ISSN 2085-4552

Machine Learning for Chili Pepper Price Forecasting Using Exogenous Public-Attention Signals and Bayesian Hyperparameter Optimization

Wresti Andriani[1], Gunawan Gunawan[2], Naella Nabila Putri Wahyuning Naja[3]
[1] Departemen of Informatika, Universitas Bima Sakapenta, Tegal, Indonesia
[2] Depertemen of Sistem Informasi, Universitas Pancasakti, City, Indonesia
[3]Departemen of Ekonomi Manajemen, Universitas Negeri Semarang, Semarang, Indonesia
[1] wresty.andriani@gmail.com, [2] Waeisul.wbn@bsi.ac.id, [3] naellanabilap@students.unnes.ac.id

Accepted 21 June 2026
Approved 28 June 2026

***Abstract*— Chili pepper prices in Indonesia are highly volatile due to seasonal production, distribution constraints, and short-term market reactions to public issues. This study proposes a machine-learning approach for short-term chili pepper price forecasting by integrating public-attention signals from Google Trends and news volume into a composite Shock Index. A LightGBM model optimized using Bayesian hyperparameter optimization was evaluated through rolling time-series cross-validation across multiple forecasting horizons. The results show that the attention-augmented scenario, which combines price features, Google Trends, news volume, and the Shock Index, consistently outperforms the price-only baseline. After optimization, the best scenario achieved an average sMAPE of 12.47% and maintained MASE below 1, indicating better performance than the naïve benchmark. These findings suggest that public-attention signals provide useful early-warning information for forecasting chili price movements, particularly when market uncertainty increases. The proposed approach can support nowcasting and price-monitoring systems for policymakers and market participants, although its performance remains dependent on the stability and representativeness of online attention data.**

***Index Terms*— Bayesian hyperparameter optimization ; chili price forecasting; LightGBM; public-attention signals; Shock Index.**

I. Introduction

Chili peppers are a strategic horticultural commodity in Indonesia because they affect household expenditure, food inflation, and food price stability[1]. Seasonal production dynamics, together with supply chains that are sensitive to weather, logistics costs, and distribution disruptions, make chili prices highly volatile [2]. In addition to these structural factors, public perception and media attention toward specific issues, such as food-safety concerns, imported chili, or distribution shortages, may amplify short-term market expectations [3]. When such issues receive high online search interest and news coverage, they can influence consumer concern, trader sentiment, and anticipatory market behavior. Therefore, chili price forecasting should not rely only on historical price patterns but also consider high-frequency external signals that may function as early-warning indicators.

Prior studies provide an important foundation for chili price forecasting in Indonesia. Pangesti et al.[4] compared ARIMA and a Gated Recurrent Unit (GRU) model for forecasting chili prices in East Java and found that GRU produced better forecasting accuracy. Prihandi et al.[5] applied ARIMA with weather-related variables, including rainfall and sunlight duration, to examine the relationship between production conditions and red chili prices in North Sumatra. Meanwhile, Gunadi and Perdana [6]. used a univariate Long Short-Term Memory (LSTM) model to forecast chili prices in Metro City. These studies indicate that nonlinear models and external covariates can improve forecasting performance, but their implementations remain limited in terms of feature scope, evaluation design, and optimization strategy.

Based on these prior studies, three research gaps can be identified. First, many chili price forecasting models still rely mainly on historical price data and do not systematically include high-frequency exogenous signals that represent public attention, such as online search interest and news intensity. Second, although ARIMA-based models can incorporate external variables, they are limited in capturing nonlinear relationships and rapidly changing attention dynamics. Third, many studies still rely on a single holdout evaluation, while rolling or expanding multi-fold backtesting, Bayesian hyperparameter optimization, ablation analysis, and interpretability testing are not consistently applied. These gaps limit the reproducibility and operational reliability of existing forecasting models.

This study aims to: (1) improve short-horizon chili pepper price forecasting accuracy by integrating public-attention signals as exogenous features; (2) evaluate the contribution of Google Trends, news volume, and their composite Shock Index through an ablation study; (3) apply rolling time-series cross-validation to obtain more reliable out-of-fold performance estimates; and (4) optimize LightGBM configurations using Bayesian hyperparameter optimization. The forecasting performance is evaluated using sMAPE, RMSE, and MASE across multiple horizons, namely H = 1, 2, and 4 weeks.

To achieve these objectives, this study collects chili price data by region and combines them with public-attention signals from Google Trends and news volume. The feature engineering process includes price lags, moving averages, rolling volatility, calendar features, and a composite Shock Index representing the intensity of public issues. A LightGBM model is then trained and optimized using Bayesian hyperparameter optimization under a rolling cross-validation scheme. Feature contributions are examined through ablation scenarios and interpretability analysis.

The contribution of this study is twofold. Academically, it provides empirical evidence on the role of public-attention signals in improving machine-learning-based forecasting of horticultural commodity prices. Practically, it offers a reproducible forecasting framework that can support nowcasting and early-warning systems for policymakers, market participants, and other stakeholders concerned with food price stability.

II. Literature Review

Price volatility in chili pepper markets has received considerable attention in Indonesia because chili is closely related to household consumption, farmer income, and food inflation. At the market-structure level, Sukiyono and Asriani [7] found that chili prices in Bengkulu showed high volatility at the producer and wholesaler levels, with strong price transmission across vertical markets. This finding indicates that chili price movements are not determined only by production seasonality, but also by market-chain interactions, distribution mechanisms, and expectations formed by market actors.

