COMPARISON OF ARIMA, LSTM, AND GRU MODELS FOR FORECASTING SALES OF HIT AEROSOL PRODUCTS

Nendi Sunendar*; Yan Rianto

Computer Science
Universitas Nusa Mandiri, Depok, Indonesia
www.nusamandiri.ac.id
14240027@nusamandiri.ac.id*, yanr001@gmail.com
(*) Corresponding Author

Abstract*—A more accurate forecasting model, such as LSTM, can significantly enhance business efficiency by providing more reliable predictions of future sales, allowing for better inventory management, optimized production schedules, and more precise distribution planning. This leads to reduced costs, minimized stockouts, and improved customer satisfaction. This study evaluates the forecasting performance of ARIMA, Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU) models using sales data from 2021 to 2023. The models are assessed based on Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Results show that LSTM outperforms the other models with a MAPE of 10.76%, followed by ARIMA at 11.23% and GRU at 11.47%. These findings highlight the advantages of deep learning methods, particularly LSTM, in capturing complex patterns and trends in time series data. The study demonstrates the potential of these models to optimize sales forecasting, aiding decision-making processes in production and distribution planning.*

Keywords: *ARIMA, GRU, LSTM, sales prediction, time series forecasting.*

Abstrak*—Model peramalan yang lebih akurat, seperti LSTM, dapat secara signifikan meningkatkan efisiensi bisnis dengan memberikan prediksi penjualan yang lebih dapat diandalkan, memungkinkan pengelolaan persediaan yang lebih baik, penjadwalan produksi yang lebih optimal, dan perencanaan distribusi yang lebih tepat. Hal ini dapat mengurangi biaya, meminimalkan kekurangan stok, dan meningkatkan kepuasan pelanggan. Penelitian ini mengevaluasi kinerja peramalan model ARIMA, Long Short-Term Memory (LSTM), dan Gated Recurrent Unit (GRU) menggunakan data penjualan dari tahun 2021 hingga 2023. Model-model tersebut dievaluasi berdasarkan Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), dan Mean Absolute Percentage Error (MAPE). Hasil penelitian menunjukkan bahwa LSTM mengungguli model lainnya dengan MAPE sebesar 10,76%, diikuti oleh ARIMA sebesar 11,23% dan GRU sebesar 11,47%. Temuan ini menyoroti keunggulan metode deep learning, khususnya LSTM, dalam menangkap pola dan tren kompleks pada data deret waktu. Penelitian ini menunjukkan potensi model-model tersebut untuk mengoptimalkan peramalan penjualan, yang dapat mendukung proses pengambilan keputusan dalam perencanaan produksi dan distribusi.*

Kata Kunci: *ARIMA, GRU, LSTM, prediksi penjualan, peramalan deret waktu.*

INTRODUCTION

The evolution of data science and machine learning has transformed how businesses analyze and predict future trends. In the retail and manufacturing sectors, predicting product demand is crucial for optimizing inventory, production, and distribution strategies (Bilgili & Pinar, 2023). Accurate sales forecasting enables companies to reduce operational costs, prevent overstocking, and improve customer satisfaction by meeting market demands efficiently. Time series forecasting, a widely adopted technique, plays a pivotal role in sales prediction. It utilizes historical data to predict future values, helping businesses anticipate demand fluctuations and seasonal trends. Among the popular methods, traditional statistical models like ARIMA (Auto-Regressive Integrated Moving Average) and machine learning models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have shown significant potential in capturing temporal patterns. These models differ in their approach to processing and analyzing data, with ARIMA relying on linear relationships and LSTM and GRU leveraging deep learning techniques to understand nonlinear and complex dependencies (Yavasani & Wang, 2023).

A more accurate forecasting model, such as LSTM, can significantly enhance business efficiency by providing more reliable predictions of future sales, allowing for better inventory management, optimized production schedules, and more precise distribution planning. This leads to reduced costs, minimized stockouts, and improved customer satisfaction. The ARIMA model has been a cornerstone of time series forecasting for decades due to its simplicity and effectiveness in analyzing linear patterns. However, it often struggles with nonlinear relationships and requires extensive preprocessing, such as stationarity adjustments and differencing. On the other hand, LSTM and GRU, as variants of recurrent neural networks (RNNs), excel in capturing long-term dependencies and handling large datasets with intricate patterns. These deep learning models can learn from sequential data and provide better accuracy for complex time series datasets (Wu, Du, Zhang, Zhou, & Liu, 2023).

