying unusual price behaviors. Together, this framework enables both long-term accurate price forecasting and early detection of abnormal price changes, providing actionable insights for stakeholders.

3.1. Data collection

The proposed methodology started with the acquisition of agricultural commodity price data from the Indonesian Price Information Centre Database via the official website (bi.go.id). While the database offers a broad range of commodities, this study specifically concentrated on agricultural commodities relevant to MSMEs, specifically rice, red chili, and shallots. This is in accordance with Indonesian Presidential Regulation No. 59 of 2020, which emphasizes strategic policies for stabilizing national food prices and supplies, particularly for key agricultural products [67].



```
Proposed Model Algorithm: Prediction and Anomaly Detection
    Input:
    Historical price dataset $\mathcal{D}$ with features $f_1, f_2, \ldots, f_n$
    (e.g., commodity type, historical prices, time).
    Prediction horizon: $T$ future time steps.

    Output:
    Predicted prices: $\hat{y}_t$ for $t = T+1, T+2, \ldots, T+k$.
    Detected anomalies: $\mathcal{A} = \{t_1, t_2, \ldots\}$ (timestamps where anomalies occur).

    Prediction Model (Transformer-based):
    Input:
    Historical features $f_1, f_2, \ldots, f_n$ for $t = 1, 2, \ldots, T$.
    Output:
    Predicted future prices $\hat{y}_t$ for $t = T+1, T+2, \ldots$.
    Process:
    Use the Transformer encoder to extract temporal patterns from the input features.
    Use the Transformer decoder to generate predictions for future prices.

    Anomaly Detection (Attention-Boosted Autoencoder):
    Input:
    Historical prices $y_t$ for $t = 1, 2, \ldots, T$.
    Output:
    Anomaly timestamps $\mathcal{A} = \{t_1, t_2, \ldots\}$
    where the reconstruction error exceeds a predefined threshold.
    Process:
    Train the attention-boosted autoencoder to reconstruct input prices.
    Compute the reconstruction error: $E_t = |y_t - \hat{y}_t|$.
    Identify anomalies: $t$ is an anomaly if $E_t > Threshold$.
```

Fig. 2. Proposed algorithm for commodity price prediction and anomaly detection.

3.2. Data preprocessing

In the data preprocessing phase, two key tasks are performed: normalization of the data and handling of missing values. These steps are crucial to facilitate convergence of the model, as the optimization algorithm can then make consistent updates across all features [68]. Normalization is the process of scaling the data such that it falls within a specific range. In this case, the data was normalized to a range of $[-1, 1]$. The equation for this normalization is given by Equation (1):

$$x' = 2 \cdot \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) - 1 \tag{1}$$

where.

- x' is the normalized value;
- x is the original value;
- $x_{\min}$ is the minimum value in the dataset;
- $x_{\max}$ is the maximum value in the dataset.

To address the issue of missing data, linear interpolation was applied. Linear interpolation entails estimating the value of missing data points by assuming that the change between two consecutive data points follows a linear pattern. The equation representing this linear interpolation is provided in Equation (2):

$$y = y_1 + \frac{(x - x_1)(y_2 - y_1)}{(x_2 - x_1)} \tag{2}$$

in which the following variables are defined.

- x is the point at which the data is missing;
- y is the interpolated value;
- x_1 and x_2 are the known data points surrounding the missing value;
- y_1 and y_2 are the corresponding known values at x_1 and x_2.

To prevent overfitting, the dataset was split into 80 % for training and 20 % for testing, allowing the model to learn from the majority of the data while reserving an unseen portion for final evaluation. To further enhance model robustness, 10-fold cross-validation was applied within the 80 % training dataset. This technique divides the training data into 10 subsets, where the model is iteratively trained on 9 subsets and validated on the remaining one. The process repeats until each subset has been used for validation, ensuring a stable and unbiased performance evaluation. Finally, the trained model was tested on the 20 % unseen test data to assess its generalization capability.

3.3. Prediction

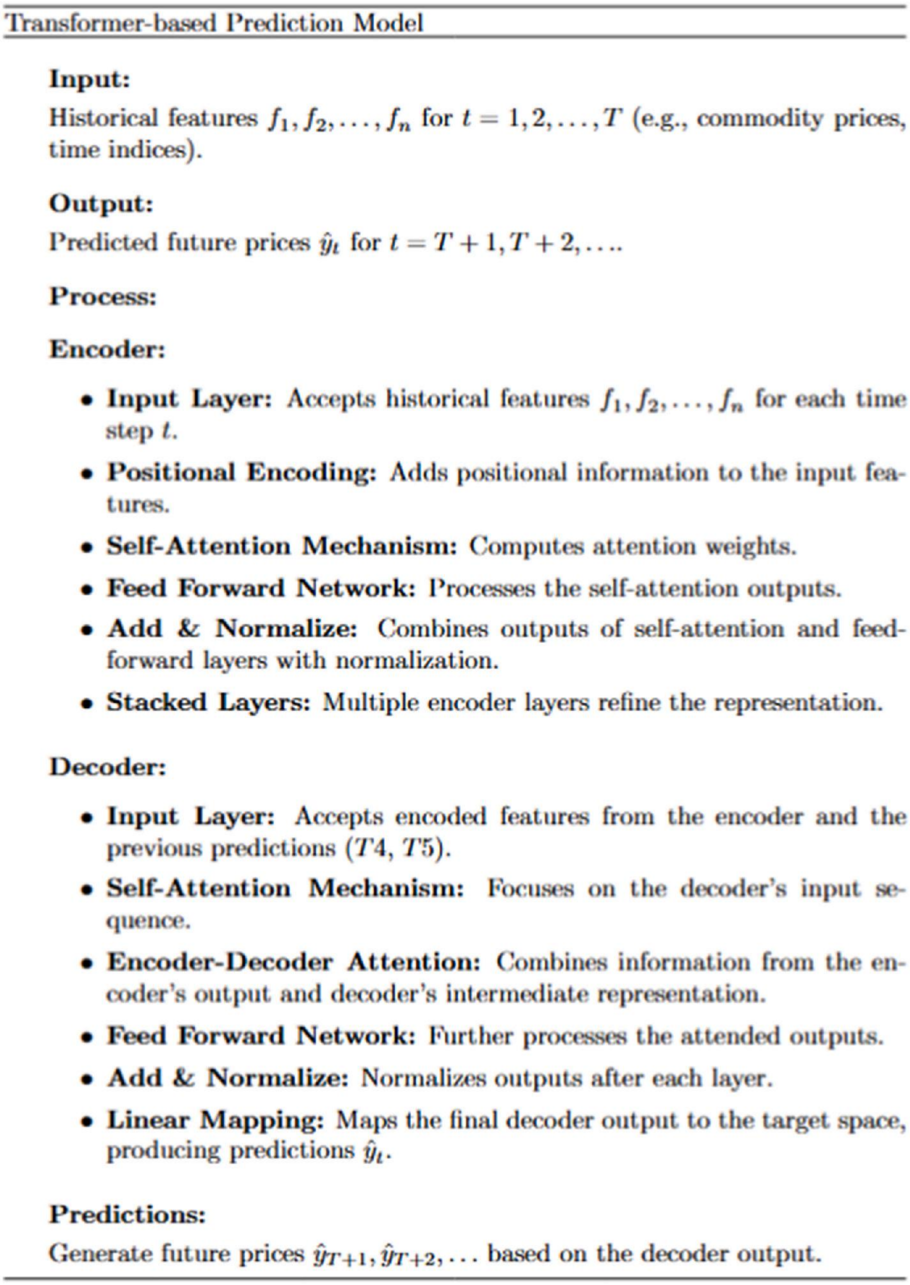
In the prediction phase of the proposed methodology, various deep learning models were evaluated for forecasting agricultural commodity prices. While traditional RNN-based models have been widely used, they often struggle to capture long-term dependencies and manage complex temporal patterns due to vanishing gradient issues [49,50]. To address these limitations, Transformer models were selected for their superior ability to retain long-range dependencies between input data points through an attention mechanism, enabling them to effectively process volatile time series data and improve predictive accuracy. This architecture, originally proposed by Vaswani et al. [51], relies on a self-attention mechanism that enables the model to weigh the relevance of distant elements within the input sequence when formulating predictions. In scenarios where the relationships between distant points in a sequence are of importance, such as in commodity price prediction, where historical data points over extended periods can exert influence, the attention mechanism is crucial.