Classical time-series models remain important in agricultural price forecasting. Windhy and Jamil [8] showed that ARIMA can be used to forecast Indonesian red chili prices by relying on historical price patterns and standard diagnostic procedures, including stationarity testing and ACF/PACF analysis. However, ARIMA-based approaches generally assume relatively linear relationships and may have limitations when dealing with nonlinear changes, sudden shocks, and regime shifts that frequently occur in horticultural commodity markets. Therefore, more flexible forecasting methods are needed to capture complex and changing price behavior.

Recent studies have increasingly adopted machine-learning and deep-learning methods for chili price prediction. Pangesti et al. [4] compared ARIMA and GRU models for chili price forecasting in East Java and found that GRU produced better forecasting performance. Gunadi and Perdana [6] applied LSTM to forecast chili prices in Metro City and reported promising predictive accuracy. These studies show that nonlinear models can improve price forecasting compared with purely classical models. Nevertheless, many deep-learning applications still rely mainly on historical price data, making them less sensitive to external shocks that may appear outside the price series itself.

For tabular time-series forecasting, gradient boosting models such as LightGBM have become attractive because they can handle nonlinear relationships, interactions among variables, and heterogeneous feature structures efficiently. Oikonomou and Damigos [9] demonstrated that AutoReg-LightGBM and LightGBM–ARIMA ensembles can perform competitively for short-term commodity price forecasting under rolling-origin evaluation. This suggests that LightGBM is suitable for forecasting problems involving lagged price features and external variables. However, its use in Indonesian chili price forecasting remains relatively limited, especially when combined with public-attention indicators and systematic hyperparameter optimization.

In addition to historical prices and market fundamentals, public-attention signals can provide useful information for nowcasting and short-term forecasting. Macias et al. [10]. showed that high-frequency online price information can improve food inflation nowcasting compared with traditional data sources. In commodity markets, Chi et al. [3] found that media emotion intensity is associated with commodity return dynamics, indicating that news and public sentiment may influence short-term market behavior. In the context of chili prices, increases in search interest and news coverage about issues such as imported chili, food safety, supply disruption, or price spikes may reflect consumer concern and trader expectations. These signals may not directly determine prices, but they can capture market attention before or during price movements.

Based on the reviewed literature, three main research gaps can be identified. First, many Indonesian chilies price studies still focus on univariate forecasting or limited covariates, while high-frequency public-attention signals such as Google Trends and news volume are rarely integrated systematically. Second, although LightGBM has shown potential in commodity forecasting, its application to chili price forecasting with public-attention features remains underexplored. Third, previous studies often rely on a single holdout evaluation, while rolling time-series cross-validation, Bayesian hyperparameter optimization, ablation analysis, and interpretability testing is not consistently applied. This study addresses these gaps by combining LightGBM, Bayesian hyperparameter optimization, rolling backtesting, and a composite Shock Index derived from Google Trends and news volume to improve short-term chili pepper price forecasting.

III. METHOD

This chapter outlines the methodological design employed to forecast short-term chili pepper prices by leveraging exogenous features that capture the intensity of public issues/attention, as in Figure 1.

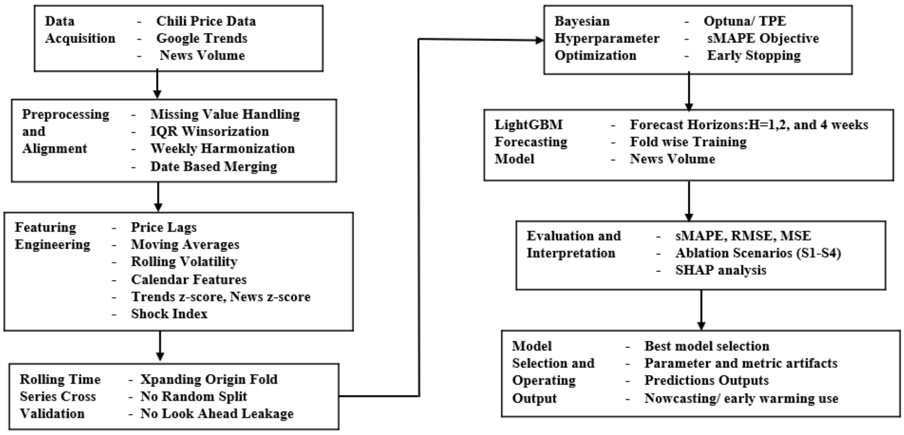

Fig. 1. Research framework for chili pepper price forecasting using public-attention signals, Bayesian hyperparameter optimization, and LightGBM

Figure 1 illustrates the proposed research framework. The process begins with the collection of chili price data and public-attention signals from Google Trends and news volume. The data are then preprocessed through missing-value handling, outlier treatment, frequency harmonization, and date-based alignment. Feature engineering is conducted by constructing price lags, moving averages, rolling volatility, calendar variables, rolling z-scores of public-attention signals, and the composite Shock Index. The forecasting model is trained using LightGBM under an expanding-origin rolling cross-validation scheme to avoid look-ahead leakage. Bayesian hyperparameter optimization is applied using sMAPE as the main objective function. Model performance is evaluated across 1-, 2-, and 4-week horizons using sMAPE, RMSE, and MASE, followed by ablation analysis and SHAP-based interpretation. The best configuration is then selected and prepared for nowcasting and early-warning applications.

a) Data acquisition

Price data were collected from official Indonesian food-price information sources, including the Bank Indonesia Food Price Information Center and the Ministry of Trade SP2KP dashboard [11], [12]. The observed commodities include red cayenne chili and red curly chili. The regional scope covers selected Indonesian regions, namely DKI Jakarta, West Java, East Java, and North Sumatra, to allow cross-regional comparison of chili price behavior.