In the context of this research, the sales data of HIT Aerosol products is analyzed to evaluate the performance of ARIMA, LSTM, and GRU models. The data spans three years, from 2021 to 2023, capturing daily sales from multiple distribution channels. The objective is to identify the most suitable model for sales forecasting, which can assist businesses in making informed decisions related to production planning and supply chain management.

The importance of accurate sales forecasting cannot be overstated. For consumer goods companies, the ability to predict sales accurately impacts not only inventory management but also production efficiency and market responsiveness. HIT Aerosol, a product distributed widely in various regions, exhibits demand patterns influenced by seasonal trends and market dynamics (Darmawan, Alfarisi, & Hozairi, 2024).

Understanding these patterns and predicting future sales effectively can provide a competitive edge to businesses. Furthermore, the increasing availability of sales data and advancements in computational power have paved the way for implementing sophisticated machine learning algorithms. While traditional models like ARIMA are still widely used, the emergence of LSTM and GRU has introduced opportunities to achieve higher accuracy in forecasting . This study seeks to explore these opportunities by comparing the strengths and weaknesses of these models in a real-world scenario (Rusman, Chunady, Makmud, Setiawan, & Hasani, 2023).

The primary objectives of this research are to evaluate the accuracy of ARIMA, LSTM, and GRU models in forecasting HIT Aerosol sales, compare the performance of these models using metrics such as Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), identify the most suitable model for capturing the seasonal and nonlinear patterns in sales data, and provide practical recommendations for businesses in optimizing their forecasting processes based on the findings (Karunasingha, 2022).

This study focuses on three forecasting models: ARIMA, LSTM, and GRU. The sales data of HIT Aerosol products is used as the dataset, spanning three years (2021-2023). The analysis includes preprocessing steps such as handling missing values, detecting and addressing outliers, and normalizing data for deep learning models. The performance of the models is evaluated based on their ability to predict sales accurately over a specified period (Si, Nadarajah, Zhang, & Xu, 2024). However, previous studies have shown that Prophet tends to be less optimal in capturing complex nonlinear patterns in sales data compared to deep learning models such as LSTM and GRU . Since the dataset in this study exhibits more complex patterns beyond mere seasonal trends, Prophet is considered less suitable for achieving high forecasting accuracy. Hybrid models such as ARIMA-LSTM or ARIMA-GRU have shown promising results in recent forecasting studies (Jain, Agrawal, Mohapatra, & Srinivasan, 2024). Previous studies have indicated that the accuracy improvement from hybrid models is often only marginal compared to standalone deep learning models, especially when data preprocessing is performed effectively (Kurniawan, Parhusip, & Trihandaru, 2024).

This research contributes to the field of sales forecasting by highlighting the practical applications of machine learning models in a real-world context. By comparing traditional statistical methods with advanced deep learning techniques, this study provides insights into the advantages and limitations of each approach. The findings are expected to assist businesses in selecting appropriate forecasting models, thereby improving decision-making processes and operational efficiency (Abubaker & Ala'Khalifeh, 2023).

The rest of the paper is structured as follows: Section 2 provides an overview of related work, discussing previous studies on sales forecasting and the models used. Section 3 describes the methodology, including data collection, preprocessing, and model implementation. Section 4 presents the results and analysis, comparing the performance of ARIMA, LSTM, and GRU. Section 5 discusses the implications of the findings and practical recommendations for businesses. Section 6 concludes the study, summarizing the key insights and suggesting directions for future research (Teixeira et al. 2024).

MATERIALS AND METHODS

The dataset used in this study consists of daily sales data for HIT Aerosol products collected over a three-year period (2021–2023). The data, obtained from distribution channels, includes key attributes such as product category, sales volume, and timestamps. Four product categories are covered: Blooming Tea, Citrus, Sweet Flower, and Lily Blossom. These categories capture the diversity of demand patterns and provide a comprehensive basis for the analysis. A summary of the dataset, including average daily sales, total data points, and detected outliers, is presented in Table 1.

