Hybrid ICEEMDAN–SVD–LSTM Model for Short-Term Stock Closing Price Forecasting: A Multi-Sector Evaluation on the Indonesia Stock Exchange

Yustian Faza[1], Vemmie Nastiti Lestari[2]

[1,2]Department of Mathematics, Universitas Gadjah Mada

Bulaksumur, Sleman, Daerah Istimewa Yogyakarta, 55281, Indonesia

e-mail: yustianfaza@mail.ugm.ac.id[1], vemmie.lestari@ugm.ac.id[2]

ABSTRAK

Penelitian ini mengembangkan dan mengevaluasi model hibrida ICEEMDAN–SVD–LSTM untuk meramalkan harga penutupan harian saham BBRI, UNVR, dan UNTR selama periode 15 tahun (September 2010–September 2025). Kebaruan penelitian ini terletak pada penerapan ICEEMDAN–SVD–LSTM untuk peramalan saham di Bursa Efek Indonesia yang masih jarang diteliti, serta pada perbandingan sistematis antarmodel berbasis dekomposisi dan denoising dalam satu kerangka eksperimen yang seragam pada tiga sektor dengan karakteristik volatilitas berbeda. ICEEMDAN dipilih karena menghasilkan dekomposisi yang lebih bersih, dengan mode semu (*spurious mode*) dan *noise residual* yang lebih sedikit dibandingkan EMD, EEMD, dan CEEMDAN, sementara *denoising* berbasis SVD melalui matriks Hankel meningkatkan rasio sinyal terhadap *noise* pada tiap komponen sebelum tahap peramalan LSTM. Kinerja model dibandingkan dengan tujuh model pembanding, yaitu LSTM, EMD–LSTM, EEMD–LSTM, CEEMDAN–LSTM, ICEEMDAN–LSTM, EEMD–SVD–LSTM, dan CEEMDAN–SVD–LSTM. Berdasarkan RMSE, MAE, dan MAPE, ICEEMDAN–SVD–LSTM menghasilkan eror terendah pada ketiga saham, dengan RMSE 16.4605 (BBRI), 24.3086 (UNVR), dan 70.8494 (UNTR), serta MAE dan MAPE yang juga lebih rendah dari seluruh pembanding. Uji Diebold–Mariano mengonfirmasi bahwa keunggulan tersebut signifikan secara statistik terhadap semua model pembanding. Hasil ini menunjukkan model yang diusulkan efektif untuk peramalan harga saham jangka pendek dan dapat menjadi alat bantu bagi investor maupun analis dengan tetap memperhatikan risiko pasar.

Kata kunci: Peramalan Harga Saham; ICEEMDAN; *Singular Value Decomposition*; LSTM; Runtun Waktu

ABSTRACT

This study develops and evaluates a hybrid ICEEMDAN–SVD–LSTM model for forecasting the daily closing prices of BBRI, UNVR, and UNTR over an observation period of 15 years (September 2010–September 2025). The novelty lies in applying ICEEMDAN–SVD–LSTM to stock forecasting on the still underexplored Indonesia Stock Exchange, and in systematically comparing decomposition- and denoising-based models within one uniform framework across three sectors with different volatility profiles. ICEEMDAN was chosen because it yields a cleaner decomposition, with fewer spurious modes and less residual noise than EMD, EEMD, and CEEMDAN, while Hankel-matrix-based SVD denoising increases the signal-to-noise ratio of each component before the LSTM forecasting stage. The model's performance was compared with seven benchmarks: LSTM, EMD–LSTM, EEMD–LSTM, CEEMDAN–LSTM, ICEEMDAN–LSTM, EEMD–SVD–LSTM, and CEEMDAN–SVD–LSTM. In terms of RMSE, MAE, and MAPE, ICEEMDAN–SVD–LSTM produced the lowest errors for all three stocks, with RMSE values of 16.4605 (BBRI), 24.3086 (UNVR), and 70.8494 (UNTR), and lower MAE and MAPE than every benchmark. A Diebold–Mariano test confirmed that this advantage is statistically significant against all benchmark models. These results indicate that the proposed model is effective for short-term stock price forecasting and can serve as a decision-support tool for investors and analysts, while acknowledging market risk.