Public-attention signals were collected from Google Trends and GDELT News Timeline [13], [14]. Google Trends was used to capture online search interest, while GDELT-based news volume was used to represent the intensity of public news coverage related to chili-market issues. The search terms included issue-related keywords such as imported chili, Chinese chili, toxic chili, pesticide chili, and spelling variants of cabai and cabe. These keywords were selected because they represent public issues that may influence consumer concern, trader expectations, and short-term market responses.

The observation period covers January 2024 to December 2025. All variables were harmonized into weekly frequency before being merged into a single modeling dataset. The forecasting horizons were set to 1, 2, and 4 weeks ahead. A summary of the dataset structure is presented in Table I.

TABLE I. DATASET SUMMARY

Component	Description
Price data	Chili pepper price series from official Indonesian food-price sources
Commodities	Red cayenne chili and red curly chili
Regions	DKI Jakarta, West Java, East Java, and North Sumatra
Observation period	January 2024–December 2025
Frequency	Weekly after temporal harmonization
Public-attention signals	Google Trends score and news volume
Exogenous keywords	Imported chili, Chinese chili, toxic chili, pesticide chili, and spelling variants of cabai/cabe
Forecast horizons	1, 2, and 4 weeks ahead
Evaluation design	Expanding-origin rolling time-series cross-validation

Because this study uses time-series data, random train-test splitting was not applied. Instead, an expanding-origin rolling cross-validation scheme was used to preserve chronological order and reduce the risk of look-ahead leakage [15], [16].

b) Preprocessing and Alignment

The preprocessing stage was conducted to ensure consistency across price and public-attention variables. Missing values in the price series and exogenous variables were handled using conservative time-series imputation, including short interpolation and forward filling where appropriate, following common time-series preprocessing practices [13]. Extreme values were treated using IQR-based winsorization to reduce the effect of non-informative spikes while preserving meaningful price movements.

After missing-value and outlier handling, all data sources were harmonized into weekly frequency. Price data, Google Trends scores, and news-volume indicators were then merged by date. Exogenous variables were lagged where necessary to ensure that only information available at time t was used to forecast future prices. This procedure was applied to prevent look-ahead bias and maintain the validity of the time-series forecasting design [14], [17].

c) Feature Engineering

Two groups of predictors were constructed: endogenous price-based features and exogenous public-attention features. Endogenous features include price lags, moving averages, rolling volatility, and calendar variables. Price lags were used to capture short-term persistence in chili price movements, while moving averages and rolling volatility were used to represent trend and instability patterns [10], [18]

Exogenous features were derived from Google Trends and news volume. These variables were transformed using past-only rolling z-scores so that the intensity of public attention could be compared over time. Several lag structures were applied to capture delayed market responses, since online searches or news coverage may influence prices with a time lag rather than immediately [19].

We also define a composite Shock Index like formula 1.

$$Shock\ Index = z\ (Trends) + z\ (News) \quad (1)$$

where z(Trends) denotes the rolling z-score of Google Trends search interest at time t, and z(News) denotes the rolling z-score of news volume at time t. The Shock Index serves as a numerical proxy for the intensity of public issues that may influence short-term chili price movements [20].

d) Rolling Time Series Cross-Validation

Model evaluation was performed using an expanding-origin rolling cross-validation scheme. This method is suitable for time-series forecasting because it preserves the chronological order of observations and avoids information leakage from future periods into the training process [15] [16]. In each fold, the model was trained using historical observations and validated using the next chronological period. The validation window then moved forward sequentially across folds.

For each forecasting horizon h∈{1,2,4}, the target variable was defined by shifting the chili price series forward by h weeks. This means that the model used information available at time t to predict chili prices at time t+h. Out-of-fold predictions from all validation windows were aggregated to compute the final evaluation metrics.

- Fold notation and aggregation.

For horizon h ∈ {1,2,4} and validation indices V_k in fold k, we form predictions and aggregate out of fold matrc across fold, as the formula 2.

$$V_h = U_{k-1}^{k} V_k, \quad n_h = [V_h] \quad (2)$$

All metrics below are computed on V_h. Primary accuracy measures, like formula 3, 4 and 5.

- sMAPE

$$sMAPE = \frac{200}{n_h} \sum_{t \in V_h} \frac{[y_t - y_{t.h}]}{[y_t] + [y_{y.}]} \quad (3)$$

- RMSE:

$$\text{RMSE} = \sqrt{\frac{1}{n_h} \sum_{t \in V_h} (y_t - \bar{y}_{t.h})^2} \quad (4)$$

- MASE

$$MASE = \frac{\frac{1}{nh} \sum t \in V_h [y_t - \vec{y}t.h]}{\frac{1}{N-1} \sum_{t=2}^{N} [y_t - \vec{y}t.1]} \quad (5)$$

where N is the in-sample length used to compute the naïve MAE in the denominator.