Fig. 3 illustrates the Transformer-based prediction model algorithm formulated in a mathematically structured framework. It processes historical features $f_1, f_2, \ldots, f_n$ over time steps $t = 1, 2, \ldots, T$ to predict future values $\hat{y}_t$ for $t = T+1, T+2, \ldots$. The encoder transforms input features through positional encoding, introducing temporal dependencies, and refines them using feed-forward networks (FFNs) and normalization operations across multiple stacked layers. The decoder complements this process by incorporating the encoder's outputs and previous predictions, further refining the information with FFNs and normalization steps. Finally, the decoder output is linearly transformed to produce the forecasted values $\hat{y}_t$, enabling the model to generate predictions based on the temporal relationships captured from the input data. The self-attention operation is mathematically defined by Vaswani et al. [51] as presented in Equation (3):

$$\text{Self-attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

In this context, the following variables are defined.

- Q (query), K (key), and V (value) represent different transformations of the input sequence;
- d_k is the dimension of the key vectors.



Transformer-based Prediction Model

Input:

Historical features $f_1, f_2, \ldots, f_n$ for $t = 1, 2, \ldots, T$ (e.g., commodity prices, time indices).

Output:

Predicted future prices $\hat{y}_t$ for $t = T+1, T+2, \ldots$.

Process:

Encoder:

- **Input Layer:** Accepts historical features $f_1, f_2, \ldots, f_n$ for each time step t.
- **Positional Encoding:** Adds positional information to the input features.
- **Self-Attention Mechanism:** Computes attention weights.
- **Feed Forward Network:** Processes the self-attention outputs.
- **Add & Normalize:** Combines outputs of self-attention and feed-forward layers with normalization.
- **Stacked Layers:** Multiple encoder layers refine the representation.

Decoder:

- **Input Layer:** Accepts encoded features from the encoder and the previous predictions ($T4$, $T5$).
- **Self-Attention Mechanism:** Focuses on the decoder's input sequence.
- **Encoder-Decoder Attention:** Combines information from the encoder's output and decoder's intermediate representation.
- **Feed Forward Network:** Further processes the attended outputs.
- **Add & Normalize:** Normalizes outputs after each layer.
- **Linear Mapping:** Maps the final decoder output to the target space, producing predictions $\hat{y}_t$.

Predictions:

Generate future prices $\hat{y}_{T+1}, \hat{y}_{T+2}, \ldots$ based on the decoder output.

Fig. 3. The algorithm of Transformer-based model proposed by Wu et al. [69].

The prediction process in the Transformer model is tuned by several key hyperparameters that balance model complexity and predictive performance. The number of encoder and decoder layers is selected to capture important relationships in the data without over-complicating the model. The attention heads enable the model to focus on multiple aspects of the input sequence simultaneously, facilitating the model's understanding of the interrelationship across different time steps. The input dimension reflects the key features being used, such as price and volume, ensuring that the model processes all relevant information. The hidden layer sizes in the FFN and before and after the Transformer blocks are configured to provide the model with sufficient capacity to learn patterns in the data while maintaining efficiency. These hidden layers transform the input data as it traverses the network, ensuring that the model can make accurate predictions.

The model's performance is evaluated using two standard metrics: MAPE and MAE. These metrics are commonly employed for the assessment of the prediction accuracy in regression models. The MAPE calculates the average percentage error between the predicted values and the actual values, providing an indication of the prediction error relative to the magnitude of the actual values. In contrast, the MAE calculates the average absolute error between the predicted and actual values, giving a more direct measure of the model's prediction accuracy in the same unit as the data.

In Equation (4), the MAPE is defined as the average percentage error between the predicted values and the actual values, indicating the relative prediction error.

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{4}$$

In this equation, y_i represents the actual values, $\hat{y}_i$ represents the predicted values, and n is the total number of observations.

In Equation (5), the MAE is defined as the average absolute difference between the predicted and actual values, thereby providing a direct measure of the magnitude of errors in the same unit as the data.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

These metrics assist in quantifying the accuracy of the model's prediction. The MAPE focuses on relative errors, while the MAE provides the absolute error.

To ensure model generalization and mitigate overfitting, 10-fold cross-validation was applied during the training phase, where the dataset was iteratively split into training and validation subsets. Final evaluation was conducted on an independent 20 % test set. This approach ensures unbiased performance assessment across multiple partitions.

3.4. Proposed attention-boosted LSTM-VAE anomaly detection

This section presents an anomaly detection method for time series data that employs a combination of LSTM layers and an attention mechanism to enhance its ability to capture long-term dependencies. The objective is to identify deviations from typical patterns in the data, such as sudden spikes or drops in agricultural commodity prices. This model leverages the ability of LSTM to capture temporal dependencies, while the attention mechanism selectively highlights important time steps for further analysis.

The architecture of the proposed model is illustrated in Fig. 4. It consists of LSTM layers that function as both the encoder and decoder, with an attention mechanism applied in between. The LSTM encoder processes the input data, to extract sequential dependencies. The attention mechanism then generates a context vector by weighting the significance of each time step. This context vector is passed to the LSTM decoder, which reconstructs the time series. Finally, a time-distributed dense layer produces the predicted sequence.

The attention mechanism in this architecture enhances the anomaly detection model's ability to focus on time steps most relevant for identifying anomalies, allowing it to efficiently capture sequence deviations indicative of potential anomalies. By integrating the attention mechanism with LSTM layers, the model highlights critical sections of the time series while preserving the sequential dependencies of the entire input. The anomaly detection process relies on the reconstruction error, which represents the difference between predicted and actual values. As shown in Fig. 4, the model flags data points as anomalies when their reconstruction error exceeds a predefined threshold.

Fig. 5 presents the attention-boosted LSTM model algorithm, which processes a batch of input sequences by first passing them through a series of LSTM encoder layers to capture temporal dependencies. Subsequently, the encoder's output is then fed into an attention mechanism, which computes attention weights that emphasize the most relevant time steps. The attention weights are then employed to generate a context vector by taking a weighted sum of the encoder's outputs, accentuating important features within the sequence. The context vector is then repeated across all time steps and passed through LSTM decoder layers, which attempt to reconstruct the original sequence. A time-distributed dense layer then generates the final predicted sequence Y, which is compared to the original input X. The model calculates the reconstruction error, and anomalies are identified when this error exceeds a predefined threshold, signalling significant deviations from expected patterns. Anomalies were flagged when the reconstruction error exceeded a predefined threshold. The threshold was determined empirically by analyzing the distribution of reconstruction losses and optimizing the balance between precision and recall, ensuring minimal false positives. This approach enhances the model's ability to distinguish between normal price fluctuations and genuine anomalies, improving detection reliability. This anomaly detection model was trained using 10-fold cross-validation applied to 80 % of the dataset, ensuring robust