Table 1. Dataset Summary

No	Product Category	Average Daily Sales	Total Data Points	Outliers Detected
1	Blooming Tea	21.89	1095	15
2	Citrus	15.76	1095	12
3	Sweet Flower	16.69	1095	18
4	Lily Blossom	24.31	1095	10

Source: (Research Results, 2025)

Each product category represents a distinct group of HIT Aerosol products with varying demand patterns. "Average Daily Sales" indicates the mean daily sales volume for each category over the three-year period. "Total Data Points" reflects the number of daily sales records collected, while "Outliers Detected" highlights the number of abnormal data points identified and addressed during preprocessing. The Blooming Tea category had an average daily sales volume of 21.89 units, with 15 outliers detected. Citrus exhibited the lowest average daily sales (15.76 units) and the fewest outliers (12). Sweet Flower had an average daily sales volume of 16.69 units and 18 outliers, the highest among the categories. Finally, Lily Blossom showed the highest average daily sales (24.31 units) and 10 outliers, the lowest among all categories. These variations in sales and outlier counts reflect the diverse demand dynamics of each product category, making the dataset suitable for evaluating forecasting models (Xiang, 2024).

To ensure the quality and usability of the data, several preprocessing steps were performed. Missing entries in the dataset were identified and addressed by interpolating values using linear interpolation (AlSalehy & Bailey, 2025). Outliers were detected using the Interquartile Range (IQR) method and replaced with the median values of the respective product categories. Additionally, the data was normalized using MinMaxScaler to scale values between 0 and 1, ensuring compatibility with deep learning models (Brykin, 2024).

For model optimization, hyperparameter tuning was conducted using grid search, systematically exploring different combinations of parameters such as batch size, learning rate, dropout rate, and the number of LSTM/GRU units. Cross-validation helps prevent overfitting and provides a more reliable assessment of model performance, which is essential for time series forecasting.

By incorporating both grid search for hyperparameter tuning and k-fold cross-validation for validation, this study ensures that the selected models are both optimized and capable of delivering consistent performance across different data subsets.

Three models were implemented to forecast sales: ARIMA, LSTM, and GRU. ARIMA (Auto-Regressive Integrated Moving Average) is a statistical model widely used for time series forecasting. The auto_arima function was utilized to determine the optimal parameters. The data was checked for stationarity using the Augmented Dickey-Fuller (ADF) test, and non-stationary data was differenced to achieve stationarity before fitting the ARIMA model to the training data and making predictions. LSTM (Long Short-Term Memory) is a variant of recurrent neural networks (RNNs) capable of capturing long-term dependencies in sequential data. Key hyperparameters such as the number of LSTM units, batch size, and epochs were optimized using grid search. The data was divided into training and testing sets with a ratio of 80:20, and LSTM layers were constructed with dropout regularization to prevent overfitting. The model was trained using the Adam optimizer and Mean Squared Error (MSE) as the loss function (Abumohsen, Owda, Owda, & Abumihsan, 2024).

GRU (Gated Recurrent Unit) simplifies the architecture of LSTM while maintaining comparable performance. Similar to LSTM, key hyperparameters were optimized. The GRU model was constructed with layers similar to LSTM, using fewer parameters for faster training, and was trained and evaluated using the same dataset split and metrics as LSTM (Casado-Vara et. al, 2021).

RESULTS AND DISCUSSION

The ARIMA model is applied to several products, each with parameters tailored based on data analysis. For the BLOOMINGTEA product, the parameters used are (4, 1, 5), with p = 4 for autoregression, d = 1 for differencing, and q = 5 for moving average. The CITRUS product uses the parameters (2, 1, 1), SWEETFLOWER uses (0, 1, 2), and LILYBLOSSOM uses (5, 0, 5). Each ARIMA model is customized to the characteristics of the training data, which is then used to predict the test data. The complete ARIMA parameter settings for each product are summarized in Table 2.

Table 2. Parameter ARIMA

No	Product Category	p (AR order)	d (Differencing order)	q (MA order)
1	Blooming Tea	4	1	5
2	Citrus	2	1	1
3	Sweet Flower	0	1	2
4	Lily Blossom	5	0	5

Source: (Research Results, 2025)

The LSTM and GRU parameter tables outline the key configurations used for each model applied to various products. Both models use two layers with 50 units each, ensuring that the models have sufficient capacity to capture complex patterns in the data. The Optimizer used for both LSTM and GRU models is Adam, known for its efficiency and adaptability in training deep learning models. The Loss Function for both models is Mean Squared Error (MSE), which is commonly applied in regression tasks to evaluate prediction accuracy by measuring the squared difference between predicted and actual values.