Optional accuracy comparison (Diebold–Mariano test). To compare two models, define the loss differential

$$d_t = L(e_{1,t}) - L(e_{2.t}) \quad (6)$$

Use either expanding or rolling training windows as appropriate and always evaluate OOF predictions to obtain unbiased performance estimates.

e) Bayesian Hyperparameter Optimization

Bayesian hyperparameter optimization was applied to improve the LightGBM model configuration. The optimization process was conducted using a Tree-structured Parzen Estimator approach, which is commonly used to search for efficient hyperparameter configurations by modeling the relationship between candidate parameters and objective performance [21] [22] [23] . n each trial, a candidate hyperparameter set was evaluated using out-of-fold predictions generated from the rolling cross-validation scheme. This ensured that the optimization objective reflected realistic forecasting performance rather than performance from a random split.

The main objective function for optimization was the average sMAPE across forecasting horizons. Early stopping was applied during model training to reduce overfitting and improve computational efficiency. The optimized hyperparameters included learning rate, number of estimators, maximum depth, number of leaves, minimum child samples, subsampling parameters, and regularization parameters.

f) Training per Cross-Validation Fold)

For each fold k in the rolling CV scheme and each horizon h∈ {1,2,4}[15]. The model used is a LightGBM Regressor with the best hyperparameters θ obtained from Bayesian HPO [16]. Training is performed on T_k and evaluated with early stopping on the validation window V_k the model is saved at its best iteration and produces predictions [18]. Predictions from all folds are then combined into out-of-fold (OOF) predictions for metric computation. Here's the pseudocode at figure 2.

```
for each fold k in rolling_CV:
    split data → T_k (train), V_k (validation)
    for each horizon h ∈ {1, 2, 4}:
        y_target = shift (price, -h); X = feature_matrix_at_t
        model = LightGBM (params = θ*); model.fit(X[T_k], y_target[T_k], early_stopping_on=V_k)
        ŷ_OOF [V_k, h] = model.predict(X[V_k]); save (best_iteration, feature_importance)
aggregate ŷ_OOF across folds → compute sMAPE, RMSE, MASE; select the best configuration
```

Fig. 2. Pseudocode Training per Cross-Validation Fold)

g) LightGBM Model Training

The forecasting model used in this study was LightGBM Regressor. LightGBM was selected because gradient boosting models are suitable for nonlinear tabular data and can effectively handle interactions between lagged price features and exogenous variables [9], [24]. For each fold in the rolling cross-validation scheme and each forecasting horizon, the model was trained using the selected feature set and the best hyperparameter configuration obtained from Bayesian optimization. The model was then evaluated on the validation window and used to generate out-of-fold predictions.

h) Ablation Scenarios

To assess the marginal contribution of public-attention signals to forecasting accuracy, we compare four feature scenarios under an identical training configuration (same folds, random seeds, HPO procedure, early stopping, and evaluation metrics). Scenarios:

1. S1, Price-only**:** endogenous features only (price lags, moving averages, volatility, calendar) [24].
2. S2, Trends**:** S1 plus Google Trends (lags 0–7; rolling z-score)[19], [25] .
3. S3, News: S1 plus news volume (lags 0–7; rolling z-score) [3], [26].
4. S4, Both (Shock Index): S1 plus both signals and the Shock Index [27], [28].

i) Model Selection & Artifacts

The best model configuration was selected based on the lowest average sMAPE across the forecasting horizons of 1, 2, and 4 weeks using out-of-fold predictions. When two or more configurations produced similar sMAPE values, RMSE and MASE were used as supporting criteria, with preference given to models that maintained MASE below 1, indicating better performance than the naïve benchmark [29]. When necessary, statistical comparison between top-performing configurations can be conducted using the Diebold–Mariano test to examine whether differences in predictive accuracy are statistically meaningful [30].

After Bayesian hyperparameter optimization, the top configurations were rechecked through the same rolling validation procedure to confirm performance stability. The final model was then retrained using the selected configuration, and key experimental artifacts were recorded to support reproducibility. These artifacts include metric summaries, best hyperparameters, out-of-fold predictions, prediction plots, SHAP summaries, training logs, the final LightGBM model file, and configuration metadata such as random seed and package versions [17], [31]. This artifact-logging process ensures that the experiment can be verified, repeated, and prepared for future integration into nowcasting or early-warning systems.

IV. Result and Discussions

This section presents the empirical results of the proposed chili pepper price forecasting model. The analysis includes data representation, feature construction, baseline evaluation before Bayesian hyperparameter optimization, post-optimization performance, and interpretation of the contribution of public-attention signals. The evaluation was conducted across three forecasting horizons: 1, 2, and 4 weeks ahead. Four feature scenarios were compared: S1: Price-only, S2: Price + Google Trends, S3: Price + News Volume, and S4: Price + Google Trends + News Volume + Shock Index.

A. Data Representation and Feature Construction

The empirical dataset was constructed from chili price series and public-attention signals. Price data were obtained from official Indonesian food-price information sources, while public-attention signals were represented by Google Trends scores and news-volume indicators. To avoid presenting an excessively large raw dataset in the manuscript, Table II provides a representative sample of the merged weekly data. This sample illustrates how chili price data, online search interest, and news intensity were aligned before feature engineering.

TABLE II. REPRESENTATIVE SAMPLE OF MERGED PRICE AND PUBLIC-ATTENTION DATA

Week	Chili type	Price per kg	Trends score	News count
2025-10-06	Red Cayenne	49.200	38	1
2025-10-13	Red Cayenne	50.800	62	2
2025-10-20	Red Cayenne	53.600	71	3
…	…	…	…	…
2025-10-27	Red Cayenne	52.400	58	2

The representative sample indicates that increases in Google Trends score, and news count occur during periods of higher chili prices. Although this pattern should not be interpreted as direct causality, it suggests that public attention may contain useful short-term information related to market concern, consumer expectation, or issue salience. These signals were therefore transformed into rolling z-scores and combined into the Shock Index.