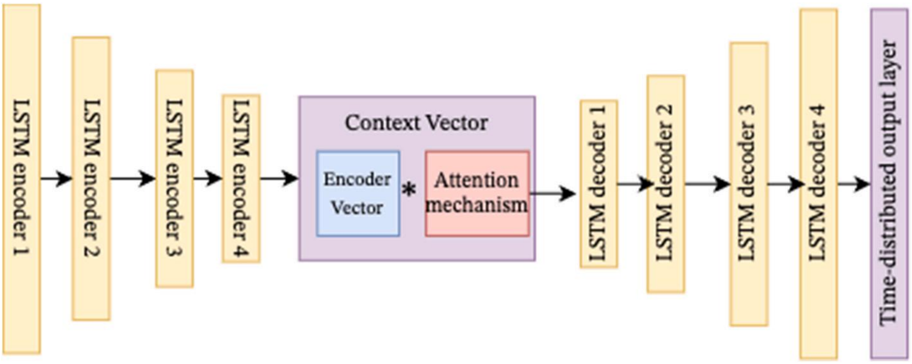
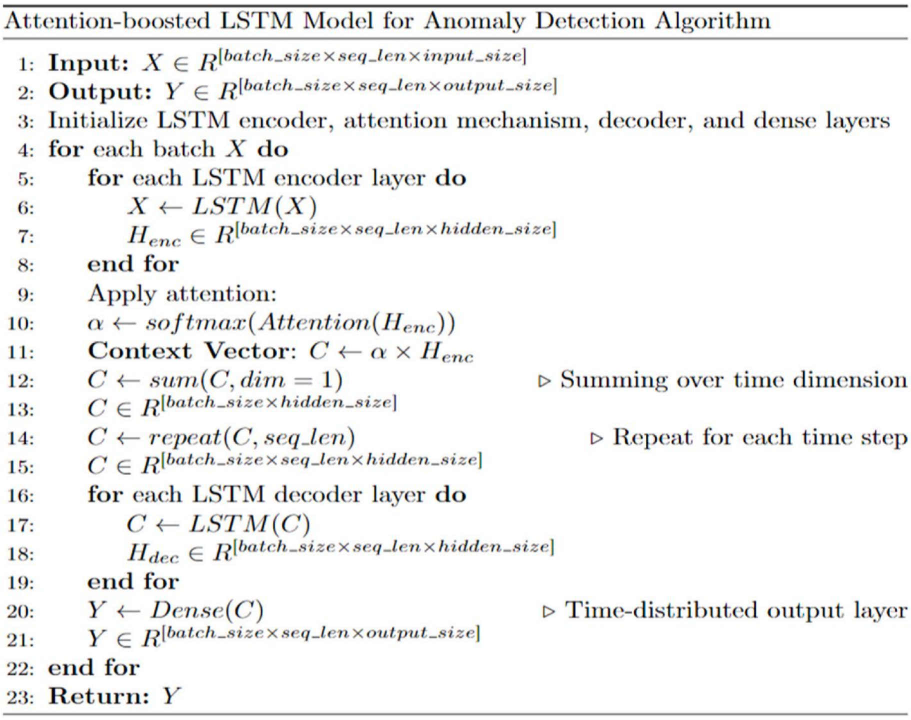


Fig. 4. Architecture of the proposed attention-boosted LSTM model.



```
Attention-boosted LSTM Model for Anomaly Detection Algorithm
Input: $X \in R^{[batch\_size \times seq\_len \times input\_size]}$
Output: $Y \in R^{[batch\_size \times seq\_len \times output\_size]}$
Initialize LSTM encoder, attention mechanism, decoder, and dense layers
for each batch $X$ do
    for each LSTM encoder layer do
        $X \leftarrow LSTM(X)$
        $H_{enc} \in R^{[batch\_size \times seq\_len \times hidden\_size]}$
    end for
    Apply attention:
    $\alpha \leftarrow softmax(Attention(H_{enc}))$
    Context Vector: $C \leftarrow \alpha \times H_{enc}$
    $C \leftarrow sum(C, dim = 1)$                ▷ Summing over time dimension
    $C \in R^{[batch\_size \times hidden\_size]}$
    $C \leftarrow repeat(C, seq\_len)$            ▷ Repeat for each time step
    $C \in R^{[batch\_size \times seq\_len \times hidden\_size]}$
    for each LSTM decoder layer do
        $C \leftarrow LSTM(C)$
        $H_{dec} \in R^{[batch\_size \times seq\_len \times hidden\_size]}$
    end for
    $Y \leftarrow Dense(C)$                       ▷ Time-distributed output layer
    $Y \in R^{[batch\_size \times seq\_len \times output\_size]}$
end for
Return: $Y$
```

Fig. 5. Proposed attention-boosted VAE anomaly detection algorithm.

validation across multiple partitions. The final model was then evaluated on the remaining 20 % unseen test data to assess its generalization performance. All computations were conducted using Python, with PyTorch for deep learning implementation and Scikit-learn for data preprocessing and evaluation.

4. Results and discussion

This section presents the results of the proposed framework for forecasting future prices and detecting anomalies in agricultural commodity markets using advanced deep learning techniques. The focus is on evaluating the effectiveness of the Transformer model for price prediction, followed by the use of an attention-boosted LSTM-VAE to detect anomalies in highly volatile commodities such as red chili, shallots, and rice. The results are organized into four main sub-sections: data collection, data preprocessing, prediction, and anomaly detection, each highlighting key findings and their implications. Additionally, the section compares the performance of the proposed models against traditional methods using key evaluation metrics to determine the advantages of attention-based approaches in handling complex temporal dependencies. The insights gained are expected to provide valuable support for MSMEs in managing price volatility, optimizing inventory, and enhancing strategic decision-making.

All experiments were conducted on a Windows system equipped with an Intel Core i7-10750H CPU operating at 2.60 GHz, 16 GB of RAM, and a GeForce GTX 1650 GPU with 4 GB of dedicated memory. The predictive and anomaly detection models were developed utilizing the Torch deep learning framework along with the PyTorch forecasting

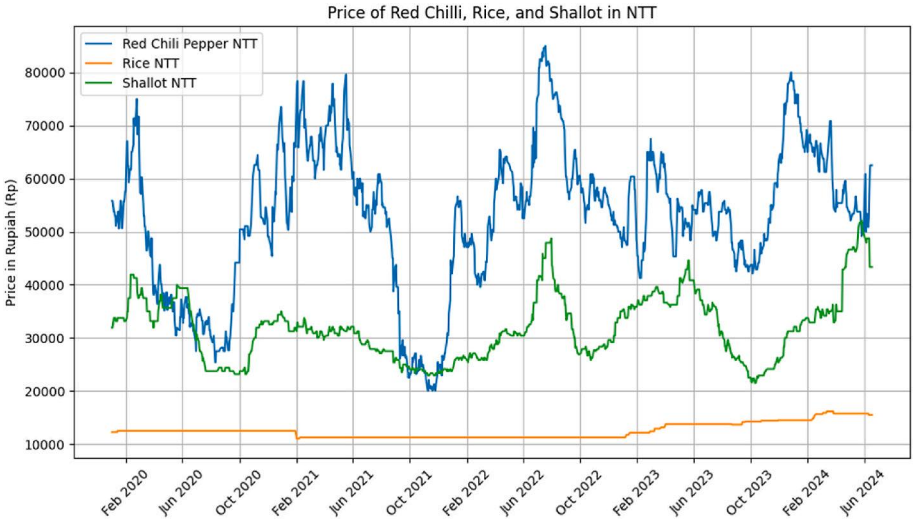


Fig. 6. Price trends of red chili, shallots, and rice in NTT (East Nusa Tenggara) from early 2020 to mid-2024. For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.

library.

4.1. Data collection

The data used in this study consisted of daily agricultural commodity prices for red chili, shallots, and rice, sourced from the Indonesian Price Information Centre Database. An example of the price trends can be seen in Fig. 6. The dataset encompasses a four-year period (2020–2024) and was recorded for five working days each week, with the exception of weekends and public holidays, during which no surveys were conducted. Consequently, the dataset exhibited a number of missing values, particularly during major national holidays. This disrupted the continuity of observed trends. To address this issue, linear interpolation was applied to estimate the missing data points, ensuring a more consistent temporal sequence. Linear interpolation was selected for its simplicity and effectiveness in maintaining trend consistency without introducing artificial volatility.

Table 2 reveals distinct patterns of price variability among the three agricultural commodities. Red chili pepper exhibits the highest level of price fluctuation, as indicated by its relatively high standard deviation, indicating greater instability in its pricing compared to shallot and rice. The median and mean values for rice being closely aligned suggest a more symmetric price distribution, whereas the greater variability in red chili pepper may indicate the presence of outliers or more frequent extreme price changes. These insights highlight underlying differences in the market behavior of these commodities.

