Optimization of Mineral Fuel Export Forecasting Using Attention-based Long Short-Term Memory

Ananda Prasetya[1*], Jatmiko Endro Suseno[2], Sutikno[3]

[1]Master of Information Systems, Universitas Diponegoro, Indonesia
[2] Department of Physics, Universitas Diponegoro, Indonesia
[3] Department of Informatics, Universitas Diponegoro, Indonesia

Abstract.

Purpose: This study aims to optimize the forecasting of the Net Value of Indonesia's mineral fuel exports using the Attention-based Long Short-Term Memory (LSTM) model, supported by Dropout and Recurrent Dropout techniques that are combined to produce an optimal model.

Methods: Modeling uses an LSTM architecture equipped with an Attention mechanism, as well as Dropout and Recurrent Dropout. The research procedure uses the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. The research material used is the Indonesian mineral fuel export dataset with HS code 27 from 2014 to 2025. Model was built using the Random Search method to optimize hyperparameters such as the number of neurons (units), activation functions (Tanh, ReLu), and optimizers (Adam, Nadam, RMSprop).

Result: The Attention-based LSTM model with Dropout and Recurrent Dropout techniques achieved a MAPE of 7.76%, which was better than the other models tested. Attention analysis shows that lag 12 has the greatest dominance, while lags 11 to 10 also contribute significantly, indicating an annual seasonal pattern. Projections for the next 12 months show a moderate decline in Net Value, in line with seasonal trends and historical data.

Novelty: The main contribution of this research is the optimization of an Attention-based LSTM model using a combination of Dropout and Recurrent Dropout techniques, which is effective in forecasting Indonesia's mineral fuel export values because it is able to capture annual seasonal patterns, thereby improving the accuracy and stability of the forecast results.

Keywords: Deep Learning, LSTM, Attention Mechanism, Dropout, Recurrent Dropout.
Received December 2025 / **Revised** December 2025 / **Accepted** January 2026

INTRODUCTION

Indonesia is known to have three leading export commodities in the global market, namely mineral fuels, animal/vegetable fats and oils, and iron and steel [1]. The mineral fuel [HS 27] products industry has become the second largest sector in international trade, with Russia, the United States, and Canada as the main exporters in 2021 [2]. In the context of Indonesia's mineral fuel exports, the data shows that Indonesia faces significant challenges competing in the global market. Therefore, to overcome challenges such as limited market information, marketing barriers, strict regulations, and price fluctuations that arise during the expansion process, this sector could consider implementing a deep learning platform to help improve operational efficiency and competitiveness [3]. This is why accurate export forecasting is so important in planning strategies and formulating policies for effective decision-making. However, forecasting mineral fuel [HS 27] export data remains difficult because it exhibits complex patterns with large numerical values that are influenced by global energy price trends. Based on existing theory, conventional statistical models such as Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) are often unable to understand these complex patterns, especially when there are structural changes and instability in large amounts of data. Recent studies show that the use of deep learning, particularly Long Short-Term Memory (LSTM) networks, has proven to be more effective in modeling non-linear and long-term dependencies in complex and large multivariate time series data [4], [5], [6]. This model is also capable of understanding complex and intricate time patterns that cannot be solved by traditional statistical models. This is because LSTM is able to learn long-term and non-linear

---

relationships in data directly, while ARIMA and SARIMA are only capable of capturing linear and seasonal patterns.

Although the LSTM model has been proven to be superior to traditional statistical methods based on the above research, this LSTM model still has shortcomings, especially in terms of its ability to specifically recognize the parts of historical data that most influence the prediction results [7]. In practice, LSTM treats all historical data equally, without distinguishing which information is more important. As a result, prediction effectiveness may decrease, especially if important patterns in the time series are unevenly distributed. To overcome these limitations, several previous studies have discussed the application of various methods to improve forecasting accuracy, particularly for complex time series data using a combination of LSTM and Attention mechanisms. Previous studies applying Attention-based LSTM models have shown improved prediction accuracy in financial and multivariate time series datasets [8] [9] [10]. However, most existing studies focus on financial markets or general multivariate datasets, while the application of Attention-based LSTM models for export forecasting remains limited.

Based on these findings, it can be concluded that there are still research gaps that need to be explored further, particularly regarding the application of Attention-based LSTM in the context of Indonesian exports. This gap highlights a significant weakness in existing export forecasting research. First, previous studies have not sufficiently addressed the need for models that are clearly explainable, highly accurate, and suitable for mineral fuel [HS 27] export data in the context of unstable and complex global energy conditions. Second, efforts to address the problem of overfitting that often occurs in deep learning models when used for economic time series data have not been given sufficient Attention. Therefore, this study offers an LSTM model that uses Attention, supported by Dropout and Recurrent Dropout techniques to improve optimal accuracy and keep the model stable to prevent overfitting. The combination of Dropout and Recurrent Dropout techniques serves as a way to improve the consistency and accuracy of prediction results [11], [12], [13], [14]. Both techniques work as a way to regulate the model by introducing random noise into the connections between neurons and recurrent connections during the training process.

There are three main contributions in this study. First, this study uses an LSTM model equipped with an Attention mechanism, addressing the gaps and complexities in mineral fuel [15], dataset forecasting research and identifying influential historical periods to provide practical insights for policymakers. Second, this study applies Dropout and Recurrent Dropout simultaneously, which helps make the model more stable, reduces overfitting, and produces more accurate results compared to traditional methods such as SARIMA and Prophet. Third, this study finds the best hyperparameter settings that can provide excellent mineral fuel export prediction results.