TABLE III. REPRESENTATIVE SAMPLE OF PUBLIC-ATTENTION FEATURES

Week	Trend z	News z	Shock Indeks
2025-10-06	-0,40	-0,35	-0,7
2025-10-13	0,85	0,90	1,7
2025-10-20	1,35	1,55	2,9
…	…	…	…
2025-10-27	0,70	0,50	1,2

The Shock Index was used to summarize the intensity of public attention by combining search interest and news volume. A higher Shock Index indicates that chili-related public issues received stronger online and media attention. In the context of chili markets, such attention may emerge from issues related to imported chili, food safety, supply shortages, or sudden price increases. These public signals may influence market behavior through consumer concern, trader anticipation, and short-term expectation formation.

B. Baseline Performance Before Bayesian HPO

At this stage, we present the initial (baseline) forecasting performance prior to hyperparameter tuning via Bayesian optimization. The evaluation uses out-of-fold (OOF) predictions under a rolling cross-validation scheme across three weekly forecasting horizons (H = 1, 2, 4). To ensure a fair comparison, all scenarios employ identical fold splits, random seeds, and training procedures, with no hyperparameter optimization applied (a conservative default configuration).

Table 4 presents the baseline forecasting performance before Bayesian hyperparameter optimization. All scenarios were evaluated using the same rolling time-series cross-validation scheme, forecasting horizons, and evaluation metrics. This ensures that the comparison reflects the contribution of different feature sets rather than differences in evaluation procedure.

TABLE IV. BASELINE METRICS (BEFORE BAYESIAN HPO)

Scenario	Horizon	sMAPE	RMSE	MASE
S1: price only	1	12,9	3,150	0,9
S1: price only	2	15,6	4,020	0,9
S1: price only	4	19,8	5,180	1,0
S2: plus trends	1	11,6	2,890	0,8
S2: plus trends	2	14,2	3,720	0,9
S2: plus trends	4	18,1	4,880	1,0
S3: plus news	1	11,9	2,950	0,8
S3: plus news	2	14,0	3,690	0,9
S3: plus news	4	17,4	4,710	1,0
S4: plus both	1	10,5	2,680	0,8
S4: plus both	2	12,8	3,450	0,8
S4: plus both	4	16,6	4,550	0,9

At table 4 (OOF rolling CV; H = 1/2/4) show, Cross-horizon pattern. All scenarios exhibit larger errors as the forecast horizon lengthens: sMAPE and RMSE are lowest at H=1 and highest at H=4, which is consistent with the increasing difficulty of longer-range forecasting.

Scenario ranking (best → worst). S4 (+Trends +News +Shock Index) dominates at every horizon, followed by S3 (+News) ≈ S2 (+Trends), and then S1 (price-only). As an illustration of sMAPE: H=1: S4 10.5% vs S1 12.9%; H=2: 12.8% vs 15.6%; H=4: 16.6% vs 19.8%.

Value of public-attention signals. Relative to S1, adding a single exogenous signal already helps: S2 (Trends) is slightly better at H=1 (11.6% vs S3's 11.9%), suggesting search interest is more responsive for very short horizons. S3 (News) wins at H=2/4 (14.0% & 17.4% vs 14.2% & 18.1%), indicating news volume is more informative for longer horizons. S4 (the combination with the Shock Index) delivers the strongest synergy, not just a simple sum of S2 and S3.

Practical gains of S4 vs S1. sMAPE: −2.4 p.p. (H=1), −2.8 p.p. (H=2), −3.2 p.p. (H=4). RMSE: −470 (H=1), −570 (H=2), −630 (H=4). MASE: −0.11 (H=1), −0.10 (H=2), −0.09 (H=4). These improvements grow with the horizon, implying exogenous signals become more valuable as uncertainty increases.

Naïve benchmark (MASE). S1 fails to beat the naïve forecast at H=4 (MASE 1.07 > 1), whereas S4 stays < 1 across all horizons (0.81 / 0.88 / 0.98), underscoring that relying solely on historical prices is insufficient for longer horizons. Overall, public-attention features deliver clear benefits even before hyperparameter tuning. The grafik can be shown in Figure 3.