Both models are trained for 50 epochs, meaning the training process is repeated 50 times over the dataset to adjust the model's weights. The Batch Size is set to 1 for both models, indicating that each training example is processed individually in each update. These parameters are consistent across all products in both LSTM and GRU models, providing a clear structure for how each model processes time series data to generate predictions. These hyperparameter configurations are detailed in Table 3.

Table 3. Parameter LSTM and GRU

No	Product Category	Units (Layer 1/2)	Loss Func	Epochs	Batch Size
1	Blooming Tea	50/50	MSE	50	1
2	Citrus	50/50	MSE	50	1
3	Sweet Flower	50/50	MSE	50	1
4	Lily Blossom	50/50	MSE	50	1

Source: (Research Results, 2025)

The evaluation of ARIMA, LSTM, and GRU models was conducted using multiple performance metrics to measure prediction accuracy comprehensively. The metrics included Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics allowed for a detailed comparison of the models' capabilities in handling the sales forecasting task.

Table 4. Evaluation Metrics Comparison

No	Model	MSE	RMSE	MAE	MAPE
1	ARIMA	8.38	2.89	2.20	11.23
2	LSTM	6.26	2.50	1.64	10.76
3	GRU	6.49	2.55	1.68	11.47

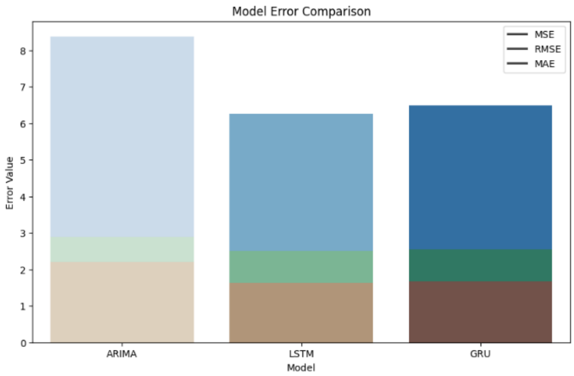
Source: (Research Results, 2025)

From the evaluation results summarized in Table 4, the LSTM model demonstrated the best overall performance, achieving the lowest MAPE of 10.76%, which indicates superior accuracy in capturing the nonlinear and seasonal trends in the sales data. LSTM also recorded the lowest MSE (6.26), RMSE (2.50), and MAE (1.64), confirming its ability to reduce error magnitudes effectively.

The GRU model, while slightly less accurate than LSTM, showed competitive performance with a MAPE of 11.47%. This suggests that GRU's simplified architecture maintains a high level of efficiency in handling complex time series data, albeit with a marginal trade-off in prediction precision. In contrast, the ARIMA model, a traditional statistical approach, achieved a MAPE of 11.23%. Although ARIMA performed adequately, its reliance on linear assumptions and challenges in handling nonlinear relationships limited its accuracy compared to the deep learning models. Additionally, ARIMA required extensive preprocessing, such as stationarity adjustments, making it less flexible for datasets with complex patterns (Ospina, Gondim, Leiva, & Castro, 2023).

This study analyzed and compared the performance of ARIMA, LSTM, and GRU models in forecasting sales of HIT Aerosol products using sales data from 2021 to 2023. Among the three models, LSTM demonstrated superior accuracy with the lowest MAPE of 10.76%, followed by ARIMA (11.23%) and GRU (11.47%). The findings highlight the ability of deep learning models, particularly LSTM, to effectively capture complex and nonlinear patterns in time series data, making them ideal for sales forecasting tasks. On the other hand, ARIMA, while simpler and computationally efficient, showed limitations in handling nonlinear relationships.

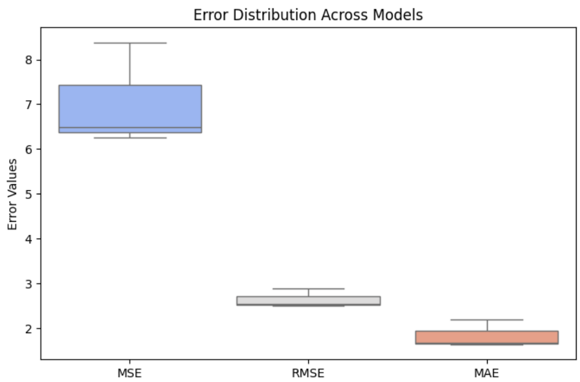
The bar chart in Figure 1 illustrates the performance of three forecasting models—ARIMA, LSTM, and GRU—evaluated using three error metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The chart highlights that ARIMA has the highest error values, indicating its lower accuracy in sales forecasting. Among the three metrics, MSE dominates, making ARIMA less effective for capturing complex patterns.