METHODS
Research Framework
The research framework uses the CRISP-DM approach, which consists of six main stages, namely Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment [16]. The framework can be seen in Figure 1. CRISP-DM was used in this study as a methodological framework that follows a goal-oriented business approach. It is a mature approach that continues to be widely accepted in data mining projects through data machine learning mining algorithms [17]. (1) Business Understanding begins with comprehensively understanding the background, context, and business needs. This business alignment is crucial to ensure that the final results are relevant and have a real impact [18]. (2) Data Understanding consists of data collection and data exploration processes aimed at examining data characteristics, temporal behavior, and potential anomalies.
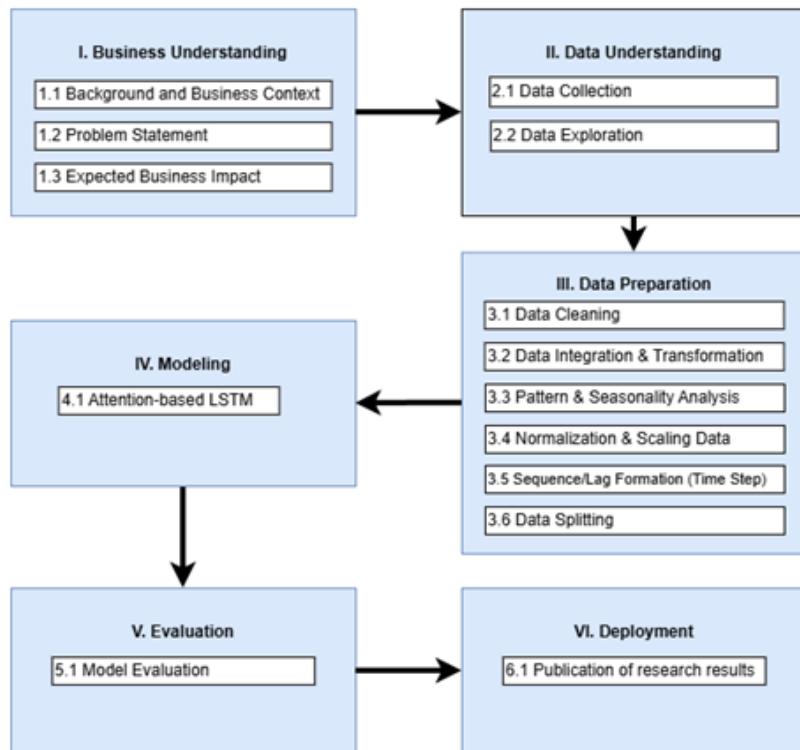
Figure 1. CRISP-DM framework stages

(3) Data Preparation, also known as the Data Preprocessing stage, aims to prepare the final dataset ready for modeling. Raw dataset was cleaned by removing incomplete records and converting all variables to numeric values. A combined monthly time index was created, Net Values and Net Weights were summed each month to reduce noise while maintaining long-term patterns. Extreme values were handled implicitly through normalization. Variables were selected based on their economic significance and correlation tests. All variables were converted using Min–Max normalization to ensure stable model learning. Time-series data was created using a 12-month sliding window method to capture temporal and seasonal relationships. The dataset was sequentially divided into 80% for training and 20% for testing to preserve the time order and prevent data leakage. (4) Modeling implements an Attention-based LSTM architecture. (5) Evaluation assesses predictive performance using statistical error metrics. (6) Deployment aims to publish research results to aid decision-making.

Dataset
The raw dataset consists of two main datasets from Indonesian Badan Pusat Statistik (BPS), which are Net Value and Net Weight from the period 2014-2025. These two datasets contain export values in USD and net export weights in kilograms for Mineral Fuel commodities with 2-Digit HS code 27. The Net Value dataset serves as the main variable in the forecasting process, while Net Weight is used as a supporting variable. The Net Value and Net Weight datasets display export values based on destination country, port of origin, HS code, and month and year of transaction. This data consists of 56,295 records, covering 128 destination countries, and a number of metadata that must be processed by cleaning the format before analysis.

Attention-based LSTM Architecture
The technical architecture of the Attention-based LSTM model can be seen in detail in Figure 2. The multivariate input sequence ($T \times F$) is processed by the LSTM layer to produce hidden states ($T \times H$). The main module calculates the Attention value, then the results are entered into the LSTM for analysis and forecasting. The output is the forecast result based on the data that has been processed by the LSTM and Attention mechanism.

In the integration stage of LSTM and the Attention mechanism, the context vector c_t (output from the Attention mechanism) is combined with the hidden state h_t (output from LSTM) to produce a new or

combined representation. This combination process is performed using the Tanh activation function, resulting in h̃_t, which contains a combination of temporal information from the LSTM and Attention Weights from the Attention mechanism. This can be seen in the following Equation (1) [19].

$$\tilde{h}_t = tanh(W_c[h_t; c_t] + b_c) \tag{1}$$

Next, the integration results are used in the output layer to generate time series forecasts. The prediction yˆ_t is obtained by multiplying the combined representation h̃_t by the weight W_y, then adding the bias b_y. In this way, the model is able to utilize both the long-term information from the LSTM and the emphasis on relevant information from Attention to produce more accurate outputs. This can be seen in the following Equation (2) [19].

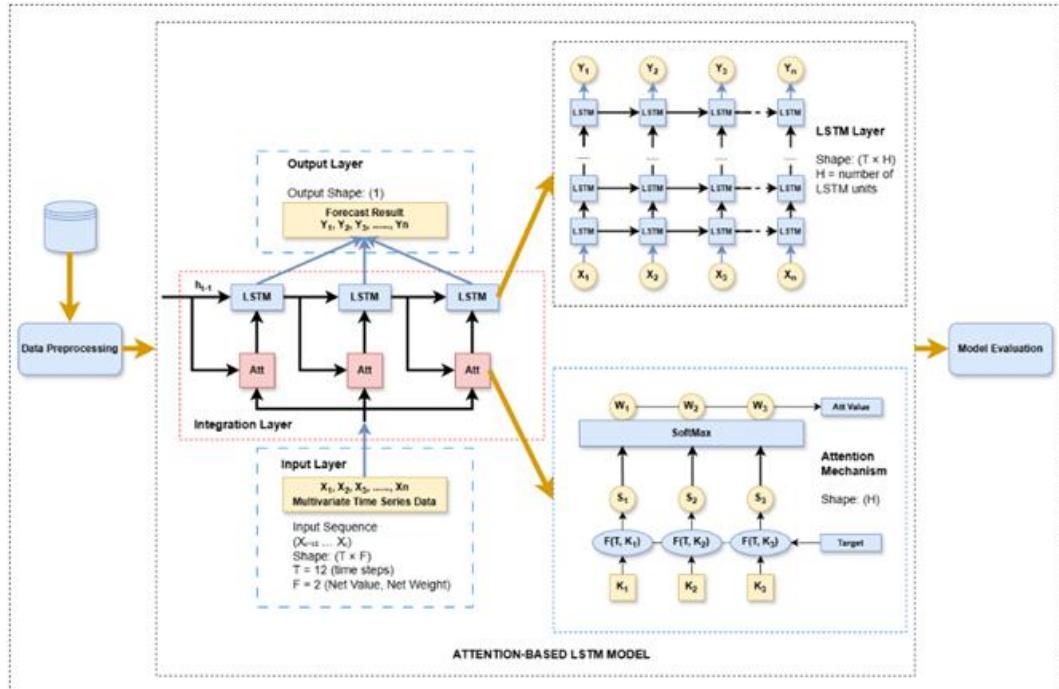$$\hat{y}_t = W_y \tilde{h}_t + b_y \tag{2}$$

Figure 2. Attention-based LSTM model architecture

Hyperparameter Optimization

The selected or combined hyperparameters are tuned using the Random Search method with a total of 10 randomized trials on the number of neuron units, activation functions (Tanh, ReLu), optimizers (Adam, Nadam, RMSprop), Dropout, and Recurrent Dropout as shown in Table 1. The number of trials is limited to 10 based on computational efficiency considerations, but this limitation still allows for sufficiently broad exploration of the hyperparameter space, thereby optimizing the search and improving validation performance efficiently [20]. The model was trained using a hold-out validation strategy by chronologically dividing the dataset into a training set (80%) and a test set (20%), while maintaining the temporal order between the training data and the unseen data [21]. Each parameter setting is performed by calculating the loss on fixed test data, and the early stopping method with a limit of 20 epochs is used to prevent overfitting [22].