Keywords: *Stock Price Forecasting; ICEEMDAN; Singular Value Decomposition; LSTM; Time Series*

INTRODUCTION

Stock price forecasting is a challenging problem because stock price data often show nonlinear and non-stationary patterns. These patterns are influenced by fundamental factors, market sentiment, and external events, causing the observed series to contain mixed-frequency components and substantial noise. Under these conditions, a single forecasting model may not be sufficient to capture the underlying signal accurately.

Long Short-Term Memory (LSTM) is widely used for sequential data because it can learn temporal dependence more effectively than conventional recurrent neural networks [1]. However, when the data contain strong noise and heterogeneous frequency components, the forecasting performance of a single LSTM model may decrease. Therefore, a decomposition stage is needed to separate the original series into simpler components before the forecasting stage.

Several decomposition methods have been developed for nonlinear and non-stationary time series. Ensemble Empirical Mode Decomposition (EEMD) was introduced to reduce the mode-mixing problem in Empirical Mode Decomposition [2]. This method was later improved into Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [3], and then into Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) [4]. Compared with EMD, EEMD, and CEEMDAN, ICEEMDAN adds noise derived from the modes of white noise rather than injecting white noise directly, which yields better mode separation, fewer spurious modes, and lower residual noise, and has therefore been adopted as a preprocessing stage in recent financial forecasting studies [5]. In addition, Singular Value Decomposition (SVD) applied to a Hankel-matrix representation can serve as a denoising stage before prediction by separating the signal subspace from the noise subspace, thereby increasing the signal-to-noise ratio of each component. Prior studies have shown that decomposition–LSTM models can outperform a single LSTM model in nonlinear time-series forecasting [6], and that hybrid frameworks combining adaptive decomposition with deep learning consistently improve accuracy for noisy financial series [5], [7]. A recent study also reported that ICEEMDAN–SVD–LSTM achieved better performance than several alternative hybrid models [8].

The need for a reliable forecasting model has become more important in the Indonesian capital market. By the end of October 2025, the number of capital market investors had reached nineteen million [9]. In addition, 54.23% of investors were under 30 years old as of August 2025 [10]. Survey results also indicate strong interest in sectors such as finance and fast-moving consumer goods [11]. Deep-learning models such as LSTM and GRU have been applied to price forecasting on the Indonesia Stock Exchange (IDX) [12], [13], yet a consensus on the most suitable model has not been reached, and these studies have generally been conducted without a decomposition or denoising stage [12]. Based on these considerations, this study focuses on three strategic sectors that are among those most favored by young investors and that display distinctly different volatility characteristics: banking, fast-moving consumer goods (FMCG), and industry. Banking sits at the center of the national financial system and is sensitive to monetary policy, benchmark interest rates, and credit quality. FMCG reflects domestic consumption patterns and is exposed to inflation and shifts in consumer behavior, whereas industry is driven by commodity prices, global trade policy, and international supply chains. Selecting sectors with such different dynamics allows the proposed model to be evaluated across a broad range of volatility conditions within a single framework. Each sector is represented by one of its most prominent and actively traded companies on the Indonesia Stock Exchange, namely PT Bank Rakyat Indonesia (Persero) Tbk (BBRI), PT Unilever Indonesia Tbk (UNVR), and PT United Tractors Tbk (UNTR).