4.2. Data preprocessing

To prepare the data for analysis, several preprocessing techniques were employed, as depicted in Fig. 7. First, all price values were normalized using a min-max scaling technique, which maps the data into a standardized range of -1 to 1, thus reducing the influence of extreme values that might distort the learning process. This effect is evident in Fig. 7, where rice prices in the NTT province initially appear stable, but fluctuations become apparent after normalization. Missing data points, which frequently occur due to reporting gaps or regional inconsistencies, were addressed using linear interpolation. This method was selected based on its capacity to maintain trend continuity without introducing artificial variance, while preserving the underlying original pattern. The preprocessed dataset was then split into training (80 %), validation (10 %), and testing (10 %) sets to ensure robust model evaluation. The 80 % training set was further utilized in a 10-fold cross-validation process, in which the data was iteratively partitioned into multiple training and validation subsets to enhance model generalization and prevent overfitting.

4.3. Data prediction

In this study, hyperparameter tuning and prediction were performed using 10-fold random state variations to identify the optimal configuration for forecasting the prices of red chili, shallot, and rice. The final results of this process are presented in Table 3. The volatility of prices for red chili and shallot was similar, which necessitated the use of identical hyperparameters. These included a model dimension (Dmodel) of 128, a single encoder and decoder layer, an FFN size of 30, and a time lag of 21. In contrast, rice prices exhibited minimal volatility, with only slight fluctuations at the end of the period in 2024. This necessitated a different configuration with a time lag of 14 and a larger FFN size of 100 to more effectively address its price trend. To accommodate these distinct temporal patterns observed in rice prices, the pre-layer and post-layer sizes were adjusted to 50 from 30. This was done to ensure alignment across the forecasting models while maintaining consistency with the other hyperparameters.

Table 2
Descriptive statistics of the NTT province’s three agricultural commodities.

		Red Chili Pepper	Shallot	Rice
(Price)	Mean	52,927.7	31,787.1	12,505
Rp/kg	Median	54,362.5	31,250	12,500
	Standard Deviation	13,730.4	6225.7	1419.8

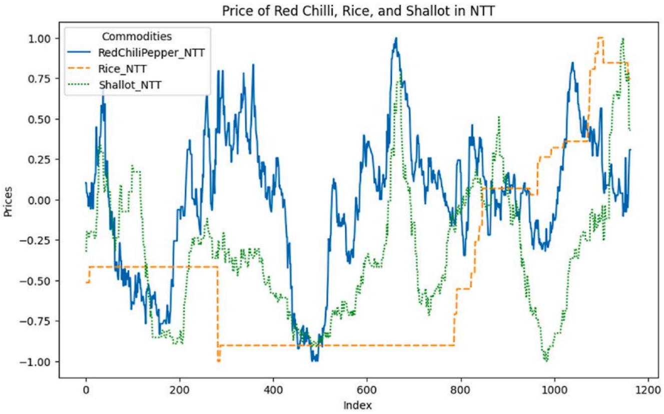


Fig. 7. Normalized price trends of red chili, shallots, and rice in NTT (East Nusa Tenggara). For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.

Table 3
Hyperparameters of the Deep Transformer for each dataset.

Hyperparameter	Commodity		
	Red Chili	Shallot	Rice
Dmodel	128	128	128
Encoder	1	1	1
Decoder	1	1	1
FFN	30	30	100
PRE_LAYER	30	30	50
POST_LAYER	30	30	50
TIMELAG	21	21	14

Fig. 8 illustrates the loss function trends for the Transformer model when applied three commodities including red chili, shallots, and rice over 100 training epochs. In the case of red chili, both the training and validation losses decreased rapidly within the first ten epochs and converged close to zero by epoch 20, indicating efficient learning with minimal divergence between the two curves. A comparable pattern was observed in shallot prices, with the validation loss initially reaching a value of approximately 0.04 and then declining sharply to nearly zero by epoch 40, while the training loss showed a steady decline, with minimal fluctuations. In contrast, the graph of the rice commodity shows a distinctive pattern, in which the training loss exhibited a significant decline in the initial epochs, while the validation loss experienced a surge at approximately epoch 10 and subsequently stabilized at a value that remained higher than the training loss. This indicates a disparity between the training and validation performance, suggesting that while the model demonstrated the capacity to fit the training data for all three commodities, its performance on unseen data was less consistent for rice than for red chili and shallots.

Fig. 9 illustrates the predicted and actual price trends for red chili, shallots, and rice using the Transformer model over a 100-timestep period. For red chili, the model accurately captured overall trends and multiple price fluctuations, with a prediction range closely aligned with the true values. This indicates strong performance in both high and low volatility conditions. Similarly, the model effectively followed the gradual increases and sharp spikes in shallot prices, maintaining a narrow prediction range that reflects high confidence and strong alignment with actual movements. In contrast, the prediction range for rice was

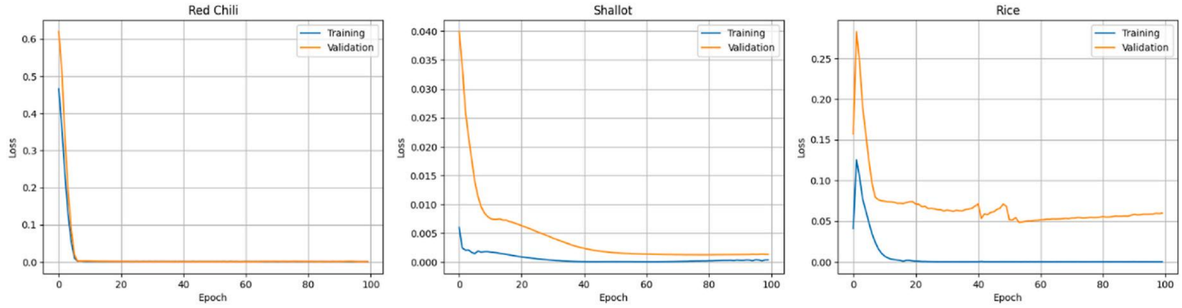


Fig. 8. Training and validation loss trends for the Transformer model applied to red chili, shallots, and rice over 100 epochs. For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.

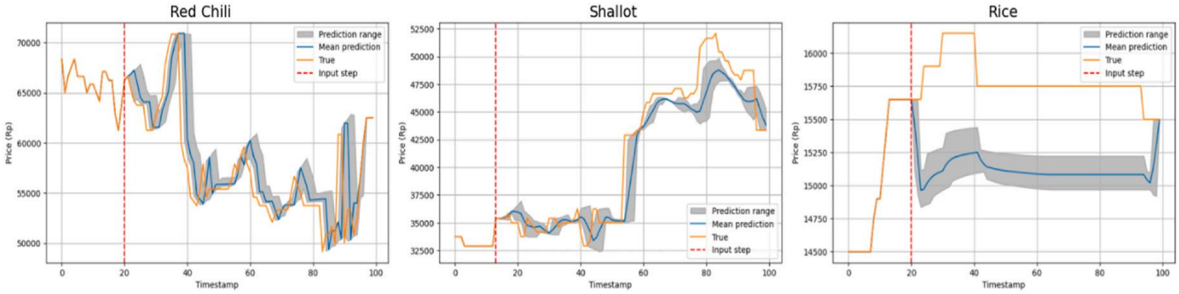


Fig. 9. Actual vs. predicted prices for red chili, shallots, and rice using the Transformer model. For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.

slightly broader, and the accuracy was comparatively lower. This can be partly explained by the relatively stable nature of the training data, recorded between February 2020 and December 2022, whereas the validation and test sets featured more abrupt changes. Despite these challenges, the model still managed to adapt to new patterns and identified key shifts in price trends during the evaluation phase. The test data specifically assessed the model's ability to predict previously unseen patterns, including significant price shifts between 2023 and mid-2024, when NTT experienced El Niño-induced droughts. These conditions delayed local market supply and coincided with increasingly strict rice import policies. The transition from stable training conditions to highly volatile test data reflects the real-world unpredictability observed in other commodities as well. Although the model initially struggled to detect price movement between timestamps 20 and 30, it successfully identified the upward trend between 30 and 50, followed by general price stability from 50 to 100. This contributed to the model's lower accuracy for rice relative to its more consistent performance with red chili and shallots.

