A Comparative Study of GNN and LSTM for Multivariate Forecasting of Premium Rice Prices Across Indonesian Provinces Using Limited Temporal and Spatial Features

Muh. Naufal Fahri Salim, A. Tasdik Bijaksana, Muh. Aipun Pratama, Fara Rahmasari Fahirun
Hasanuddin University , Indonesia
nfahrisalim@gmail.com, bijaksanaat23h@student.unhas.ac.id, aipunpratama12@gmail.com, fararahmasari3001@gmail.com

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

The prices of food commodities in Indonesia are highly volatile, making them difficult to predict accurately. This price fluctuation poses significant challenges for the government in formulating strategic policies to mitigate price surges that could affect economic stability and public welfare. Therefore, the ability to forecast food prices for the upcoming month is crucial for enabling the government to take effective preventive measures before undesirable consequences occur.

Previous studies have mainly focused on food price prediction in other countries, with limited research addressing this issue in the Indonesian context, especially given the scarcity of local data. In this study, we propose the implementation of Long Short-Term Memory (LSTM) and Graph Neural Network (GNN) models on a limited Indonesian food price dataset. Unlike other methods that require a large number of features to achieve high prediction accuracy, our approach aims to produce accurate predictions using only a limited set of features.

The main contributions of this research are: (1) implementing and comparing the performance of LSTM and GNN models on Indonesian food price data with limited features; and (2) providing recommendations on the potential use of food price prediction models to support government policy formulation. The remainder of this paper is organized as follows: Section 2 presents the literature review, Section 3 describes the methodology, Section 4 discusses the results and analysis, and Section 5 concludes the paper.

II. Literature Review

Forecasting premium rice prices involves capturing complex interactions across time and space. In previous literature, various approaches have been proposed, ranging from classical econometric models to advanced deep learning architectures. This section summarizes recent relevant works that explore time series prediction, particularly for agricultural commodities, with a focus on rice price forecasting using LSTM and GNN.

A. Traditional Models and Their Limitations

Classical time series models such as ARIMA and VECM have long been used in commodity price forecasting. For instance, Firmansyah et al. [1] used the VECM model to analyze the integration of beef markets between producer and consumer regions in Indonesia. Although effective for understanding long-term cointegration, these models struggle with non-linear relationships and multivariate complexity. They also assume stationarity and typically cannot capture dynamic spatial dependencies or higher-order temporal interactions.

B. Deep Learning for Commodity Price Forecasting

Recurrent Neural Networks (RNN), especially Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), are widely used to model sequential data. Harshith and Kumari [2] applied several deep learning architectures—including DNN, RNN, GRU, and Stacked LSTM—to forecast cumin prices in Gujarat. Their findings showed that Stacked LSTM achieved superior results, especially during critical agricultural periods. Though the study focused on cumin, the modeling strategy and evaluation metrics are applicable to rice and similar agricultural products.

Yang [3] also demonstrated the utility of LSTM and GRU in financial time series forecasting, highlighting their robustness against volatility and complex temporal patterns. These findings support the use of deep learning in price prediction tasks, especially when nonlinearity and long-term dependencies are involved.

C. Graph Neural Networks for Spatio-Temporal Forecasting

Recent studies have introduced Graph Neural Networks (GNN) to model both temporal and spatial dependencies. Özden and Bulut [4] employed a Spectral Temporal Graph

Neural Network (StemGNN) to predict vegetable prices in Turkey. Compared to CNN, LSTM, and Random Forest models, StemGNN achieved the lowest MAE (1.37) and RMSE (1.94), demonstrating its strength in capturing intra- and inter-series correlations.

Similarly, Min et al. [5] applied GNN-based models (StemGNN and T-GCN) to forecast agricultural prices in South Korea using multivariate time series, including weather data. They showed that GNN outperformed LSTM in terms of accuracy and stability across different forecasting horizons. Moreover, they incorporated rolling average smoothing to handle short-term volatility, which is highly relevant for daily rice price data.

D. Dynamic Graph-Based Models and Market Structure

To capture evolving market structures, Kim and Park [6] proposed the STAD-GCN (Spatial-Temporal Attention-based Dynamic Graph Convolutional Network), applied to South Korea's retail gasoline market. The model dynamically updated edge weights based on attention mechanisms, allowing it to reflect changes in spatial competition. Although not tested on agricultural data, the architecture is promising for rice price forecasting across provinces where market relationships may shift over time.