The number of units or neurons combined ranges from 32-256 with an increment (step) of 32 for each experiment. The model is trained for up to 200 epochs with a batch size of 16. In addition, the time step structure is set at 12 months, which represents 12 lagged inputs from the previous period to capture temporal patterns in export data. The best combination is selected based on the lowest validation loss value on 20% of the test data to ensure optimal predictive performance and prevent overfitting [23]. Meanwhile, Dropout is used with values between 0.1 and 0.4 to remove some units from the input during training, thereby reducing the possibility of the model overfitting the data [11]. Furthermore, the model also uses Recurrent

Dropout with a value between 0.0 and 0.3, which is applied to internal connections between hidden LSTM states [24]. The specific range of Dropout and Recurrent Dropout was selected based on actual experiments in which excessively high dropout rates actually disrupted the sequential learning process [25]. This mechanism helps maintain the stability of the learning process, especially in time series data that has long-term relationships.

Table 1. Attention-based LSTM tuning

Metric	Value
Validation Loss (Best)	mse (Mean Squared Error)
LSTM Units (Neurons)	32 – 256 (step 32)
Activation Function	Tanh, ReLU
Optimizer	Adam, Nadam, RMSprop
Epochs	200
batch_size	16
max_trials	10
Validation Split	(x_test, y_test)
Early Stopping	patience = 20
time_step	12
Lagged Input	Lag 1, Lag 2, …, Lag 12
Dense Units	8 – 64 (step 8)
Dropout	0.1 – 0.4 (step 0.1)
Recurrent Dropout	0.0 – 0.3 (step 0.1)

RESULT AND DISCUSSION
Business Understanding
Mineral fuel is one of Indonesia's leading export commodities in the global market[26], [27], [1]. However, fluctuations in the export value of mineral fuel commodities are difficult to forecast due to rapid changes in international market trends. Therefore, a deep learning-based forecasting method is needed that is capable of understanding long-term patterns and highlighting the most influential time periods to improve the accuracy of export, reduce market-related risks, and assist in data-driven decision-making [15], [28].

Data Understanding
Data understanding is conducted to gain an in-depth understanding of the raw dataset used to determine whether the data is appropriate and relevant to the research objectives. The process includes data collection and data exploration [29]. Data collection stage was carried out by analyzing two main datasets from Badan Pusat Statistik Indonesia (BPS). Data exploration stage provides a comprehensive overview of the data dynamics before moving on to the next stage, which can be seen in Figure 3. The correlation heatmap shows a strong positive linear relationship between Net Value and Net Weight (r = 0.69) [30], while the relationship between the (Net Value and Net Weight) and the Year variable is very weak, indicating that export dynamics are more influenced by global market conditions than by Year factors. This pattern shows that an increase in Net Weight is generally followed by an increase in Net Value.
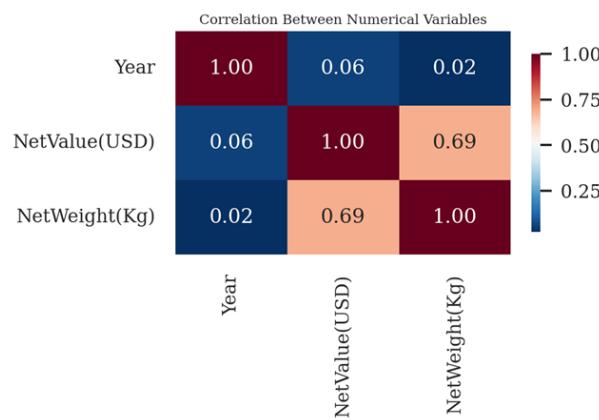

Figure 3. Correlation heatmap

The Compound Annual Growth Rate (CAGR) analysis shows a downward trend, with Net Value falling by -4.87% per year and Net Weight falling by -2.21% per year. The Augmented Dickey-Fuller (ADF) test shows that the data is not stationary (ADF statistic = -2.4577, p-value = 0.1261), as shown in Table 2. Non-stationarity is handled during Data Preparation stage using the Min-Max normalization method, which is

suitable for LSTM because its Attention mechanism is capable of learning non-linear temporal patterns from non-stationary time series data [31].

Table 2. CAGR and ADF test results

Analysis Aspect	Indicator	Value
Long-Term Growth (CAGR)	CAGR Net Value	-4.87% per year
	CAGR Net Weight	-2.21% per year
Stationarity Test (ADF Test)	ADF Statistic	-2.4577
	p-value	0.1261
	Critical Values	1%: -4.3316; 5%: -3.2330; 10%: -2.7487

Data Preparation

This stage aims to prepare the final dataset ready for modeling. Errors in this stage can have a significant impact on model accuracy, so it must be done carefully and systematically [32]. The process begins with cleaning the data, where missing values are handled to ensure that the dataset is consistent [33]. Figure 4 shows the matrix of missing values before cleaning, and Figure 5 shows the matrix after cleaning and transformation, resulting in several new columns being added to the DataFrame.
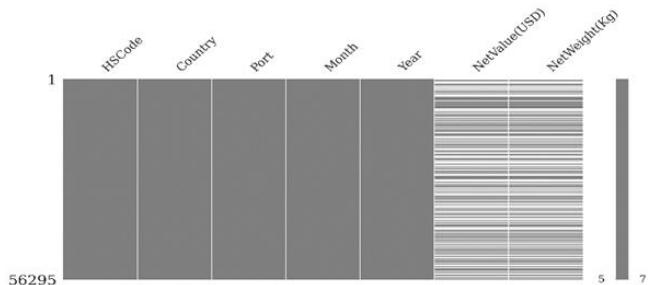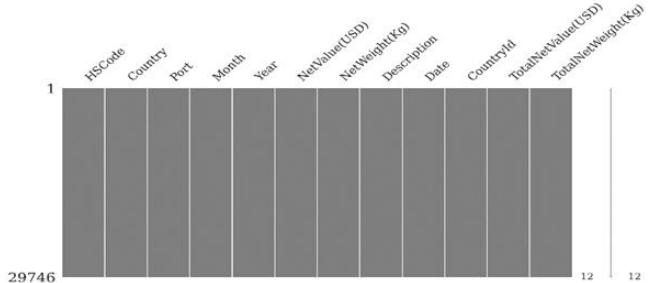

Figure 4. Matrix missing value before cleaning

Figure 5. Matrix missing value after cleaning

Pattern and seasonality analysis stage are conducted to understand how the behavior of mineral fuel export data changes. As shown in Figure 6, where the seasonal pattern box plot shows that Net Value has a seasonal pattern every month but is moderate, while the number of outliers shows large fluctuations due to factors such as changes in global mineral fuel prices. This seasonal pattern arises due to recurring cycles of production, demand, and trade contracts each year. Meanwhile, the cause of significant fluctuations, especially from unusual values, is due to factors outside the normal scope, such as changes in global mineral fuel prices, pandemics, disruptions in the supply chain, or sudden macroeconomic conditions. This seasonal pattern, which is not too high with unusual values, is in line with previous research on the volatility of energy commodity exports [34], [35].