The predictive performance of the Transformer model was further quantified using two primary metrics: MAPE and MAE. These results are shown in Table 4. The results clearly demonstrate that the Transformer model has superior performance compared to traditional models, including BiLSTM and GRU models, across all three commodities. With regard to red chili, the Transformer achieved a MAPE of 0.015921 and an MAE of 973, which were notably lower than those obtained by the BiLSTM (MAPE: 0.01966; MAE: 1129) and GRU (MAPE: 0.01861; MAE: 1078) models. This indicates that the Transformer is more adept at navigating the high volatility and sudden price shifts that are common in the red chili market. Similarly, with regard to shallots and rice, the Transformer demonstrated consistent superiority error reduction, showcasing its adaptability across diverse market conditions.

Table 4
Comparison of model performance for red chili, shallots, and rice using MAPE and MAE metrics across the Transformer, BiLSTM, and GRU models.

Commodity	Evaluation Metric	Model		
		Transformer	BiLSTM	GRU
Red Chili	MAPE	**0.015921**	0.01966	0.01861
	MAE	**973**	1129	1078
Shallot	MAPE	**0.01413**	0.02452	0.0385
	MAE	**520**	975	1033
Rice	MAPE	**0.07155**	0.071501	0.080266
	MAE	**1038**	1042	1163

4.4. Anomaly detection

The attention VAE adopts an architecture similar to the attention-boosted LSTM layer, relying on a hierarchical encoding and decoding approach to capture complex temporal dynamics in sequential data. The network is structured with multiple encoder and decoder layers, enabling the model to compress and reconstruct high-dimensional features while using attention mechanisms to prioritize relevant time steps.

Table 5 presents the layer-wise distribution of LSTM neurons in the attention-boosted LSTM layer. The number of neurons decreases progressively to enhance feature extraction and improve interpretability. The encoder's role is to summarize the input price sequences and pass them into the latent space of the attention-boosted VAE. In addition to relying on LSTM, the attention mechanism is applied to further capture contextual patterns. Instead of treating each timestamp equally, the attention mechanism assigns greater weights to time steps that are more relevant or informative, enabling the model to focus on significant changes or trends in the price sequence. This weighted combination of

Table 5
Layer-wise configuration of the attention-boosted LSTM network.

Layer	Number of LSTM neurons
Encoder 1/Decoder 4	512
Encoder 2/Decoder 3	256
Encoder 3/Decoder 2	128
Encoder 4/Decoder 1	64

hidden states produces a more meaningful latent representation before entering the VAE's probabilistic layer. The architecture gradually reduces the number of neurons from 512 to 64 across four layers to compress information effectively. The decoder then reconstructs the data from this latent space back to its original shape. The reconstructed values are compared to the predicted values, and if the difference exceeds a predefined threshold, the instance is identified as an anomaly.

Fig. 10 illustrates the distribution of MAE reconstruction losses generated by the attention-boosted VAE model during anomaly detection. The *x*-axis represents the MAE reconstruction loss values, while the *y*-axis indicates the frequency of each loss value across the test data. The distribution exhibited a right-skewed pattern, with the majority of data points exhibiting low reconstruction losses (below 0.5), indicating that the model accurately captures the normal price patterns for the majority of observations. A smaller number of data points exhibited reconstruction losses exceeding 1.5, indicating the potential presence of anomalies. These elevated MAE values are indicative of instances where the model has encountered difficulty in reconstructing the authentic patterns, likely due to significant deviations from expected behavior, such as sudden price spikes or declines. The smooth curve overlaying the histogram provided a clearer view of the loss distribution trend, which peaked at approximately 0.3. This further confirms that the majority of data points were well-represented by the model, with only a few points falling into the anomaly range. The results demonstrate the effectiveness of the attention-boosted VAE in detecting anomalous price patterns by identifying data points with unusually high reconstruction errors.

Table 6 presents a comparison of the anomaly detection performance of three models: the attention-boosted LSTM-VAE, standard LSTM-VAE, and ANN-VAE, using the MAE loss for red chili, shallots, and rice. The attention-boosted LSTM-VAE consistently achieved the lowest average MAE loss across all commodities, indicating superior anomaly detection capabilities. With regard to the commodity of red chili, the attention-boosted LSTM-VAE achieved an MAE of 0.7714989, which is lower than the 0.812345 recorded by both the standard LSTM-VAE and ANN-VAE. Similarly, for shallots, the attention-boosted LSTM-VAE significantly outperformed the other models, with an MAE of 0.79628307, in contrast to 0.9327428 (LSTM-VAE) and 1.2243275 (ANN-VAE). In the case of rice, the attention-boosted LSTM-VAE also achieved the lowest MAE at 0.86234256, exhibiting a slight superiority over the LSTM-VAE (0.8678327) and a notable improvement over the ANN-VAE (0.9362434). These results illustrate the enhanced performance of the attention-boosted LSTM-VAE, which benefits from its ability to focus on critical time points through the attention mechanism, thereby enhancing its sensitivity to anomalies in comparison to its counterparts.

Fig. 11 displays the anomaly detection results for red chili, shallots, and rice using the attention-boosted LSTM-VAE model. In each plot, the blue line represents the true price values over time, while the red dots indicate the anomalies that have been identified. In the case of red chili, multiple anomalies were identified during periods of sharp price drops and recoveries, particularly between timestamps 20–50. This suggests that the model is effective in capturing abrupt changes caused by market fluctuations. With regard to shallots, the model identified anomalies predominantly during periods of accelerated price appreciation and subsequent stabilization, such as between timestamps 50–100, where sudden price surges are marked as deviations. This indicates that the model is responsive to unanticipated upward trends, which could potentially signal market disruptions. In the case of rice, a considerable number of anomalies were detected in the flatter segments between timestamps 30–70, where the prices remained relatively stable. This suggests that the model identifies the absence of expected variability as an anomaly, highlighting the challenge of distinguishing between normal price stability and anomalous stasis in low-volatility commodities. Overall, the results confirm that the attention-boosted LSTM-VAE is effective in identifying unusual patterns in high-volatility commodities like red chili and shallots but may generate false positives during periods of relative stability, as observed in the "Rice" plot.

Table 6
Comparison of the average MAE loss for anomaly detection models in the context of Indonesian commodities.

Commodity	Model's Average MAE Loss		
	Proposed attention-boosted LSTM-VAE	BiLSTM-VAE	ANN-VAE
Red Chili	**0.7714989**	0.812345	0.812345
Shallot	**0.79628307**	0.9327428	1.2243275
Rice	**0.86234256**	0.8678327	0.9362434

The proposed model has demonstrated its capability not only in forecasting future prices but also to identify sudden price changes up to 100 days in advance. These advancements could be valuable for policymakers in mitigating price volatility and developing timely interventions. For example, if the model predicts a price increase during dry periods, which may result from drought conditions, a potential response could involve enhancing irrigation infrastructure to stabilize supply. In addition, MSMEs could use the model's forecasts to anticipate market trends and manage their inventory more effectively. This would enable them to adjust their inventory levels in anticipation of expected price increases or decreases, thereby reducing the negative impact of market fluctuations.

5. Conclusion

This study proposed a comprehensive deep learning framework combining Transformer models for price forecasting and the attention-

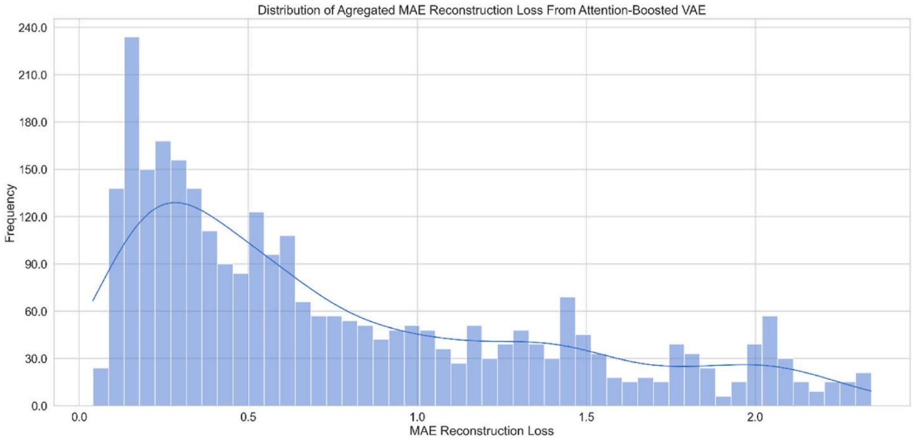


Fig. 10. Distribution of aggregated MAE reconstruction loss from the attention-boosted VAE model.