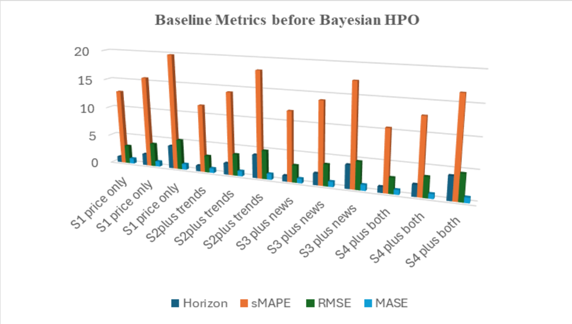

Fig. 3. Baseline Metrics before Bayesian HPO

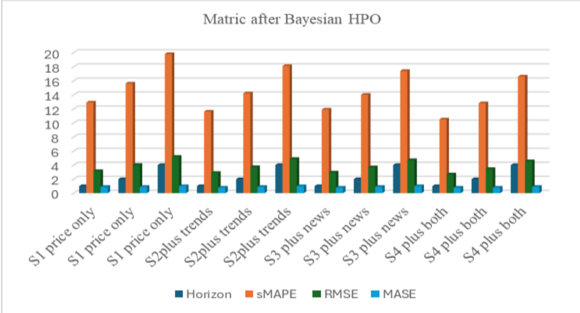

Fig. 4. Metrics after Bayesian HPO

C. Performance After Bayesian HPO

At this stage, the models are retrained with Bayesian hyperparameter optimization (HPO) using the same rolling cross-validation scheme as the baseline. The search space (depth, learning rate, number of iterations/trees, regularization, and leaf size) is applied uniformly across all scenarios (S1–S4) with identical seeds, folds, and early-stopping procedures, so performance differences reflect the benefit of tuning rather than procedural variation. The evaluation metrics remain sMAPE (primary), RMSE, and MASE at H = 1/2/4 weeks.

Overall, HPO reduces error in every scenario: both sMAPE and RMSE are consistently lower than pre-HPO, while MASE moves closer to or below 1 (indicating performance beyond the naïve benchmark). S4 (+Trends +News +Shock Index) remains the best configuration across all horizons, suggesting that combining public-attention signals summarized in the Shock Index provides a structural advantage even after optimization. The relative gains are typically more pronounced at longer horizons, indicating that exogenous information becomes more valuable as uncertainty increases. The result as tabel 5.

TABLE 5. BASELINE METRICS (AFTER BAYESIAN HPO)

Scenario	Horizon	sMAPE	RMSE	MASE
S1: price only	1	12,6	3,080	0,9
S1: price only	2	15,2	3,950	0,9
S1: price only	4	19,2	5,100	1,0
S2: plus, trends	1	11,3	2,830	0,8
S2: plus, trends	2	13,8	3,660	0,8
S2: plus, trends	4	17,7	4,800	1,0
S3: plus, news	1	11,5	2,890	0,8
S3: plus, news	2	13,6	3,620	0,9
S3: plus, news	4	16,9	4,630	0,9
S4: plus, both	1	9,8	2,550	0,7
S4: plus, both	2	12,0	3,350	0,8
S4: plus, both	4	15,6	4,400	0,9

And the grafic for metric after Bayesian HPO can be seen at Figure 4.

D. Average Performance Across Forecasting Horizons

Table 6 summarizes the average post-HPO performance across the 1-, 2-, and 4-week horizons.

TABLE VI. THE AVERAGES ACROSS HORIZONS (H = 1, 2, AND 4 WEEKS)

Skenario	sMAPE (%)	RMSE	MASE
S1_price_only	15,67	4,043	0,9
S2_plus_trends	14,27	3,763	0,9
S3_plus_news	14,00	3,713	0,9
S4_plus_both	12,47	3,433	0,8

The average results show that S4 produced the lowest sMAPE, RMSE, and MASE. Its average sMAPE of 12.47% indicates the strongest overall accuracy among the four scenarios. Compared with S1, the S4 model reduced average sMAPE by 3.20 percentage points. This confirms that the combination of price-based features and public-attention signals provides better forecasting performance than relying only on historical price information.

The results also show that S3 slightly outperformed S2 on average. This suggests that news volume may provide more persistent information than search interest in the observed dataset. Google Trends can reflect immediate public curiosity or concern, while news coverage may represent broader and more sustained issue salience. However, the best performance was achieved when both signals were combined through the Shock Index, indicating that search attention and news attention are complementary rather than interchangeable.

E. Interpretation of Public-Attention Effects

The main empirical finding is that Google Trends and news volume improve chili price forecasting because they capture information that is not fully reflected in historical price movements. Chili markets are sensitive not only to supply and demand fundamentals, but also to public issues such as imported chili, food-safety concerns, and distribution shortages. When such issues receive high search activity or news coverage, they may affect consumer concern, trade expectations, and short-term market behavior.

Google Trends is useful because it captures direct public search behavior. A rise in search intensity may indicate that consumers or market participants are paying attention to chili-related issues before price changes become fully visible in official price data. This helps the model detect early signals of market concern.

News volume, on the other hand, reflects the intensity of media coverage. When news coverage increases, the issue may gain broader public visibility and influence expectations for a longer period.

The Shock Index strengthens forecasting performance because it combines these two types of attention signals. Search interest captures public curiosity and concern, while news volume captures issue amplification through media coverage. Their combination provides a more comprehensive representation of public-attention dynamics. This explains why S4 consistently outperformed S2 and S3 individually.

These findings support the argument that public-attention signals can serve as early-warning indicators in agricultural commodity forecasting. However, the relationship should be interpreted carefully. Public attention does not necessarily cause price changes directly. Instead, it can act as a proxy for market concerns, information diffusion, and expectation formation, which may accompany or precede short-term chili price movements.

F. Summary of Findings

Overall, the results demonstrate three key findings. First, LightGBM with Bayesian HPO improves chili price forecasting performance compared with the non-optimized baseline. Second, public-attention signals from Google Trends and news volume provide additional predictive value beyond historical price features. Third, the full model that combines price features, Google Trends, news volume, and the Shock Index achieves the best performance across all forecasting horizons.

The post-HPO S4 model achieved the best average performance, with an average sMAPE of 12.47%, RMSE of 3,433 IDR/kg, and MASE of 0.87. These results indicate that the model outperformed the naïve benchmark and is suitable for supporting short-term chili price monitoring. The widening performance gap at longer horizons further suggests that public-attention signals are particularly useful when uncertainty increases.