Meanwhile as shown in Figure 7, the monthly export trend graph shows the development of exports from 2014 to 2025. There was a decline until 2016, followed by a recovery from 2016 to 2019, the impact of the pandemic from 2020 to 2021, and a significant increase from 2021 to 2022. The Covid-19 pandemic has weakened Indonesia's exports, including strategic commodities such as mineral fuels [36]. The difference between the Net Value and Net Weight scales shows that changes in value are more influenced by commodity prices than by the quantity of goods traded. This is the basis for the modeling process, in which Net Value is used as the main variable and Net Weight is used as a supporting variable. This is reinforced

by the results of correlation in Figure 3, where the heatmap shows a strong positive linear relationship between Net Value and Net Weight (r = 0.69).
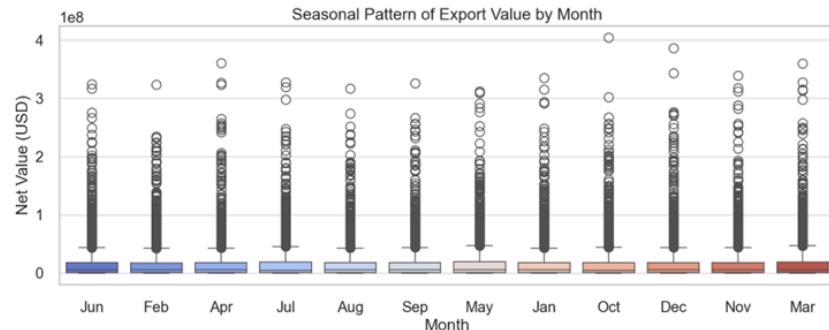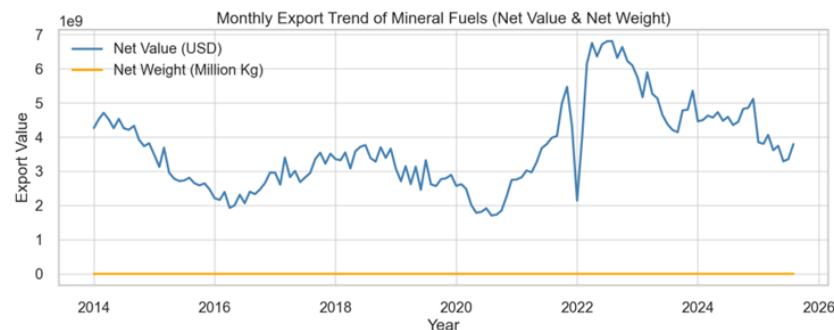

Figure 6. Seasonal pattern boxplot

Figure 7. Monthly export trend

Modeling

Based on Figure 8, the average Attention weight per lag shows that the lag 12 has the highest weight (0.7), followed by lags 11 and 10, which shows that the annual seasonal pattern is very strong in the monthly export data. This pattern is consistent with economic research showing that monthly data often experiences seasonal changes that repeat every 12 months due to the influence of the calendar, demand in certain seasons, and annual cycles in various economic activities [37]. Attention mechanism allows the model to focus on the most relevant lags from the input sequence, thereby improving accuracy [38], [39], [40]. In the LSTM model, the Attention mechanism can also help detect long-term patterns and extreme events better in complex time series data [41], [42]. This indicates that the previous year's export value has a major influence on forecasting, suggesting that the annual seasonal cycle is a dominant factor.
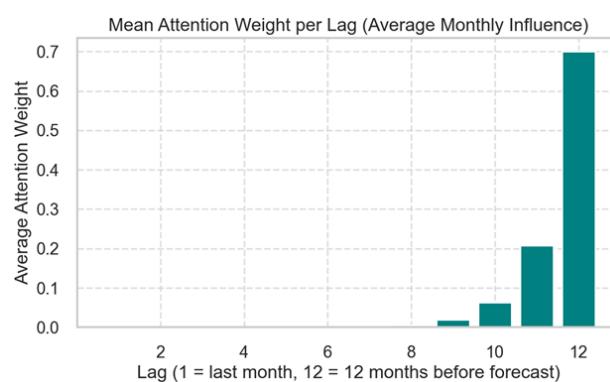

Figure 8. Attention weights

Then, a comparison of the forecast model with historical data that can be seen in Figure 9 shows that LSTM with an Attention mechanism is able to describe general trends, seasonal fluctuations, and long-term patterns well, although there are still slight deviations at some points [43]. The cause of these slight

deviations is due to the complexity of the model [44], [45] challenges in handling long-term dependencies [46], or the dynamic nature of Attention weights [47]. Moreover, 12-month future forecast describes a downward trend in export values until 2026, reflecting previous dynamics and seasonal patterns, with no signs of recovery in the short term, which can be seen in Figure 10.
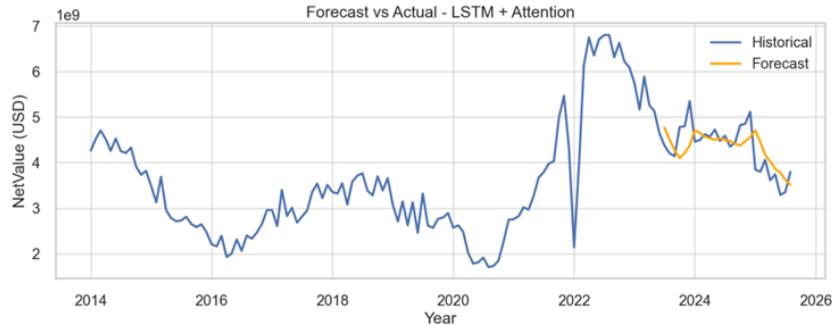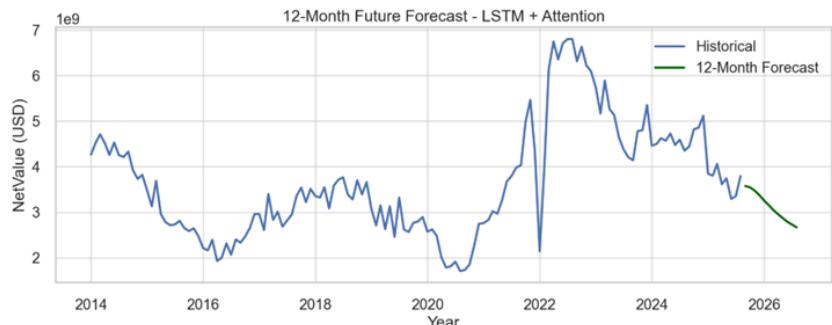