E. Comparative Analysis of GNN and LSTM

Among the reviewed studies, Min et al. [5] provide a direct comparison between GNN and LSTM for forecasting agricultural commodity prices. Their experiments indicate that GNN architectures—particularly StemGNN—are better suited for handling multivariate spatio-temporal data, achieving lower forecasting errors compared to LSTM. This reinforces the potential of GNN models for tasks involving provincial rice prices, which are influenced by both geographical and seasonal factors.

F. Identified Gaps and Research Opportunities

Despite promising results, most GNN applications are limited to international datasets and other commodities like vegetables or fuel. There is a lack of studies applying GNN to forecast daily premium rice prices in Indonesia. Furthermore, many evaluations focus solely on MAE and RMSE, while metrics like MAPE and horizon-specific degradation are rarely explored. The use of dynamic graph structures, such as STAD-GCN, remains underutilized in agriculture, representing a key area for further research.

G. Summary

The reviewed literature highlights that while LSTM is a powerful tool for time series modeling, GNN models offer significant advantages in handling spatial and temporal dependencies jointly. For multivariate forecasting tasks such as predicting daily premium rice prices across Indonesian provinces, the integration of GNN with temporal smoothing and dynamic graph adaptation presents a promising direction.

III. Methodology

A. Research Design

This research adopts an experimental comparative approach to evaluate several deep learning models for food price prediction. The workflow consists of several sequential stages: (1) data scraping from official government sources, (2) data cleaning and preprocessing, (3) time series formatting and graph construction (for GNN), (4) feature engineering, (5) data splitting into training and testing sets, (6) model training and architecture experimentation, (7) hyperparameter tuning, (8) evaluation, and (9) results analysis.

B. Dataset

The primary dataset was obtained through web scraping from the official government website https://panelharga.badanpangan.go.id/beranda. The dataset contains 41,648 rows and 4 columns: date, commodity (only premium rice), province, and price. The data spans from 19 May 2022 to 16 May 2025.

C. Preprocessing and Feature Engineering

1) GNN Preprocessing: The preprocessing steps for the Graph Neural Network (GNN) model include:

- **Data Cleaning:** Checking and removing missing values.
- **Datetime Conversion:** Converting the date column to datetime format.
- **Pivoting:** Transforming the data with date as index, provinces as columns, and price as values.
- **Sorting and Interpolation:** Sorting dates in ascending order, applying linear interpolation for missing values, and forward/backward filling for edge missing data.
- **Province Indexing:** Mapping province names to indices for graph construction.
- **Graph Construction:** Defining graph edges based on geographical proximity between provinces.
- **Sequence Preparation:** Creating input sequences of 30 days per province as features and the next 30 days as labels.
- **Normalization:** Applying MinMaxScaler normalization per province.
- **Temporal Dataset:** Creating a temporal graph dataset, where each snapshot represents one day with the label being the next 30 days.
- **Additional Features:** Experimenting with added features such as week_sin, month_sin, and quarter_sin.

2) LSTM Preprocessing: Preprocessing for the LSTM model follows similar steps: data cleaning, datetime conversion, feature engineering (adding month_sin, month_cos, and moving average of price with window size 3), and normalization.

3) Data Splitting: Both models use an 80:20 time series split for training and testing to preserve temporal integrity.

D. Model Architecture

The models compared in this study include:

- **LSTM (Long Short-Term Memory):** Used for its effectiveness in sequential time series data modeling, as supported by previous literature.
- **GNN (Graph Neural Network):** Specifically, GCNConv and GAT/GRU architectures are implemented to capture spatial and temporal dependencies across provinces.
- **Feature Engineering:** Additional temporal features (e.g., sine/cosine of month, moving averages) are explored to enhance predictive capability.
- **Loss Function:** Huber loss is utilized for GNN models due to its robustness to outliers.

E. Experiment Steps

The research steps are as follows: (1) data scraping from official government sources, (2) data cleaning and preprocessing, (3) time series formatting and graph construction (for GNN), (4) feature engineering, (5) data splitting into training and testing sets, (6) model training and architecture experimentation, (7) hyperparameter tuning, (8) evaluation, and (9) results analysis.