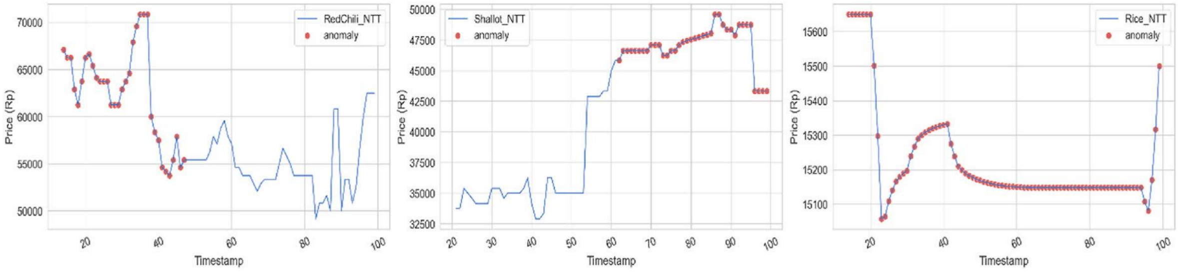


Fig. 11. Anomaly detection results for Indonesian commodities using the attention-boosted LSTM-VAE.

boosted LSTM-VAE for anomaly detection, designed for the volatile agricultural markets of Indonesia. The results demonstrated that the Transformer model significantly outperformed other models such as the BiLSTM and GRU in forecasting accuracy, as indicated by lower MAPE and MAE scores for red chili, shallots, and rice. This highlights the Transformer's superior ability to capture complex temporal dependencies and handle both high and low market volatility. The anomaly detection component, leveraging the attention-boosted LSTM-VAE, proved effective in identifying unexpected price spikes and drops, offering early warnings to mitigate risks and support proactive decision-making by stakeholders.

The findings highlight the transformative potential of attention-based mechanisms in improving forecasting accuracy and enhancing the resilience of MSMEs that depend on agricultural commodities. The proposed framework provides practical value by equipping MSMEs with a strong tool to navigate market uncertainties, stabilize operations, and optimize inventory and pricing strategies. However, additional factors that may significantly impact agricultural production should also be taken into account, including annual rainfall, sunlight duration, and other relevant geospatial variables. Currently, these features are not integrated into the model.

Future research should aim to incorporate sparsity to improve model efficiency and reduce computational complexity without sacrificing accuracy. Moreover, including multiple relevant inputs, such as temperature, rainfall, soil moisture, input costs, and market prices, could further improve model performance. In real-world applications, developing anomaly detection systems may enable quicker responses to market disruptions. Evaluating the performance of the proposed framework across various sectors or regions would also be valuable to determine its broader applicability (e.g., stock market prediction [31, 70]). Taken together, these steps would strengthen the role of machine learning-based tools in addressing price volatility and market uncertainty.

CRediT authorship contribution statement

Eko Sediyono: Writing – review & editing, Validation, Supervision, Investigation, Formal analysis, Conceptualization. **Kristoko Dwi Hartomo:** Visualization, Validation, Methodology, Investigation, Formal analysis. **Christian Arthur:** Writing – original draft, Visualization, Investigation, Formal analysis, Data curation. **Intiyas Utami:** Validation, Investigation, Formal analysis, Data curation, Conceptualization. **Ronny Prabowo:** Validation, Investigation, Formal analysis, Conceptualization. **Raymond Chiong:** Writing – review & editing, Supervision, Methodology, Investigation.

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.

Acknowledgment

The authors would like to acknowledge financial support from the Directorate of Research, Technology, and Community Service, Directorate General of Higher Education, Research, and Technology, and Ministry of Education, Culture, Research, and Technology, Indonesia (Grant No. 108/E5/PG.02.00.PL/2024).

Data availability

Data will be made available on request.

References

[1] S. Yang, Y. Fu, Interconnectedness among supply chain disruptions, energy crisis, and oil market volatility on economic resilience, Energy Econ. 143 (2025) 108290, https://doi.org/10.1016/j.eneco.2025.108290.

[2] R. Karkowska, S. Urjasz, Importance of geopolitical risk in volatility structure: new evidence from biofuels, crude oil, and grains commodity markets, Journal of Commodity Markets 36 (2024) 100440, https://doi.org/10.1016/j.jcomm.2024.100440.

[3] A.A. Patunru, Trade policy in Indonesia: between ambivalence, pragmatism and nationalism, Bull Indones Econ Stud 59 (3) (2023) 311–340, https://doi.org/10.1080/00074918.2023.2282821.

[4] E. Valerio, N. Hilmiati, J. Prior, D. Dahlanuddin, Analysis of the agricultural innovation system in Indonesia: a case study of the beef sector in Nusa Tenggara Barat, Agric. Syst. 203 (2022) 103529, https://doi.org/10.1016/j.agsy.2022.103529.

[5] W. Hermawan Fitrawaty, M. Yusuf, I. Maipita, A simulation of increasing rice price toward the disparity of income distribution: an evidence from Indonesia, Heliyon 9 (3) (Mar. 2023) e13785, https://doi.org/10.1016/j.heliyon.2023.e13785.

[6] Y.N. Muflikh, C. Smith, C. Brown, A.A. Aziz, Analysing price volatility in agricultural value chains using systems thinking: a case study of the Indonesian chilli value chain, Agric. Syst. 192 (2021) 103179, https://doi.org/10.1016/j.agsy.2021.103179.

[7] N. Nurwaesari, P. Maulidiyah, D.M. Fitri, Inequality impact of MSME assistance in Indonesia on urban migration in the post pandemic era, in: Proceedings of the 2nd International Indonesia Conference on Interdisciplinary Studies (IICIS 2021), Atlantis Press, 2021, pp. 67–72, https://doi.org/10.2991/assehr.k.211206.010.

[8] Y. Pujowati, P.A.S. Hasibuan, S.T. Budiono, Analysis of the impact of COVID-19 on MSME income (micro, small and medium enterprises) in Nganjuk regency, Jurnal Pamator : Jurnal Ilmiah Universitas Trunojoyo 15 (2022) 100–112, https://doi.org/10.21107/pamator.v15i1.13922.

[9] R. Hartono, Moh S. Rahman, D. Retnoningsih, M.I. Shaleh, Price volatility of horticulture commodity during the pandemic in East Java, Indonesia, HABITAT 34 (3) (Nov. 2023) 277–288, https://doi.org/10.21776/ub.habitat.2023.034.3.25.

[10] K.D. Hartomo, C. Arthur, Enhanced MSME support allocation with integrated K-means and Tukey's outlier detection, in: K. Arai (Ed.), Intelligent Systems and Applications, Springer Nature Switzerland, Cham, 2024, pp. 241–257.

[11] M. P. Halim, N. Yudistira, C. Dewi, Multicommodity prices prediction using multivariate data-driven modeling: Indonesia case, IEEE Trans. Comput. Soc. Syst. 11 (6) (Dec. 2024) 8106–8117, https://doi.org/10.1109/tcss.2022.3229772.

[12] M. K. Mohanty, P. K. G. Thakurta, S. Kar, Agricultural commodity price prediction model: a machine learning framework, Neural Comput. Appl. 35 (20) (Apr. 2023) 15109–15128, doi: 10.1007/s00521-023-08528-7.