Figure 9. Forecast results by attention-based LSTM

Figure 10. Forecast results by attention-based LSTM

The results of the 12-month forecast show a consistent downward trend in the Net Value of mineral fuels. This decline began in the first month of the forecast period in September 2025 and continued sequentially until the last month in August 2026 as shown in Table 3. The model is evaluated using a time-based hold-out method, where 20% of the final data is used as a test set. This method mimics real-world forecasting conditions, as the model must be able to predict future values without access to data after the training period. This approach is often used in time series forecasting research with deep learning models to ensure a realistic assessment of the model's ability to predict new data [48].

Table 3. Forecasted net value results

Month	Forecasted Net Value (USD)
September 2025	3574440310
October 2025	3546575504
November 2025	3472535689
December 2025	3375285026
January 2026	3256562939
February 2026	3152465469
March 2026	3051832871
April 2026	2958562406
May 2026	2872559895
Juny 2026	2790832423
July 2026	2730453336
August 2026	2666438183

Evaluation
Performance of the model was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) [49], [50]. LSTM model with Attention mechanism and a combination of Dropout and Recurrent Dropout shows fairly good forecast results as shown in Table

4. The best validation loss value was recorded at 0.0068 using 160 neurons and a Tanh activation function with the Adam optimizer. The use of the Tanh activation function faces challenges such as vanishing gradients which can hinder model performance [51], and not hardware-friendly which makes it less efficient for certain applications [52]. On the other hand, several previous studies have shown that the Adam optimizer used as an optimization algorithm has proven effective for LSTM models with Attention mechanisms and is capable of improving accuracy and reliability in various cases [53], [54].

The forecasting error measures are MAE of 333,781,434.02 USD, RMSE of 420,272,857.16 USD, and MAPE of 7.76%, which indicates that the model is capable of accurately describing trends and changes in the Net Value of exports. The MAE and RMSE values are scale-dependent, so they have large values because they follow the original scale of the price data in USD, which has a large value [55], [56]. This is because MAE and RMSE are useful for providing an overview of absolute error, while MAPE is useful for comparing performance on data with different scales [57]. However, the very high MAE and RMSE values are not only due to scale-dependent reasons, but also indicate that the original data in USD has considerable variation and change. MAE and RMSE measure errors absolutely, so when the actual values are large, the errors that appear will also be large, even though the model is able to capture trend patterns well. In contrast, MAPE measures errors relatively in the form of percentages, making it more suitable for comparing model performance across different data sets with different scales [58], [59].

Table 4. Forecast tuning result configuration

Metric	Result
Validation Loss (Best)	0.0068
LSTM Units (Neurons)	160
Activation Function	Tanh
Optimizer	Adam
MAE	333,781,434.02 USD
RMSE	420,272,857.16 USD
MAPE	7.76 %

The results of the model performance comparison show that the neural network-based model (Attention-based LSTM) using Dropout and Recurrent Dropout has the best performance compared to other models, as shown in Table 5. This indicates that applying Dropout and Recurrent Dropout to LSTM can improve the accuracy of forecasting and help models adapt to data variations better without overfitting [60]. Dropout and Recurrent Dropout, which function as noise injection during the training process, can be effective regularizers if configured appropriately [13]. Dropout and recurrent dropout are only used in neural network-based models such as LSTM as a way to reduce overfitting during the training process [61], [25]. Meanwhile, traditional statistical models such as SARIMA and Prophet do not have neuron or hidden state structures, so they do not apply this type of regularization. The results of this study indicate that Dropout and Recurrent Dropout are effective ways to perform regularization on Attention-based LSTM model when used for time series multivariate data [62], [63]. This is consistent with previous research showing that neural network models such as LSTM are capable of understanding non-linear patterns and complex temporal relationships in multivariate time series data, thus generally providing more accurate results than traditional methods such as SARIMA and Prophet, which are linear and assume that the data is stationary [64].

Table 5. Model performance comparison

Model	RMSE (USD)	MAE (USD)	MAPE	Standard Deviation
LSTM + Attention (With Dropout + Recurrent Dropout)	420,272,857.16	333,781,434.02	7.76%	328,065,400
LSTM + Attention (Without Dropout + Recurrent Dropout)	714,496,808.34	579,268,274.97	12.62%	295,273,180
LSTM (With Dropout + Recurrent Dropout)	445,996,100	352,424,500	7.85%	363,598,100
LSTM (Without Dropout + Recurrent Dropout)	725,131,800	583,973,800	12.75%	277,430,200
SARIMA	1,757,205,000	1,549,631,000	38.04%	681,335,800
Prophet	2,540,723,000	2,361,519,000	57.66%	586,750,000

Residual analysis was performed using ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function), which showed that the residuals did not exhibit any significant autocorrelation patterns,

indicating that the residuals were random, similar to white noise, and did not have any specific patterns, which showed that the model successfully captured the time trend well [65], [66]. As shown in Table 6, the Ljung-Box test results reinforce that the residuals do not have significant autocorrelation, while the Breusch-Pagan test (p-value 0.740) shows that the residual variance is relatively constant [67], [68]. These findings indicate that the prediction errors mainly originate from data variation, not from an inaccurate model, so that the LSTM-Attention model successfully captures temporal relationships well.

Table 6. Ljung-box and breusch-pagan test results

Test	Lag	Test Statistic	p-value
Ljung-Box	10	7.41	0.686
Ljung-Box	15	20.17	0.165
Ljung-Box	20	24.70	0.213
Breusch-Pagan	-	LM = 0.11 (f = 0.10)	0.740 (f p = 0.753)

CONCLUSION

The Attention-based LSTM model combined with Dropout and Recurrent Dropout proved to be effective in forecasting multivariate time series data, especially complex export data. This model recognized annual seasonal patterns that were previously not clearly visible. Attention analysis shows that the previous year's value has the greatest impact. In contrast, the residual and forecast components indicate that the model can capture general trends and seasonal patterns, though it is less sensitive to unusual data. Some limitations in this study that need to be considered are that model evaluation was only carried out using the hold-out method, so there was no cross-validation, external validation, or use of other economic variables that could provide a more comprehensive assessment of the model's generalization and sensitivity capabilities. Therefore, the results obtained should be understood as an initial indication of the model's capabilities.