Despite this relevance, the application of ICEEMDAN–SVD–LSTM to stock closing-price forecasting on the Indonesia Stock Exchange remains limited, and existing work leaves three gaps. First, prior ICEEMDAN–SVD–LSTM studies have mainly demonstrated the framework on generic nonlinear series rather than on exchange-listed equities [8], while recent ICEEMDAN-based financial studies have paired the decomposition with other predictors, such as Transformer architectures, instead of LSTM [5]. Second, deep-learning forecasting on the IDX has largely been carried out without a decomposition or denoising stage, and no consensus on the best model has been reached [12], [13]. Third, few studies provide a systematic, single-framework comparison across multiple decomposition-based and denoising-based hybrid models, evaluated across sectors with different volatility characteristics and supported by a formal test of statistical significance. This study addresses these gaps rather than proposing a new algorithm. Its contribution is a controlled comparison of eight decomposition × denoising variants under identical settings, evaluated across three IDX sectors with distinct volatility profiles (banking, FMCG, and industry), with the accuracy differences validated statistically using the Diebold–Mariano test and all settings and random seeds fully reported to support reproducibility. Accordingly, this study develops and evaluates a hybrid ICEEMDAN–SVD–LSTM model for forecasting stock closing prices and compares its performance with seven decomposition- and denoising-based benchmarks using data from the banking, FMCG, and industrial sectors.

METHOD

Data and Study Design

This study used a case study design with secondary data in the form of daily stock closing prices from three companies listed on the Indonesia Stock Exchange, namely PT Bank Rakyat Indonesia (Persero) Tbk (BBRI), PT Unilever Indonesia Tbk (UNVR), and PT United Tractors Tbk (UNTR). The data represent the banking, fast-moving consumer goods, and industrial sectors, respectively. The study only used the closing price variable, so the forecasting problem was treated as a univariate time-series problem. The observation period spanned approximately 15 years and differed slightly across the three stocks: BBRI covered 27 September 2010 to 25 September 2025 (3,692 observations), while UNVR and UNTR both covered 29 September 2010 to 26 September 2025, with 3,691 and 3,690 observations, respectively. These small differences in the starting dates and the number of observations arise from differences in the availability of historical data for each stock at the time of data collection.

Preliminary Data Analysis

Before model construction, the data were examined for stationarity and linearity. Stationarity was tested using the Augmented Dickey-Fuller (ADF) test [14], while linearity was examined using the Tsay test [15]. These tests were used to confirm the characteristics of the stock closing-price series before the forecasting stage.

Data Preprocessing

The data were first transformed using Min-Max normalization so that all observations were scaled to the interval ([0,1]). For an observed value x_t, the normalized value z_t is defined as

$$z_t = \frac{x_t - x_{min}}{x_{max} - x_{min}}$$

After forecasting, the predicted values were transformed back to the original scale using the inverse Min-Max transformation.

The dataset was then divided into training and testing sets using an 80:20 chronological split, so that the earliest 80% of each series (by date) formed the training set and the most recent 20% formed the testing set. Because the data are time series, the split preserved the original time order of the observations, with no shuffling across the split boundary; the model was therefore trained only on earlier observations and evaluated on strictly later ones, preventing any leakage of future information into training.

ICEEMDAN–SVD–LSTM Procedure

The proposed model combines decomposition, denoising, and forecasting stages. First, the normalized series was decomposed using Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) [4]. This procedure decomposes the original series into several Intrinsic Mode Functions (IMFs) and one residual component, which can be written as

$$x_t = \sum_{i=1}^{M} d_{i,t} + r_t,$$

where $d_{i,t}$ denotes the i-th IMF and r_t denotes the residual component.

The decomposition was configured with an ensemble size of I = 800 realizations, a noise amplitude of α = 0.1 relative to the standard deviation of the residual at each stage, and a maximum of Kmax = 8 IMF levels. Decomposition was terminated when the standard deviation of the residual fell below 1×10^{-8}, or otherwise at Kmax.