V. Conclusions

This study proposed a machine-learning approach for short-term chili pepper price forecasting by integrating public-attention signals from Google Trends and news volume into a composite Shock Index. The results show that the attention-augmented model, represented by S4: Price + Google Trends + News Volume + Shock Index, consistently outperformed the price-only baseline across the 1-, 2-, and 4-week forecasting horizons. After Bayesian hyperparameter optimization, S4 achieved the best average performance, with an average sMAPE of 12.47%, RMSE of 3,433 IDR/kg, and MASE of 0.87. These results indicate that the proposed model performed better than the naïve benchmark and that public-attention signals provide additional predictive value beyond historical price information.

The findings suggest that Google Trends and news volume can serve as useful early-warning indicators for chili price movements. Search interest may reflect consumer concern and immediate public curiosity, while news volume may capture broader issue salience and market narratives. When combined into the Shock Index, these signals help the model identify short-term attention dynamics related to issues such as imported chili, food-safety concerns, supply disruptions, or sudden price increases. Therefore, the proposed LightGBM model with Bayesian hyperparameter optimization can support nowcasting and early-warning applications for policymakers, market participants, and stakeholders concerned with food price stability.

However, this study has several limitations. First, the model depends on online public-attention signals, which may not fully represent all market actors, especially farmers, traders, or consumers with limited digital activity. Second, Google Trends and news-volume data can be unstable because they are influenced by changes in search behavior, media reporting intensity, keyword selection, and platform-specific data availability. Third, the empirical scope is limited to selected Indonesian regions and chili commodities, so the model's performance may not generalize directly to other regions, commodities, or market structures without further validation. In addition, public-attention signals should be interpreted as proxies of market concern and information diffusion, not as direct causal determinants of price changes.

Future research may extend this approach by incorporating additional exogenous variables, such as weather, production volume, distribution costs, market supply, and regional consumption patterns. Further studies may also test the model across broader geographical areas and different agricultural commodities to evaluate its robustness. Combining public-attention signals with fundamental market indicators may provide a more comprehensive forecasting framework for food-price monitoring and policy-oriented early-warning systems.

References

[1] M. Rachmaniah, A. I. Suroso, M. Syukur, and I. Hermadi, "Strategic Food Risks – Chili'S Agrosystem Perspective," *J. Manag. Agribisnis*, vol. 18, no. 1, pp. 19–31, 2021, doi: 10.17358/jma.18.1.19.

[2] Y. N. Muflikh, C. Smith, C. Brown, N. Kusnadi, A. M. Kiloes, and A. A. Aziz, "Integrating system dynamics to value chain analysis to address price volatility in the Indonesian chilli value chain," *Food Policy*, vol. 128, p. 102713, 2024, doi: doi.org/10.1016/j.foodpol.2024.102713.

[3] Y. Chi, L. El-Jahel, and T. Vu, "Media emotion intensity and commodity futures pricing," *J. Commod. Mark.*, vol. 37, p. 100460, 2025, doi: doi.org/10.1016/j.jcomm.2025.100460.

[4] W. Pangesti, N. Syukri, K. A. Notodiputro, Y. Angraini, and L. N. A. Mualifah, "Performance Evaluation of ARIMA and GRU Models for Forecasting Chili Price in East Jawa," *JUITA J. Inform.*, vol. 13, no. 2, pp. 209–218,

2025, doi: 10.30595/juita.v13i2.26445.

[5] I. Prihandi, S. Wijono, I. Sembiring, and E. Maria, "Implementation of ARIMA with Min-Max Normalization for predicting the Price and Production Quantity of Red Chili Peppers in North Sumatra Province considering Rainfall and Sunlight Duration Factors," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 2, pp. 21876–21887, 2025, doi: 10.48084/etasr.9875.

[6] G. M. Gunadi and A. Perdana, "Forecasting Chili Prices in Metro City Using Long Short-Term Memory (LSTM)," *Int. J. Softw. Eng. Comput. Sci.*, vol. 5, no. 1, pp. 215–225, 2025, doi: 10.35870/ijsecs.v5i1.3526.

[7] K. Sukiyono and P. S. Asriani, "Chili Price Volatilities and Transmisions at Vertical Markets in Bengkulu Province," *J. Agro Ekon.*, vol. 38, no. 1, pp. 29–39, 2020.

[8] A. M. Windhy and A. S. Jamil, "Forecasting Indonesian Red Chilli Prices: The ARIMA Approach," *Agriekstensia*, vol. 20, no. 1, pp. 90–106, 2021, doi: 10.34145/agriekstensia.v20i1.1502.

[9] K. Oikonomou and D. Damigos, "Short term forecasting of base metals prices using a LightGBM and a LightGBM - ARIMA ensemble," *Miner. Econ.*, vol. 38, no. 1, pp. 37–49, 2025, doi: 10.1007/s13563-024-00437-y.

[10] P. Macias, D. Stelmasiak, and K. Szafranek, "Nowcasting food inflation with a massive amount of online prices," *Int. J. Forecast.*, vol. 39, no. 2, pp. 809–826, 2023, doi: doi.org/10.1016/j.ijforecast.2022.02.007.