This study uses the hold-out method for validation testing, so that the model's ability to generalize is only checked on one part of the data. For more accurate and stable results, future studies are recommended to use the rolling/expanding window method or time series-based cross-validation. In addition, it is advisable to combine external variables such as global prices, energy policies, or macroeconomic indicators, and analyze hybrid models between statistics and deep learning, such as LSTM-ARIMA or LSTM-SARIMA combinations, in order to obtain more precise results and be able to detect unusual changes more effectively and complex seasonal patterns.

REFERENCES

[1]   S. I. Nikensari, N. Qisthina, P. Yuniarti, and S. Y. Wong, "Export competitiveness of Indonesia to ASEAN market for three leading commodities," vol. 04024, pp. 1–10, 2024, doi: https://doi.org/10.1051/e3sconf/202456804024.

[2]   N. Saqib, I. Ozturk, A. Sharif, and D. Cichoń, "Enhancing sustainable energy: Mineral exports, financial development, and foreign investment can build a greener future?," *Resources Policy*, vol. 97, p. 105249, 2024, doi: https://doi.org/10.1016/j.resourpol.2024.105249.

[3]   F. Munawar, G. Munawar, and D. Tarmidi, "Enhancing handicraft exports in West Java: A business intelligence approach to market expansion," *International Journal of Advanced and Applied Sciences*, vol. 11, no. 3, pp. 226–237, 2024, doi: 10.21833/ijaas.2024.03.022.

[4]   E. Dave, A. Leonardo, M. Jeanice, and N. Hanafiah, "Forecasting Indonesia Exports using a Hybrid Model ARIMA- LSTM," *Procedia Comput. Sci.*, vol. 179, no. 2020, pp. 480–487, 2021, doi: 10.1016/j.procs.2021.01.031.

[5]   A. I. Kurnadipare, S. Amaliya, K. A. Notodiputro, Y. Angraini, and L. N. A. Mualifah, "Comparing Forecasts of Agricultural Sector Export Values Using Sarima and Long Short-Term Memory Models," *Barekeng*, vol. 19, no. 1, pp. 385–396, 2025, doi: 10.30598/barekengvol19iss1pp385-396.

[6]   M. L. Hossain, S. M. N. Shams, and S. M. Ullah, "Time-series and deep learning approaches for renewable energy forecasting in Dhaka : a comparative study of ARIMA , SARIMA , and LSTM models," *Discover Sustainability*, vol. 6, no. 775, 2025, doi: https://doi.org/10.1007/s43621-025-01733-5.

[7]   Y. Liu, X. Hao, B. Zhang, and Y. Zhang, "Simplified long short-term memory model for robust and fast prediction," *Pattern Recognit. Lett.*, vol. 136, pp. 81–86, 2020, doi: 10.1016/j.patrec.2020.05.033.

[8] J. Ju and F.-A. Liu, "Multivariate Time Series Data Prediction Based on ATT-LSTM Network," *Applied Sciences (Switzerland)*, vol. 11, no. 20, p. 9373, 2021, doi: https://doi.org/10.3390/app11209373.

[9] L. Zhao, Y. Lai, S. Shi, G. Cheng, Z. Qiu, and Z. Xie, "Research on Financial Time Series Prediction Based on LSTM and Attention Mechanism," *Proceedings - 2025 Asia-Europe Conference on Cybersecurity, Internet of Things and Soft Computing, CITSC 2025*, pp. 347–351, 2025, doi: 10.1109/CITSC64390.2025.00069.

[10] J. Sun and W. Guo, "Time Series Prediction Based on Time Attention Mechanism and LSTM Neural Network," *2023 IEEE International Conference on Integrated Circuits and Communication Systems, ICICACS 2023*, pp. 1–5, 2023, doi: 10.1109/ICICACS57338.2023.10099498.

[11] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout : A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014.

[12] R. Maalej, "Online Arabic Handwriting Recognition with Dropout applied in Deep Recurrent Neural Networks," *2016 12th IAPR Workshop on Document Analysis Systems (DAS)*, pp. 417–421, 2016, doi: 10.1109/DAS.2016.49.

[13] E. Laksana, M. Aczon, L. Ho, C. Carlin, D. Ledbetter, and R. Wetzel, "The impact of extraneous features on the performance of recurrent neural network models in clinical tasks ☆," *J. Biomed. Inform.*, vol. 102, no. September 2019, p. 103351, 2020, doi: 10.1016/j.jbi.2019.103351.

[14] X. Zhang, C. Li, B. Xu, Z. Pan, and T. Yu, "Dropout Deep Neural Network Assisted Transfer Learning for Bi-Objective Pareto AGC Dispatch," *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1432–1444, 2023, doi: 10.1109/TPWRS.2022.3179372.

[15] M. Schmitt, "Deep learning in business analytics: A clash of expectations and reality," *International Journal of Information Management Data Insights*, vol. 3, no. 1, p. 100146, 2023, doi: 10.1016/j.jjimei.2022.100146.

[16] K. G. Sanabia-Lizarraga, B. Carballo-Mendívil, A. Arellano-González, and A. Bueno-Solano, "Business Intelligence for Agricultural Foreign Trade: Design and Application of Power BI Dashboard," *Sustainability (Switzerland)*, vol. 16, no. 21, 2024, doi: 10.3390/su16219576.

[17] W. Y. Ayele, "Adapting CRISP-DM for Idea Mining A Data Mining Process for Generating Ideas Using a Textual Dataset," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 11, no. 6, pp. 20–32, 2020.

[18] M. Pohl *et al.*, "The Linkage to Business Goals in Data Science Projects," *ACIS 2023 Proceedings*, vol. 60, 2023, doi: https://aisel.aisnet.org/acis2023/60.

[19] F. Li *et al.*, "A hierarchical temporal attention-based LSTM encoder-decoder model for individual mobility prediction," *Neurocomputing*, vol. 403, pp. 153–166, 2020, doi: 10.1016/j.neucom.2020.03.080.

[20] J. A. Ilemobayo, O. Durodola, A. Ogungbire, and A. Osinuga, "Hyperparameter Tuning in Machine Learning : A Comprehensive Review," *Journal of Engineering Research and Reports Volume*, vol. 26, no. 6, pp. 388–395, 2024.

[21] V. Cerqueira, L. Torgo, and I. Mozetič, *Evaluating time series forecasting models : an empirical study on performance estimation methods*, vol. 109, no. 11. Springer US, 2020. doi: 10.1007/s10994-020-05910-7.

[22] M. Vilares Ferro, Y. Doval Mosquera, F. J. Ribadas Pena, and V. M. Darriba Bilbao, "Early stopping by correlating online indicators in neural networks," *Neural Networks*, vol. 159, pp. 109–124, 2023, doi: 10.1016/j.neunet.2022.11.035.