Second, each IMF was denoised using Singular Value Decomposition (SVD) through Hankel-matrix representation [16]. For a Hankel matrix H_i formed from the i-th component, the decomposition is written as

$$H_i = U_i \Sigma_i V_i^{\top},$$

where U_i and V_i are orthogonal matrices and Σ_i is the diagonal matrix of singular values. In this study, the noise subspace was identified and removed using an energy-based hard-thresholding criterion. Each IMF of length N was first embedded into a Hankel trajectory matrix $\mathbf{H_i}$ of dimension $L \times K$, where $L = \lfloor N/2 \rfloor + 1$ and $K = N - L + 1$. After applying SVD to H_i, the resulting singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ were ranked in descending order. The number of retained singular values k was determined by the following criterion:

$$k = \min\left\{k' : \frac{\sum_{j=1}^{k'} \sigma_j^2}{\sum_{j=1}^{r} \sigma_j^2} \geq 0.98\right\}$$

That is, the smallest k such that the retained singular values account for at least 98% of the total spectral energy of the component. Singular values with index $j > k$ were treated as noise and set to zero. The denoised low-rank approximation matrix was then reconstructed as $\widetilde{H_i} = U_i \widetilde{\Sigma_i} V_i^{\top}$, where $\widetilde{\Sigma_i}$ contains only the k retained singular values. The denoised time-series component was recovered from $\widetilde{H_i}$ by anti-diagonal averaging [16]. This procedure was applied independently to each IMF; the residual component was not subjected to SVD denoising, as it captures the low-frequency trend of the original series and is not expected to contain high-frequency noise.

Third, each denoised IMF and the residual component were forecast separately using Long Short-Term Memory (LSTM) [1]. The input data for each LSTM model were constructed using a sliding-window approach with lookback window $p = 10$. The hyperparameter configuration (lookback window p, hidden size, and learning rate) followed prior decomposition–LSTM studies [8] and was applied identically to all models, so that the performance differences in Table 2 reflect model structure rather than tuning effort. No per-model tuning was performed; this is acknowledged as a limitation, since every variant, including the proposed model, could improve with model-specific tuning. The network architecture consisted of one LSTM layer with input size 1 and hidden size 50, followed by one fully connected output layer. The model was trained using Mean Squared Error (MSE) loss and the Adam optimizer [17] with learning rate 0.001. Training was conducted for 100 epochs with batch size 32 and gradient clipping of 1.0.

After forecasting each component, the final forecast was obtained by summing all predicted IMFs and the predicted residual:

$$\hat{x}_{(t+h)} = \sum_{i=1}^{M} \hat{d}_{(i,t+h)} + \hat{r}_{(t+h)},$$

where h denotes the forecasting horizon.

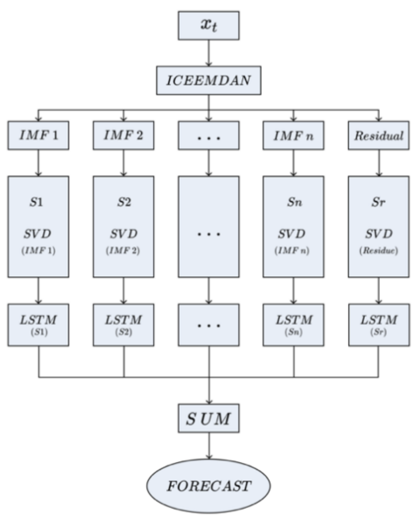

Figure 1. The proposed ICEEMDAN–SVD–LSTM forecasting framework

Model Comparison and Evaluation

To assess the effectiveness of the proposed model, ICEEMDAN–SVD–LSTM was compared with seven alternative models, namely LSTM, EMD–LSTM, EEMD–LSTM, CEEMDAN–LSTM, ICEEMDAN–LSTM, EEMD–SVD–LSTM, and CEEMDAN–SVD–LSTM. All models were evaluated on the testing set using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These measures are defined as follows [18]:

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{x}_t)^2}, \qquad \text{MAE} = \frac{1}{T}\sum_{t=1}^{T}|x_t - \hat{x}_t|, \qquad \text{MAPE} = \frac{100}{T}\sum_{t=1}^{T}\left|\frac{x_t - \hat{x}_t}{x_t}\right|.$$

The model with the smallest forecasting errors was selected as the best model and was then used to generate three-day-ahead forecasts.