F. Evaluation Metrics

Model performance is evaluated using several metrics:

- **Root Mean Squared Error (RMSE):**

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{1}$$

- **Mean Absolute Error (MAE):**

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

- **Mean Squared Error (MSE):**

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3}$$

- **Mean Absolute Percentage Error (MAPE):**

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

- **Coefficient of Determination (R^2 Score):**

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

G. Tools and Libraries

The implementation utilizes Python programming language with the following libraries: PyTorch, Matplotlib, NetworkX, NumPy, Pandas, and Scikit-learn.

IV. Results and Discussion

A. Experimental Results

The experimental evaluation involved two main models: a two-layer LSTM and a spatial-temporal GNN (StemGNN variant), each trained and tested on the same rice price dataset. Performance metrics were computed on the original price scale for interpretability. Table I summarizes the evaluation metrics of both models on the test set.

TABLE I
Performance Comparison between LSTM and GNN

Model	MAE	MAPE(%)	MSE	RMSE	R^2
LSTM (Initial)	740.07	5.31	918626.49	958.45	0.5012
LSTM (Final)	**319.01**	**2.28**	**265014.15**	**514.80**	**0.8561**
GNN (Initial)	596.89	3.22	1712991.87	1308.82	0.6900
GNN (Final)	451.42	3.00	299143.59	546.93	-5.9673

Improvements in the LSTM model were achieved after feature engineering (adding cyclical time features and moving averages) and architectural modifications (increasing the number of layers and using MAE as loss). For GNN, the introduction of the StemGNN block and additional engineered features also led to more reasonable predictions, although performance was still suboptimal compared to LSTM.

B. Prediction Visualization

Figures 1 and 2 show the comparison between actual prices and predicted prices for each model. The LSTM model's predicted prices closely followed the actual prices with minor deviations, indicating strong predictive capability. In contrast, the GNN model, while able to track overall trends, exhibited greater fluctuations and occasional systematic underestimation.

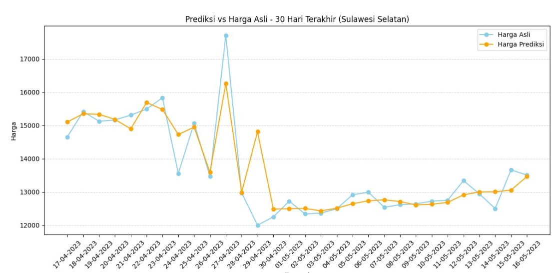

Fig. 1. LSTM: Actual vs. Predicted Rice Prices

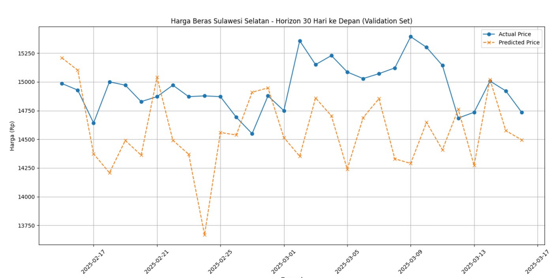

Fig. 2. GNN: Actual vs. Predicted Rice Prices (South Sulawesi, 30-day horizon)

C. Model Architecture and Feature Engineering Effects

Initial LSTM models with basic architecture (single layer) and minimal features produced less accurate results (MAE = 740.07, MAPE = 5.31%, $R^2 = 0.5012$). By applying feature engineering—including cyclical encoding of month, moving averages, and multi-layer architecture with robust loss (MAE)—the LSTM achieved significant improvements, with MAE reduced to 319.01 and R^2 increased to 0.8561. This suggests that LSTM models are highly effective for time series forecasting in cases with limited but relevant temporal features.

The initial GNN (GCNConv-based) model showed moderate performance ($R^2 = 0.69$) but the predicted values consistently overestimated actual prices, as seen in the prediction plots. Modifying the architecture to a StemGNN variant—incorporating spectral graph convolution and multiple engineered features (rolling averages, volatility, seasonality encoding)—yielded more realistic predictions, though R^2 dropped below zero, indicating inconsistency in capturing variance. The result demonstrates that, while GNNs can model spatial dependencies, their natural smoothing effect (where predictions are regularized toward the mean due to graph Laplacian smoothing) can result in under/overestimation, especially when important exogenous features (e.g., weather, policy shocks) are absent. This aligns with findings from previous studies, which suggest that GNNs benefit from richer feature sets to capture complex non-temporal influences.