[13] A. Morchid, et al., IoT-enabled smart agriculture for improving water management: a smart irrigation control using embedded systems and Server-Sent Events, Sci. Afr. 27 (Mar. 2025) e02527, doi: 10.1016/j.sciaf.2024.e02527.

[14] A. Morchid, et al., Fire detection and anti-fire system to enhance food security: a concept of smart agriculture systems-based IoT and embedded systems with machine-to-machine protocol, Sci. Afr. 27 (Mar. 2025) e02559, https://doi.org/10.1016/j.sciaf.2025.e02559.

[15] L. Madaan, A. Sharma, P. Khandelwal, S. Goel, P. Singla, A. Seth, Price forecasting & anomaly detection for agricultural commodities in India, in: Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies, Association

for Computing Machinery, New York, NY, USA, 2019, pp. 52–64, https://doi.org/10.1145/3314344.3332488, in COMPASS '19.

[16] C.O. Hurley, MSME competitiveness in small island economies: a comparative systematic review of the literature from the past 24 years, Enterpren. Reg. Dev. 30 (9–10) (Oct. 2018) 1027–1068, https://doi.org/10.1080/08985626.2018.1515822.

[17] M.A. Ghiffari, Analysis of raw material inventory control using the economic order quantity (EOQ) method: study of simping MSMEs, Purwakarta regency, Indonesia, Arkus 10 (1) (Feb. 2024) 520–524, https://doi.org/10.37275/arkus.v10i1.524.

[18] Y. Nataliani, C. Arthur, T. Wellem, K. D. Hartomo, N. H. A. Wahab, Multi-objective k-nearest neighbor for breast cancer detection, JOIV : Int. J. Inform. Visualization 9 (1) (Jan. 2025) 241, https://doi.org/10.62527/joiv.9.1.2669.

[19] A. Morchid, R. Jebabra, H. Qjidaa, R. El Alami, M.O. Jamil, Agri-tech innovations for sustainability: a fire detection system based on MQTT broker and IoT to improve environmental risk management, Results Eng. 24 (Dec. 2024) 103683, https://doi.org/10.1016/j.rineng.2024.103683.

[20] A. Morchid, Z. Oughannou, R. El Alami, H. Qjidaa, M.O. Jamil, H.M. Khalid, Integrated internet of things (IoT) solutions for early fire detection in smart agriculture, Results Eng. 24 (Dec. 2024) 103392, https://doi.org/10.1016/j.rineng.2024.103392.

[21] A. Morchid, et al., IoT-enabled fire detection for sustainable agriculture: a real-time system using flask and embedded technologies, Results Eng. 23 (Sep. 2024) 102705, https://doi.org/10.1016/j.rineng.2024.102705.

[22] A. Morchid, R. Jebabra, H.M. Khalid, R. El Alami, H. Qjidaa, M. Ouazzani Jamil, IoT-based smart irrigation management system to enhance agricultural water security using embedded systems, telemetry data, and cloud computing, Results Eng. 23 (Sep. 2024) 102829, https://doi.org/10.1016/j.rineng.2024.102829.

[23] S. Park, S. Jung, E. Hwang, S. Rho, Variational autoencoder-based anomaly detection scheme for load forecasting, in: H.R. Arabnia, K. Ferens, D. de la Fuente, E.B. Kozerenko, J.A. Olivas Varela, F.G. Tinetti (Eds.), Advances in Artificial Intelligence and Applied Cognitive Computing, Springer International Publishing, Cham, 2021, pp. 833–839, https://doi.org/10.1007/978-3-030-70296-0_62.

[24] K.D. Hartomo, Y. Nataliani, A new model for learning-based forecasting procedure by combining k-means clustering and time series forecasting algorithms, PeerJ Comput. Sci. 7 (2021) 1–29, https://doi.org/10.7717/PEERJ-CS.534.

[25] X. Fu, et al., Simulation and forecasting of fishery weather based on statistical machine learning, Inf. Process. Agric. 11 (1) (Mar. 2024) 127–142, https://doi.org/10.1016/j.inpa.2023.05.001.

[26] C. Deina, et al., A methodology for coffee price forecasting based on extreme learning machines, Information Processing in Agriculture 9 (4) (Dec. 2022) 556–565, https://doi.org/10.1016/j.inpa.2021.07.003.

[27] M.A. Nanda, et al., Higuchi fractal dimension and deep learning on near-infrared spectroscopy for determination of free fatty acid (FFA) content in oil palm fruit, J. Agric. Food Res. 18 (2024) 101437, https://doi.org/10.1016/j.jafr.2024.101437.

[28] K. Dwi Hartomo, C. Arthur, Y. Nataliani, A novel weighted loss TabTransformer integrating explainable AI for imbalanced credit risk datasets, IEEE Access 13 (2025) 31045–31056, https://doi.org/10.1109/access.2025.3541878.

[29] A.H. Utomo, M.A. Gumilang, A. Ahmad, Agricultural commodity sales recommendation system for farmers based on geographic information systems and price forecasting using probabilistic neural network algorithm, in: IOP Conference Series: Earth and Environmental Science, IOP Publishing Ltd, Feb. 2022, https://doi.org/10.1088/1755-1315/980/1/012061.

[30] Z. Niu, G. Zhong, H. Yu, A review on the attention mechanism of deep learning, Neurocomputing 452 (Sep. 2021) 48–62, https://doi.org/10.1016/j.neucom.2021.03.091.

[31] R. Chiong, Z. Fan, Z. Hu, S. Dhakal, A novel ensemble learning approach for stock market prediction based on sentiment analysis and the sliding window method, IEEE Trans Comput Soc Syst 10 (5) (Oct. 2023) 2613–2623, doi:10.1109/TCSS.2022.3182375.

[32] M. Devi, J. Kumar, D.P. Malik, P. Mishra, Forecasting of wheat production in Haryana using hybrid time series model, J. Agric. Food Res. 5 (2021) 100175, https://doi.org/10.1016/j.jafr.2021.100175.

[33] B. Xu, L. Chai, C. Zhang, Research and application on corn crop identification and positioning method based on Machine vision, Information Processing in Agriculture 10 (1) (Mar. 2023) 106–113, https://doi.org/10.1016/j.inpa.2021.07.004.

[34] K. Hameed, D. Chai, A. Rassau, Texture-based latent space disentanglement for enhancement of a training dataset for ANN-based classification of fruit and vegetables, Information Processing in Agriculture 10 (1) (Mar. 2023) 85–105, https://doi.org/10.1016/j.inpa.2021.09.003.

[35] N. Harshith, P. Kumari, Memory based neural network for cumin price forecasting in Gujarat, India, J. Agric. Food Res. 15 (Mar. 2024) 101020, doi: 10.1016/j.jafr.2024.101020.

[36] K. Kethineni, P. Gera, Iot-based privacy-preserving anomaly detection model for smart agriculture, Systems 11 (6) (Jun. 2023) 304, https://doi.org/10.3390/systems11060304.

[37] F. Zhang, J. Upton, L. Shalloo, P. Shine, M.D. Murphy, Effect of introducing weather parameters on the accuracy of milk production forecast models, Information Processing in Agriculture 7 (1) (Mar. 2020) 120–138, https://doi.org/10.1016/j.inpa.2019.04.004.

[38] H. Yin, D. Jin, Y.H. Gu, C.J. Park, S.K. Han, S.J. Yoo, STL-ATTLSTM: vegetable price forecasting using stl and attention mechanism-based LSTM, Agriculture (Switzerland) 10 (12) (Dec. 2020) 1–17, https://doi.org/10.3390/agriculture10120612.

[39] E. Eldele, M. Ragab, Z. Chen, M. Wu, X. Li, TSLANet: rethinking transformers for time series representation learning [Online]. Available: http://arxiv.org/abs/2404.08472, Apr. 2024.

[40] W. Cheng, T. Ma, X. Wang, G. Wang, Anomaly detection for internet of things time series data using generative adversarial networks with attention mechanism in smart agriculture, Front. Plant Sci. 13 (Jun) (2022), 890563, doi:10.3389/fpls.2022.890563.