[23] F. Chollet, *Deep Learning with Python*, Second Edi. Shelter Island, NY 11964: Manning Publications Co., 2018.

[24] Y. Gal and Z. Ghahramani, "A Theoretically Grounded Application of Dropout in Recurrent Neural Networks," *30th Conference on Neural Information Processing Systems*, no. Nips, 2016, doi: https://doi.org/10.48550/arXiv.1512.05287.

[25] I. Salehin and D.-K. Kang, "A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain," *Electronics 2023*, vol. 12, no. 3106, 2023.

[26] H. Jayadianti, S. Saptono, Hafsah, and Juwairiah, "Ontology - Based Data Integration of Minerals in Indonesia," *International Information Institute (Tokyo). Information*, vol. 19, no. 6B, pp. 2089–2097, 2016.

[27] Y. Astor, W. N. Sulasdi, and S. Hendriatiningsih, *The Evaluation of Marine Cadastre Definitions Among Australia , Canada and United States of America Based on Indonesia ' s Perspective as an Archipelagic State*. 2017. doi: 10.1007/978-3-319-51216-7.

[28] F. Zito, V. Cutello, and M. Pavone, "Data-driven forecasting and its role in enhanced decision-making," *Eng. Appl. Artif. Intell.*, vol. 154, no. March, p. 110934, 2025, doi: 10.1016/j.engappai.2025.110934.

[29] W. Quan, H. Nan, and P. Wang, "An Original Data Understanding Process," *CSAE '18*, no. 10, pp. 1–5, 2018, doi: https://doi.org/10.1145/3207677.3277974.

[30] M. Yay, "' Relationship between right atrial pressure and the Model For End-Stage Liver Disease (MELD) score in patients with advanced heart failure ': Correlation analysis and ROC curve method," *Turk Gogus Kalp Damar Cerrahisi Derg*, vol. 30, no. 1, pp. 8–10, 2022, doi: 10.5606/tgkdc.dergisi.2022.40073.

[31] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for interpretable multi-horizon time series forecasting," *Int. J. Forecast.*, vol. 37, no. 4, pp. 1748–1764, 2021, doi: 10.1016/j.ijforecast.2021.03.012.

[32] R. Nisbet, K. McCormick, and G. Miner, "A data preparation cookbook," in *Handbook of Statistical Analysis: AI and ML Applications, Third Edition*, University of California, Irvine, United States: Elsevier, 2024, pp. 125–134. doi: 10.1016/B978-0-443-15845-2.00012-8.

[33] P. C. Austin, I. R. White, D. S. Lee, and S. Van Buuren, "Missing Data in Clinical Research : A Tutorial on Multiple Imputation," *Canadian Journal of Cardiology*, vol. 37, no. 9, pp. 1322–1331, 2021, doi: 10.1016/j.cjca.2020.11.010.

[34] M. Qadan, D. Y. Aharon, and R. Eichel, "Seasonal patterns and calendar anomalies in the commodity market for natural resources," *Resources Policy*, vol. 63, no. June, p. 101435, 2019, doi: 10.1016/j.resourpol.2019.101435.

[35] C. Frau and V. Fanelli, "Seasonality in commodity prices : new approaches for pricing plain vanilla options," *Ann. Oper. Res.*, vol. 336, no. 1, pp. 1089–1131, 2024, doi: 10.1007/s10479-022-05128-x.

[36] Ngatno and A. E. Prihatiningsih, "Analysis of Indonesia's Exports & Imports in the Asian Region: Before and During the Covid-19 Pandemic," *Asian Economic and Financial Review*, vol. 11, no. 11, pp. 923–937, 2021, doi: 10.18488/journal.aefr.2021.1111.923.937.

[37] T. Proietti and D. J. Pedregal, "Seasonality in High Frequency Time Series," *CEIS Working Paper No. 508*, vol. 19, no. 2, 2021, doi: https://dx.doi.org/10.2139/ssrn.3802611.

[38] L. Zhang *et al.*, "Time-Series Neural Network : A High-Accuracy Time-Series Forecasting Method Based on Kernel Filter and Time Attention," *Information*, vol. 14, no. 500, pp. 1–18, 2023, doi: https://doi.org/10.3390/info14090500.

[39] C. Cao, X. Xie, and S. Ju, "Wind Power Prediction Based on BiLSTM- Attention Modeling," *2023 IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE)*, pp. 1391–1396, 2023, doi: 10.1109/ICEACE60673.2023.10442200.

[40] J. Yuan, H. Dengxin, W. Yufeng, Y. Xueting, D. Huige, and Y. Qing, "Attention mechanism based CNN-LSTM hybrid deep learning model for atmospheric ozone concentration prediction," *Sci. Rep.*, vol. 15, no. 21260, pp. 1–16, 2025, doi: https://doi.org/10.1038/s41598-025-05877-2.

[41] M. A. Lahyani and M. Amayri, "Explainable Hybrid Deep Learning Model with Attention Mechanism for Short-Term Load Forecasting," *Sustainable Cities and Society: Advances*, vol. 1, no. 1, p. 100003, 2025, doi: 10.1016/j.scsadv.2025.100003.

[42] H. Liu and L. Yang, "A Comparative Study of CNN-sLSTM-Attention-Based Time Series Forecasting : Performance Evaluation on Data with Symmetry and Asymmetry Phenomena," *Symmetry (Basel).*, vol. 17, no. 1846, pp. 1–27, 2025.

[43] X. Zhou, A. Pranolo, and Y. Mao, "AB-LSTM : Attention Bidirectional Long Short-Term Memory for Multivariate Time-Series Forecasting," *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)*, pp. 1–6, 2023, doi: 10.1109/IC2E357697.2023.10262559.

[44] Z. Zhang, F. Zhou, and J. Sun, "Rolling bearing fault diagnosis based on Attention LSTM," *2023 38th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 602–607, 2023, doi: 10.1109/YAC59482.2023.10401378.

[45] S. Sang and L. Li, "A Novel Variant of LSTM Stock Prediction Method Incorporating Attention Mechanism," *Mathematics*, vol. 12, no. 945, 2024, doi: https://doi.org/10.3390/math12070945.

[46] T. Grósz and M. Kurimo, "LSTM-XL: Attention Enhanced Long-Term Memory for LSTM Cells BT - Text, Speech, and Dialogue," K. Ekštein, F. Pártl, and M. Konopík, Eds., Cham: Springer International Publishing, 2021, pp. 382–393.

[47] G. Zhong, X. Lin, K. Chen, and Q. Li, "Attention mechanism-based long short-term memory model," in *Attention Augmented Learning Machines: Theory and Applications*, College of Computer Science and Technology, Ocean University of China, Qingdao, China: Nova Science Publishers, Inc., 2023, pp. 39–47.