D. Comparative Analysis and Interpretation

Based on quantitative metrics, the final LSTM outperformed GNN in almost all aspects. LSTM's architecture is inherently adept at modeling sequential dependencies and works well with a limited set of time-based features. GNN's spatial modeling power was underutilized in this context, likely due to the lack of exogenous spatial features (such as weather, logistics, or socio-economic variables). The smoothing property of GNNs, while beneficial for denoising, can obscure local spikes or drops in price, leading to less responsive forecasts.

E. Limitations

This study is limited by time constraints that precluded more extensive hyperparameter tuning and exploration of alternative GNN architectures or integration of additional exogenous features. Furthermore, the dataset was limited to a single commodity (premium rice), and the scope for generalization to other commodities or regions is yet to be validated.

F. Implications and Future Work

The results indicate that LSTM models are suitable for short-term price forecasting in settings with limited but high-quality temporal data. These forecasts can be used as input for policy-making, such as determining intervention timing or resource allocation. For richer spatial-temporal modeling, future research should incorporate additional exogenous features (e.g., weather, supply chain data) and experiment with more advanced or hybrid GNN architectures. This study also provides a foundation for developing early warning systems or decision support tools for food security management.

G. Evaluation Metrics Formulation

For completeness, the evaluation metrics used in this study are defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{7}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{8}$$

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{10}$$

where y_i denotes the actual price, $\hat{y}_i$ the predicted price, and $\bar{y}$ the mean of the actual prices.

V. Conclusion

This study comparatively evaluated the effectiveness of Long Short-Term Memory (LSTM) and Graph Neural Network (GNN) models for forecasting premium rice prices in Indonesia using a limited set of temporal and spatial features. The experimental results demonstrated that the two-layer LSTM model, enhanced through feature engineering and architectural refinement, consistently outperformed the GNN approach in terms of all key evaluation metrics, achieving an MAE of 319.01, MAPE of 2.28%, RMSE of 514.80, and R^2 score of 0.8561 on the test set. The LSTM model was able to closely approximate actual price trends with minimal error, making it suitable for short-term price forecasting scenarios with limited data.

In contrast, the GNN-based models, even after implementing advanced architectures and additional temporal features, exhibited less stable performance and were more sensitive to the absence of rich exogenous information. The results highlighted the natural smoothing effect of GNNs, which may lead to under- or over-estimation when critical features—such as weather or macroeconomic indicators—are missing. Moreover, the negative R^2 score in the final GNN configuration suggests that further work is needed to appropriately capture complex spatial-temporal dependencies in food price data.

The insights from this research can inform government stakeholders in developing proactive food pricing policies and highlight the value of LSTM models for operational early warning systems. For future work, we recommend the integration of additional exogenous features (e.g., weather, logistics, policy shocks) and the exploration of more sophisticated or hybrid GNN architectures. Expanding the dataset to include multiple commodities and regions would also help generalize the findings and further validate the models' applicability.

REFERENCES

[1] Firmansyah, P. Maruli, and A. Harahap, "Analysis of beef market integration between consumer and producer regions in indonesia," *Open Agriculture*, vol. 8, no. 1, p. 20220221, 2023.

[2] N. Harshith and P. Kumari, "Memory based neural network for cumin price forecasting in gujarat, india," *Journal of Agriculture and Food Research*, vol. 15, p. 101020, 2024.

[3] S. Yang, "A novel study on deep learning framework to predict and analyze the financial time series information," *Future Generation Computer Systems*, vol. 125, pp. 812–819, 2021.

[4] C. Özden and M. Bulut, "Spectral temporal graph neural network for multivariate agricultural price forecasting," *Ciência Rural*, vol. 54, no. 1, p. e20220677, 2024.

[5] Y. Min, Y. R. Kim, Y. Hyon, T. Ha, S. Lee, J. Hyun, and M. R. Lee, "Rnn and gnn based prediction of agricultural prices with multivariate time series and its short-term fluctuations smoothing effect," *Scientific Reports*, vol. 15, p. 13681, 2025.

[6] S. Kim and E. Park, "Stad-gcn: Spatial–temporal attention-based dynamic graph convolutional network for retail market price prediction," *Expert Systems With Applications*, vol. 255, p. 124553, 2024.