[11] B. K. and H. Muhammad Ridwansyah Matondang, "Price Fluctuations and Volatility of National Strategic Food Commodities," vol. 8, no. 1, p. 134, 2023, [Online]. Available: http://ejournal2.undip.ac.id/index.php/agrisocionomics

[12] M. David, I. Cholissodin, and N. Yudistira, "Chili Price Prediction Using the Long Short-Term Memory Method (Case Study: Malang City)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 1214–1219, 2023, [Online]. Available: http://j-ptiik.ub.ac.id

[13] N. Niako, J. D. Melgarejo, G. E. Maestre, and K. P. Vatcheva, *Effects of missing data imputation methods on univariate blood pressure time series data analysis and forecasting with ARIMA and LSTM*, vol. 24, no. 1. 2024. doi: 10.1186/s12874-024-02448-3.

[14] X. Wang, M. Zhang, and J. Su, "Robust temporal alignment for multivariate time series forecasting," *Expert Syst. Appl.*, vol. 289, p. 128299, 2025, doi: doi.org/10.1016/j.eswa.2025.128299.

[15] A. Deng, "Time series cross validation: A theoretical result and finite sample performance," *Econ. Lett.*, vol. 233, p. 111369, 2023, doi: doi.org/10.1016/j.econlet.2023.111369.

[16] A. Vamsikrishna and E. V Gijo, "New Techniques to Perform Cross-Validation for Time Series Models," *Oper. Res. Forum*, vol. 5, no. 2, p. 51, 2024, doi: 10.1007/s43069-024-00334-8

[17] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, p. 100804, 2023, doi: doi.org/10.1016/j.patter.2023.100804.

[18] T. Hall and K. Rasheed, "A Survey of Machine Learning Methods for Time Series Prediction," *Appl. Sci.*, vol. 15, no. 11, 2025, doi: 10.3390/app15115957.

[19] A. Chawla, "Real-Time Nowcasting of CPI and Food Inflation with Google Trends-Evidence from the US, UK, and India (2012–2025," *Int. J. Multidiscip. Res.*, vol. 7, no. 5, pp. 1–14, 2025, doi: 10.36948/ijfmr.2025.v07i05.55463.

[20] A. Pareto, "Normalization (of an Indicator) in Composite Indicator Construction BT - Encyclopedia of Quality of Life and Well-Being Research," F. Maggino, Ed., Cham: Springer International Publishing, 2023, pp. 4717–4719. doi: 10.1007/978-3-031-17299-1_104641.

[21] M. A. K. Raiaan *et al.*, "A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks," *Decis. Anal. J.*, vol. 11, p. 100470, 2024, doi: doi.org/10.1016/j.dajour.2024.100470.

[22] S. Hanifi, A. Cammarono, and H. Zare-Behtash, "Advanced hyperparameter optimization of deep learning models for wind power prediction," *Renew. Energy*, vol. 221, p. 119700, 2024, doi: doi.org/10.1016/j.renene.2023.119700.

[23] B.-N. Jiang, Y.-Y. Zhang, Z.-Y. Zhang, Y.-L. Yang, and H.-L. Song, "Tree-structured parzen estimator optimized-automated machine learning assisted by meta–analysis for predicting biochar–driven N2O mitigation effect in constructed wetlands," *J. Environ. Manage.*, vol. 354, p. 120335, 2024, doi: doi.org/10.1016/j.jenvman.2024.120335.

[24] J. Gong, Z. Qu, Z. Zhu, H. Xu, and Q. Yang, "Ensemble models of TCN-LSTM-LightGBM based on ensemble learning methods for short-term electrical load forecasting," *Energy*, vol. 318, p. 134757, 2025, doi: doi.org/10.1016/j.energy.2025.134757.

[25] J. Heikkinen and K. Heimonen, "Nowcasting the nowcasting - Forecasting ISM Business surveys (PMI and NSI) with weekly Google trends," *Appl. Econ.*, vol. 56, no. 51, pp. 6300–6313, 2024, doi: 10.1080/00036846.2023.2273235.

[26] Y. Chi, L. El-Jahel, and T. Vu, "Novel and old news sentiment in commodity futures markets," *Energy Econ.*, vol. 140, p. 108006, 2024, doi: doi.org/10.1016/j.eneco.2024.108006.

[27] D. Wen, Z. Zhang, J. Nie, and Y. Cao, "Investor attention and anomalies: Evidence from the Chinese stock market," *Int. Rev. Financ. Anal.*, vol. 96, p. 103775, 2024, doi: doi.org/10.1016/j.irfa.2024.103775.

[28] F. T. Lima and V. M. A. Souza, "A Large Comparison of Normalization Methods on Time Series," *Big Data Res.*, vol. 34, p. 100407, 2023, doi: doi.org/10.1016/j.bdr.2023.100407.

[29] E. Strøm and O. E. Gundersen, "Performance metrics for multi-step forecasting measuring win-loss, seasonal variance and forecast stability: an empirical study," *Appl. Intell.*, vol. 54, no. 21, pp. 10490–10515, 2024, doi: 10.1007/s10489-024-05715-4.

[30] L. Coroneo and F. Iacone, "Testing for equal predictive accuracy with strong dependence," *Int. J. Forecast.*, vol. 41, no. 3, pp. 1073–1092, 2025, doi: doi.org/10.1016/j.ijforecast.2024.11.003.

[31] J. Pineau *et al.*, "Improving Reproducibility in Machine Learning Research," *J. Mach. Learn. Res.*, vol. 22, no. 1, pp. 1–20, 2021