To assess whether differences in forecasting accuracy between the proposed model and each benchmark are statistically significant rather than attributable to random variation, the Diebold–Mariano (DM) test with the Harvey–Leybourne–Newbold small-sample correction was applied at forecast horizon h = 1, using both squared-error and absolute-error loss differentials. The null hypothesis of equal predictive accuracy is rejected in favor of the proposed model when the DM statistic is negative and the associated p-value is below 0.05.

Implementation and Reproducibility

All experiments were implemented in Python 3.9.6 on macOS 14.4.1 (arm64). Although an Apple Silicon MPS accelerator was available, all computations were executed on CPU with deterministic algorithms enabled to ensure exact reproducibility of results. A global random seed of 42 was fixed for all data splitting and model-initialization operations, and a separate seed of 123 was fixed for the noise-assisted ICEEMDAN and EEMD decomposition procedures. Key package versions were: NumPy 2.0.2, pandas 2.3.3, SciPy 1.13.1, scikit-learn 1.5.1, statsmodels 0.14.6, PyTorch 2.8.0, and PyEMD 1.6.4.

On the CPU-based deterministic setup described above, training the proposed ICEEMDAN–SVD–LSTM model took under three minutes per stock (162.9 s on average), and the full eight-variant comparison across the three stocks completed in 77.4 minutes, confirming that the decomposition and denoising overhead remains practical for daily short-term forecasting.

RESULT AND DISCUSSION

Preliminary Analysis

To confirm the characteristics of the data, stationarity and linearity tests were conducted. The results are presented in Table 1. All three series are non-stationary at level based on the Augmented Dickey-Fuller test and nonlinear based on the Tsay test. Therefore, the three stock-price series are suitable for analysis using a hybrid model designed for nonlinear and non-stationary time-series data.

Table 1. Summary of stationarity and linearity tests

Stock	Test	p-value	Conclusion
BBRI	ADF (stationarity)	0.764276	Non-stationary at level
BBRI	Tsay (linearity)	0.000000	Reject linearity (nonlinearity detected)
UNVR	ADF (stationarity)	0.704254	Non-stationary at level
UNVR	Tsay (linearity)	0.000000	Reject linearity (nonlinearity detected)
UNTR	ADF (stationarity)	0.748950	Non-stationary at level
UNTR	Tsay (linearity)	0.006200	Reject linearity (nonlinearity detected)

After Min-Max normalization, each stock-price series was decomposed using ICEEMDAN into several IMF components and one residual component. The early IMFs represent higher-frequency movements, while the later IMFs and the residual represent lower-frequency movements and trend information. Each component was then denoised using Hankel-based SVD before the forecasting stage with LSTM.

Comparison of Forecasting Performance

The comparison results on the testing set are presented in Table 2.

Table 2. Comparison of model performance on the testing set

Stock	Model	RMSE	MAE	MAPE (%)
BBRI	ICEEMDAN–SVD–LSTM	16.4605	13.6240	0.3216
BBRI	ICEEMDAN–LSTM	19.0338	14.5238	0.3422
BBRI	EEMD–SVD–LSTM	30.4936	23.8628	0.5684
BBRI	EEMD–LSTM	33.0354	26.6762	0.6363
BBRI	CEEMDAN–LSTM	41.0700	30.5890	0.7360
BBRI	EMD–LSTM	44.1885	33.2335	0.7892
BBRI	CEEMDAN–SVD–LSTM	66.1082	51.7076	1.2223
BBRI	LSTM	82.0406	62.7367	1.4671
UNVR	ICEEMDAN–SVD–LSTM	24.3086	22.1793	0.8118
UNVR	CEEMDAN–SVD–LSTM	30.4535	24.4663	0.9137
UNVR	EMD–LSTM	31.0453	24.1239	0.9511
UNVR	EEMD–SVD–LSTM	36.7870	31.4631	1.2553
UNVR	ICEEMDAN–LSTM	39.7518	37.3209	1.4794
UNVR	EEMD–LSTM	43.1318	34.5046	1.4377
UNVR	CEEMDAN–LSTM	50.0817	44.1058	1.7674
UNVR	LSTM	68.4451	49.4002	1.9957
UNTR	ICEEMDAN–SVD–LSTM	70.8494	53.1826	0.2479
UNTR	EEMD–SVD–LSTM	83.5485	62.9207	0.2944
UNTR	EEMD–LSTM	158.1520	120.6963	0.5678
UNTR	ICEEMDAN–LSTM	158.6192	114.4342	0.5306
UNTR	CEEMDAN–SVD–LSTM	188.2239	136.7109	0.6382
UNTR	CEEMDAN–LSTM	217.0705	151.1222	0.7082
UNTR	EMD–LSTM	225.7032	162.8586	0.7633
UNTR	LSTM	522.5877	387.4158	1.7637