[41] A. Wijaya, P. Glasbergen, P. Leroy, A. Darmastuti, Governance challenges of cocoa partnership projects in Indonesia: seeking synergy in multi-stakeholder arrangements for sustainable agriculture, Environ. Dev. Sustain. 20 (1) (Feb. 2018) 129–153, https://doi.org/10.1007/s10668-016-9874-8.

[42] S. Soetriono, D. Soejono, E.S. Hani, A. Suwandari, B.S. Narmaditya, Challenges and opportunities for agribusiness development: lesson from Indonesia, The Journal of Asian Finance, Economics and Business 7 (9) (Sep. 2020) 791–800, https://doi.org/10.13106/jafeb.2020.vol7.no9.791.

[43] N. Ngadi, et al., Challenge of agriculture development in Indonesia: rural youth mobility and aging workers in agriculture sector, Sustainability 15 (2) (Jan. 2023) 922, https://doi.org/10.3390/su15020922.

[44] W. Witanti, S. Arie Anggara, M. Melina, Red chili price forecasting using attention mechanism based on long short-term memory, Journal of Applied Computer Science and Technology 5 (2) (Dec. 2024) 128–135, https://doi.org/10.52158/jacost.v5i2.875.

[45] Y. Jin, W. Li, J.M. Gil, Forecasting fish prices with an artificial neural network model during the tuna fraud, J. Agric. Food Res. 18 (2024) 101340, https://doi.org/10.1016/j.jafr.2024.101340.

[46] S. Yasmin, Md Moniruzzaman, Forecasting of area, production, and yield of jute in Bangladesh using Box-Jenkins ARIMA model, J. Agric. Food Res. 16 (2024) 101203, https://doi.org/10.1016/j.jafr.2024.101203.

[47] E. Sediyono, S. Wahyuni, I. Sembiring, Optimizing the long short-term memory algorithm to improve the accuracy of infectious diseases prediction, IAES Int. J. Artif. Intell. 13 (Sep. 2024) 2893, https://doi.org/10.11591/ijai.v13.i3.pp2893-2903.

[48] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.

[49] K. Ullah, et al., Short-term load forecasting: a comprehensive review and simulation study with CNN-LSTM hybrids approach, IEEE Access 12 (2024) 111858–111881, https://doi.org/10.1109/ACCESS.2024.3440631.

[50] M. Waqas, U.W. Humphries, A critical review of RNN and LSTM variants in hydrological time series predictions, MethodsX 13 (Dec. 2024) 102946, https://doi.org/10.1016/j.mex.2024.102946.

[51] A. Vaswani, et al., Attention is all you need, in: I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems, Curran Associates, Inc., 2017 [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[52] C. Arthur, N. Yudistira, C. Dewi, AutoCyclic: deep learning optimizer for time series data prediction, IEEE Access 12 (2024) 14014–14026, https://doi.org/10.1109/ACCESS.2024.3356553.

[53] S.C. Ibañez, C.P. Monterola, A global forecasting approach to large-scale crop production prediction with time series transformers, Agriculture (Switzerland) 13 (9) (Sep. 2023), 1855, doi:10.3390/agriculture13091855.

[54] Y. Liu, et al., Rice yield prediction and model interpretation based on satellite and climatic indicators using a transformer method, Remote Sens (Basel) 14 (19) (Oct. 2022), 5045, doi:10.3390/rs14195045.

[55] A. L'heureux, K. Grolinger, M.A.M. Capretz, Transformer-based model for electrical load forecasting, Energies 15 (14) (Jul. 2022), 4993, doi:10.3390/en15144993.

[56] K. Olorunnimbe, H. Viktor, Ensemble of temporal Transformers for financial time series, J. Intell. Inf. Syst. (2024) 1087–1111, https://doi.org/10.1007/s10844-024-00851-2.

[57] P. Nareswari, S.S. Wibowo, "Global and local commodity prices: a further look at the Indonesian agricultural commodities the creative commons attribution, CC BY NC 4.0, https://www.mfa.com.my/cmr/v28_i1_a4/, 2020.

[58] Y. Pujowati, P.A.S. Hasibuan, S.T. Budiono, Analysis of the impact of covid-19 on the income of MSMEs (micro, small and medium enterprises) in nganjuk regency, Jurnal Pamator : Jurnal Ilmiah Universitas Trunojoyo 15 (1) (Apr. 2022) 100–112, https://doi.org/10.21107/pamator.v15i1.13922.

[59] R. Lopez, J. Regier, M. I. Jordan, and N. Yosef, "Information constraints on auto-encoding variational bayes," in Advances in Neural Information Processing Systems, vol. 31, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds. Curran Associates, Inc., 2018. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2018/file/9a96a2c73c0d477ff2a6da3bf538f4f4-Paper.pdf.

[60] S.A. Hosseini, S.A. Hosseini, M. Houshmand, Variational autoencoder-based dimension reduction of Ichimoku features for improved financial market analysis, Franklin Open 8 (Sep. 2024) 100135, https://doi.org/10.1016/j.fraope.2024.100135.

[61] L. Correia, J.-C. Goos, P. Klein, T. Bäck, A.V. Kononova, MA-VAE: multi-head attention-based variational autoencoder approach for anomaly detection in multivariate time-series applied to automotive endurance powertrain testing, Proceedings of the 15th International Joint Conference on Computational Intelligence, Sep. 2023, pp. 407–418, doi:10.5220/0012163100003595.

[62] C. Ma, Y. Wang, F. Li, H. Zhang, Y. Zhang, H. Zhang, Constructing attention-LSTM-VAE power load model based on multiple features, Advances in Mathematical Physics 2024 (1) (Jan. 2024), 041791, doi:10.1155/2024/1041791.

[63] M. Ioanes, A Hybrid Approach to Scenario Generation for Stress Testing Market Fluctuations: Blending VAE and GAN Models to Generate Volatility Surfaces and

Stock Returns [Dissertation]. 2024. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-226225.

[64] S. Khadka, M. Gopinath, F.A. Batarseh, Anomalies in agricultural trade: a Bayesian classifier approach, Journal of the Agricultural and Applied Economics Association 2 (3) (Sep. 2023) 402–427, https://doi.org/10.1002/jaa2.69.

[65] L. Madaan, A. Sharma, P. Khandelwal, S. Goel, P. Singla, A. Seth, Price forecasting and anomaly detection for agricultural commodities in India, in: Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies, ACM, New York, NY, USA, Jul. 2019, pp. 52–64, https://doi.org/10.1145/3314344.3332488.

[66] W. Jiang, D. zhang, L. Ling, G. Cai, L. Zeng, Time series to imaging-based deep learning model for detecting abnormal fluctuation in agriculture product price, Soft Comput. 27 (20) (Oct. 2023) 14673–14688, https://doi.org/10.1007/s00500-023-09121-9.

[67] G. of Indonesia, Presidential regulation No. 59 of 2020 on national food price stability [Online]. Available: https://peraturan.bpk.go.id/Details/136174/perpres-no-59-tahun-2020, 2020.

[68] C. Arthur, K. D. Hartomo. Enhancing breast cancer prediction with an advanced k-nearest neighbors (KNN) algorithm integrated with feedback support mechanism, IEEE, Dec. 2023, pp. 1–5, https://doi.org/10.1109/icteca60133.2023.10491036.

[69] N. Wu, B. Green, X. Ben, S. O'Banion, Deep transformer models for time series forecasting: the influenza prevalence case [Online]. Available: http://arxiv.org/abs/2001.08317, Jan. 2020.

[70] R. Chiong, Z. Fan, Z. Hu, M.T.P. Adam, B. Lutz, D. Neumann, A sentiment analysis-based machine learning approach for financial market prediction via news disclosures, in: J.B. Sartor, T. D'Hondt, W. De Meuter (Eds.), Proceedings of Genetic and Evolutionary Computation Conference Companion (GECCO'18 Companion), ACM, New York, NY, USA, Jul. 2018, pp. 278–279, https://doi.org/10.1145/3205651.3205682.