[48] P. Hiskiawan, J. William, L. Feliepe, and T. Jansel, "A Hybrid Data Science Framework for Forecasting Bitcoin Prices using Traditional and AI Models," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 5, pp. 2089–2101, 2025.

[49] F. Zeeshan, N. Nepal, and M. Norouzifard, "Measurement and forecasting of fluctuating Cryptocurrency prices using deep learning," *Proceedings of the 9th International Conference on Computer and Communication Engineering, ICCCE 2023*, vol. 506, pp. 258–263, 2023, doi: 10.1109/ICCCE58854.2023.10246100.

[50] M. Aloudah, M. Alajmi, A. Sagheer, A. Algosaibi, B. Almarri, and E. Albelwi, "AI-Powered Trade Forecasting: A Data-Driven Approach to Saudi Arabia's Non-Oil Exports," *Big Data and Cognitive Computing*, vol. 9, no. 4, 2025, doi: 10.3390/bdcc9040094.

[51] M. H. E. Ali, A. B. Abdel-raman, and E. A. Badry, "Developing Novel Activation Functions Based Deep Learning LSTM for Classification," *IEEE Access*, vol. 10, no. September, pp. 97259–97275, 2022, doi: 10.1109/ACCESS.2022.3205774.

[52] Y. Zhang *et al.*, "b -sigmoid and 2 b -tanh : Low Hardware Complexity Activation Functions for LSTM," *2022 19th International SoC Design Conference (ISOCC)*, no. 1, pp. 93–94, 2022, doi: 10.1109/ISOCC56007.2022.10031500.

[53] T. Mamo and F. Wang, "Attention-Based Long Short-Term Memory Recurrent Neural Network for Capacity Degradation of Lithium-Ion Batteries," *Batteries*, vol. 7, no. 66, pp. 1–9, 2021, doi: https://doi.org/10.3390/ batteries7040066.

[54] A. Wan, C. Du, K. AL-Bukhaiti, and P. Chen, "Optimizing combined-cycle power plant operations using an LSTM-attention hybrid model for load forecasting," *Journal of Mechanical Science and Technology*, vol. 39, no. 10, pp. 6371–6380, 2025, doi: 10.1007/s12206-025-0961-3.

[55] G. Buturac, "Measurement of Economic Forecast Accuracy : A Systematic Overview of the Empirical Literature," *Journal of Risk and Financial Management*, vol. 15, no. 1, pp. 1–28, 2022, doi: https://doi.org/10.3390/jrfm15010001.

[56] R. L. Manogna, V. Dharmaji, and S. Sarang, "Enhancing agricultural commodity price forecasting with deep learning," *Sci. Rep.*, vol. 15, no. 20903, pp. 1–24, 2025, doi: https://doi.org/10.1038/s41598-025-05103-z 1.

[57] P. Manandhar, H. Rafiq, E. Rodriguez-Ubinas, and T. Palpanas, "New Forecasting Metrics Evaluated in Prophet , Random Forest , and Long Short-Term Memory Models for Load Forecasting," *Energies (Basel).*, vol. 17, no. 6131, pp. 1–30, 2024, doi: https://doi.org/10.3390/en17236131.

[58] C. Tofallis, "A better measure of relative prediction accuracy for model selection and model estimation," *Journal of the Operational Research Society*, vol. 66, no. 8, pp. 1352–1362, Aug. 2015, doi: 10.1057/jors.2014.103.

[59] S. K. Morley, T. V. Brito, and D. T. Welling, "Measures of Model Performance Based On the Log Accuracy Ratio," *Space Weather*, vol. 16, pp. 69–88, 2018, doi: 10.1002/2017SW001669.

[60] L. M. Cendani, R. Kusumaningrum, and S. N. Endah, "Aspect-Based Sentiment Analysis of Indonesian-Language Hotel Reviews Using Long Short-Term Memory with an Attention Mechanism," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 147, Department of Informatics, Diponegoro University, Semarang, Indonesia: Springer Science and Business Media Deutschland GmbH, 2023, pp. 106–122. doi: 10.1007/978-3-031-15191-0_11.

[61] N. Watt and M. C. du Plessis, "Dropout Algorithms for Recurrent Neural Networks," *SAICSIT '18: 2018 Annual Conference of the South African Institute of Computer Scientists and Information Technologists*, no. September 2018, pp. 72–78, 2018, doi: 10.1145/3278681.3278691.

[62] X. Wu, Y. Wei, A. Wang, Y. Lei, and R. Zhang, "Vehicle Trajectory Prediction Based on LSTM-GRU Integrating Dropout and Attention Mechanism," *Hunan Daxue Xuebao/Journal of Hunan University Natural Sciences*, vol. 50, no. 4, pp. 65–75, 2023, doi: 10.16339/j.cnki.hdxbzkb.2023155.

[63] L. Yang, Y. Jiang, K. Wang, P. Zhao, and K. Li, "EL-LSTM: A Multivariate Time Series Forecasting Model Combining Spiking Neurons and Long Short-Term Memory Networks BT -

Neural Information Processing," M. Mahmud, M. Doborjeh, K. Wong, A. C. S. Leung, Z. Doborjeh, and M. Tanveer, Eds., Singapore: Springer Nature Singapore, 2025, pp. 241–255.

[64] V. Kaushal, "Learning successful implementation of Chatbots in businesses from B2B customer experience perspective," *Concurr. Comput.*, vol. 35, no. 1, 2023, doi: 10.1002/cpe.7450.

[65] C. H. Weiß, B. Aleksandrov, M. Faymonville, and C. Jentsch, "Partial Autocorrelation Diagnostics for Count Time Series," *Entropy 2023*, vol. 25, no. 105, pp. 1–21, 2023, doi: https://doi.org/10.3390/ e25010105.

[66] H. Hassani, L. M. Mashhad, M. Royer-Carenzi, M. R. Yeganegi, and N. Komendantova, "White Noise and Its Misapplications : Impacts on Time Series Model Adequacy and Forecasting," *Forecasting 2025*, vol. 7, no. 8, pp. 1–14, 2025, doi: https://doi.org/10.3390/ forecast7010008.

[67] A. Halunga, C. D. Orme, and T. Yamagata, "A Heteroskedasticity Robust Breusch-Pagan Test for Contemporaneous Correlation in Dynamic Panel Data," *J. Econom.*, 2017, doi: 10.1016/j.jeconom.2016.12.005.

[68] T. Lee, "Wild bootstrap Ljung–Box test for residuals of ARMA models robust to variance change," *J. Korean Stat. Soc.*, vol. 51, no. 4, pp. 1005–1020, 2022, doi: 10.1007/s42952-022-00172-6.