Based on Table 2, ICEEMDAN–SVD–LSTM achieved the lowest RMSE, MAE, and MAPE across all three stocks, indicating the best performance among the compared models.

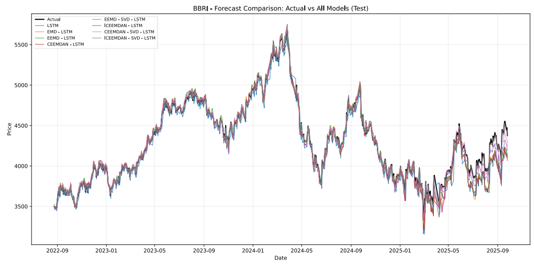

Figure 2. Actual-vs-predicted plots for BBRI

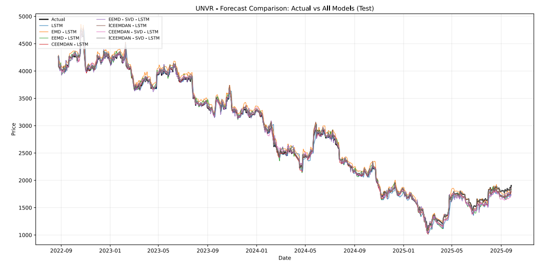

Figure 3. Actual-vs-predicted plots for UNVR

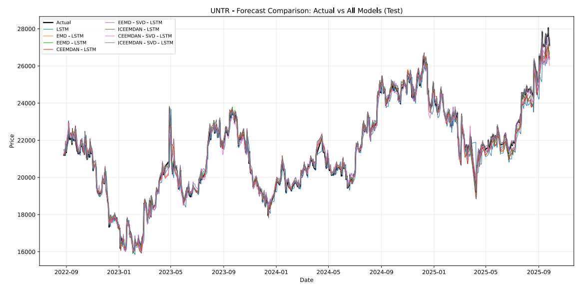

Figure 4. Actual-vs-predicted plots for UNTR

Diebold-Mariano Test

The Diebold–Mariano test described in the Method section was applied to each benchmark. For every one of the seven benchmark models and for all three stocks, the null hypothesis of equal predictive accuracy was rejected in favor of the proposed model ($p < 0.05$ in all cases). The DM statistics were consistently negative, ranging from −8.66 to −15.65, which indicates lower forecasting loss for ICEEMDAN–SVD–LSTM, and the same conclusion held under both loss functions. These results confirm that the superiority of the proposed model over each benchmark is statistically significant across all three stocks.

Three-Day-Ahead Forecasting

Because ICEEMDAN–SVD–LSTM showed the best performance, it was selected for three-day-ahead forecasting. The forecasting results are presented in Table 3.

Table 3. Three-day-ahead stock price forecasts

Stock	Date	Forecasted Price
BBRI	26/09/2025	4478.5867
BBRI	29/09/2025	4461.9314
BBRI	30/09/2025	4440.5532
UNVR	29/09/2025	1786.2141
UNVR	30/09/2025	1827.4688
UNVR	01/10/2025	1825.8862
UNTR	29/09/2025	27666.9234
UNTR	30/09/2025	27507.1132
UNTR	01/10/2025	27258.9925

The three-day-ahead forecasts show moderate short-term movements without extreme fluctuations for all three stocks.

Interpretation of Results

The substantial performance gap between standalone LSTM and decomposition-augmented models confirms that signal preprocessing substantially improves deep learning-based forecasting for noisy, non-stationary stock price series. For BBRI, the MAPE of standalone LSTM (1.4671%) was reduced to 0.3216% by the proposed model — a relative improvement of 78.1%. For UNVR and UNTR, the reductions were 59.3% and 85.9%, respectively. These findings are consistent with Hao et al. [6], who demonstrated that EMD–LSTM outperformed standalone LSTM in non-stationary wave forecasting by enabling the network to model each frequency component independently.

The incremental contribution of the SVD denoising stage is evident from a direct comparison between models with and without SVD. For BBRI, adding SVD to ICEEMDAN–LSTM reduced RMSE from 19.0338 to 16.4605, a 13.5% improvement. For the proposed ICEEMDAN-based model, SVD denoising improved accuracy for all three stocks, and the EEMD-based variant showed the same pattern, with EEMD–SVD–LSTM outperforming EEMD–LSTM for BBRI, UNVR, and UNTR. For the CEEMDAN-based variant, SVD denoising improved accuracy for UNVR and UNTR but not for BBRI, where CEEMDAN–SVD–LSTM (RMSE 66.1082) performed worse than CEEMDAN–LSTM (RMSE 41.0700). This indicates that the benefit of SVD denoising, while consistent for the ICEEMDAN and EEMD decompositions, is not universal across all decomposition methods. This finding is broadly consistent with Poongadan and Lineesh [8], who reported benefits from combining ICEEMDAN with SVD denoising in nonlinear time-series prediction.

Regarding cross-stock differences, the absolute RMSE for UNTR (70.8494) is substantially larger than for BBRI (16.4605) and UNVR (24.3086). This difference is primarily attributable to UNTR's higher absolute price level, which amplifies absolute error metrics. On a MAPE basis, which is scale-free, UNTR (0.2479%) and BBRI (0.3216%) attain the lowest values, while UNVR (0.8118%) is higher. The higher MAPE for UNVR may reflect the structural level change visible in its price series during the observation period, which increases residual non-stationarity within certain IMF components.

CONCLUSION

For BBRI, UNVR, and UNTR, the model consistently produced the smallest RMSE, MAE, and MAPE values on the testing set, and the Diebold–Mariano test confirmed that its advantage over every benchmark is statistically significant. These findings indicate that the combination of adaptive decomposition, denoising, and component-wise LSTM forecasting is more effective than forecasting the original stock-price series directly.

Based on the model comparison results, ICEEMDAN–SVD–LSTM was selected as the best model and then used to generate three-day-ahead forecasts for BBRI, UNVR, and UNTR. The resulting forecasts show moderate short-term movements without extreme fluctuations.

From a practical standpoint, accurate short-term forecasts of this kind can support financial decision-making in informing the timing of entry and exit decisions or in monitoring portfolio positions. Such applications should account for transaction costs and market risk, and the model is intended as a decision-support tool rather than a substitute for professional financial advice.

This study is limited to univariate daily closing prices from three representative stocks. Future research may evaluate the proposed model using longer forecasting horizons, additional stocks or sectors, and broader model settings. In addition, the framework may be extended to a multivariate setting by incorporating macroeconomic indicators, technical indicators, and market sentiment, and alternative sequence models such as GRU or Transformer-based architectures may be explored within the same decomposition-based framework.

REFERENCE

[1] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[2] Z. Wu and N. E. Huang, "ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOISE-ASSISTED DATA ANALYSIS METHOD," *Adv. Adapt. Data Anal.*, vol. 01, no. 01, pp. 1–41, Jan. 2009, doi: 10.1142/S1793536909000047.

[3] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic: IEEE, May 2011, pp. 4144–4147. doi: 10.1109/ICASSP.2011.5947265.

[4] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014, doi: 10.1016/j.bspc.2014.06.009.

[5] G. Kalaiarasan, B. Premjith, and N. Madhu Manjiparambil, "A Forecast-Driven Deep Learning Framework Using ICEEMDAN and PatchTST for Stock Prediction and Dynamic Portfolio Optimization," *IEEE Access*, vol. 13, pp. 218099–218122, 2025, doi: 10.1109/ACCESS.2025.3648786.

[6] W. Hao, X. Sun, C. Wang, H. Chen, and L. Huang, "A hybrid EMD-LSTM model for non-stationary wave prediction in offshore China," *Ocean Eng.*, vol. 246, p. 110566, Feb. 2022, doi: 10.1016/j.oceaneng.2022.110566.

[7] S. Li, G. Tang, X. Chen, and T. Lin, "Stock Index Forecasting Using a Novel Integrated Model Based on CEEMDAN and TCN-GRU-CBAM," *IEEE Access*, vol. 12, pp. 122524–122543, 2024, doi: 10.1109/ACCESS.2024.3452426.

[8] S. Poongadan and M. C. Lineesh, "Non-linear Time Series Prediction using Improved CEEMDAN, SVD and LSTM," *Neural Process. Lett.*, vol. 56, May 2024, doi: 10.1007/s11063-024-11622-z.

[9] "19 Juta Investor Pasar Modal dan 8 Juta Investor Saham Tercapai di Penutupan Bulan Inklusi Keuangan 2025," suit-baze. Accessed: Jun. 23, 2026. [Online]. Available: https://www.idx.co.id/

[10] "PNM Investment Management." Accessed: Jun. 23, 2026. [Online]. Available: https://www.pnmim.com/detailberita/jumlah-investor-pasar-modal-ri-tembus-18-juta-dan-didominasi-anak-muda

[11] databoks.katadata.co.id, "Deretan Sektor Saham yang Paling Disukai Anak Muda, Apa Saja? | Pusat Data Ekonomi dan Bisnis Indonesia | Databoks." Accessed: Jun. 23, 2026. [Online]. Available: https://databoks.katadata.co.id/ekonomi-makro/statistik/dcc2789e87ea5d9/deretan-sektor-saham-yang-paling-disukai-anak-muda-apa-saja

[12] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Stock price forecasting in Indonesia stock exchange using deep learning: a comparative study," *Int. J. Electr. Comput. Eng. IJECE*, vol. 14, no. 1, p. 861, Feb. 2024, doi: 10.11591/ijece.v14i1.pp861-869.

[13] B. Y. Dwiandiyanta, R. Hartanto, and R. Ferdiana, "Optimization of Stock Predictions on Indonesia Stock Exchange: A New Hybrid Deep Learning Method," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 1, pp. 19370–19379, Feb. 2025, doi: 10.48084/etasr.9363.

[14] D. Dickey and W. Fuller, "Distribution of the Estimators for Autoregressive Time Series With a Unit Root," *JASA J. Am. Stat. Assoc.*, vol. 74, Jun. 1979, doi: 10.2307/2286348.

[15] R. S. Tsay, *Nonlinear time series analysis*. Hoboken: John Wiley & Sons, 2019.

[16] H. Hassani, R. Mahmoudvand, and M. Zokaei, "Separability and window length in singular spectrum analysis," *Comptes Rendus Mathématique*, vol. 349, no. 17–18, pp. 987–990, Aug. 2011, doi: 10.1016/j.crma.2011.07.012.

[17] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Jan. 30, 2017, *arXiv*: arXiv:1412.6980. doi: 10.48550/arXiv.1412.6980.

[18] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006, doi: 10.1016/j.ijforecast.2006.03.001.