Source: (Research Results, 2025)
Figure 1. Model Error Comparison

In contrast, LSTM exhibits the lowest error values across all metrics, suggesting its superior ability to model nonlinear relationships in time series data. The GRU model performs slightly worse than LSTM but remains significantly better than ARIMA. The difference between LSTM and GRU is relatively small, implying that GRU can be a viable alternative when computational efficiency is a priority.

Overall, the findings confirm that LSTM outperforms both ARIMA and GRU in sales forecasting. The results indicate that deep learning models, particularly LSTM, are more effective in capturing intricate demand patterns, making them preferable for real-world forecasting applications. Meanwhile, ARIMA's limitations in handling nonlinear trends reinforce the need for more advanced modeling techniques. The comparison underscores the advantages of deep learning over traditional statistical approaches in time series forecasting.



Source: (Research Results, 2025)
Figure 2. Error Distribution across Model

Further insights into model performance are provided in Figure 2, which presents a box plot of error distributions across the forecasting models. The plot illustrates the variability of three error metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). MSE has the highest values and largest variability, indicating that some models exhibit significantly higher squared errors. RMSE shows a more concentrated distribution, with values between 2.5 and 3, suggesting consistent performance. MAE has the lowest and most stable error values, with minimal variability. The results indicate that MSE is more sensitive to outliers, while RMSE and MAE provide more stable error measurements. This highlights the importance of selecting appropriate evaluation metrics for model comparison in time series forecasting.

This research also underscores the importance of data preprocessing, including outlier treatment and normalization, in ensuring model accuracy. By addressing challenges such as demand variability and seasonal trends, the models provided actionable insights for optimizing inventory management, production planning, and resource allocation. The results affirm that businesses aiming for higher forecasting precision should adopt advanced deep learning techniques like LSTM or GRU, provided they have the computational resources required (Pierre, Akim, Semenyo, & Babiga, 2023).

Discussion

The comparison of ARIMA, LSTM, and GRU revealed critical trade-offs between traditional statistical methods and modern deep learning approaches. ARIMA’s reliance on linear assumptions limits its ability to model complex patterns, making it less suitable for datasets with significant nonlinearities. However, its simplicity and low computational requirements make it a practical choice for straightforward forecasting tasks or when computational resources are constrained.

In contrast, LSTM and GRU excel in capturing intricate temporal dependencies and patterns, owing to their ability to retain information over long sequences. The marginal difference in performance between LSTM and GRU indicates that GRU’s simplified architecture can be a suitable alternative when training time or computational efficiency is a priority. However, LSTM remains the preferred model for scenarios where prediction accuracy is paramount.

The evaluation metrics further validate the models’ performance. LSTM’s superior MSE, RMSE, and MAE scores highlight its precision in minimizing errors and capturing demand fluctuations. Meanwhile, ARIMA, despite being outperformed, provided a baseline for

Vol. 21, No. 2 September 2025 | DOI: 10.33480/pilar.v21i2.6412

understanding the advantages of deep learning techniques. The results of this study indicate that the LSTM model (10.76% MAPE) outperforms both ARIMA (11.23%) and GRU (11.47%) in forecasting the sales of HIT Aerosol products. These findings align with prior literature, such as (Bilgili & Pinar, 2023) and (Yavasani & Wang, 2023), which highlight the advantages of LSTM in capturing nonlinear patterns and long-term dependencies in time series data. The primary strength of this study lies in the use of a more comprehensive dataset (three years of daily data) and the implementation of dropout regularization in the LSTM model to prevent overfitting—an approach not employed in previous studies such as that of (Karunasingha, 2022). Moreover, the GRU model presents a more computationally efficient alternative to ARIMA, despite performing slightly below LSTM, offering flexibility for companies with limited computational resources, as noted by (Pierre, Akim, Semenyo, & Babiga, 2023).

CONCLUSION

In conclusion, this study highlights the significant potential of deep learning models, particularly LSTM, in advancing sales forecasting. By addressing the challenges inherent in time-series data, these models empower businesses to make data-driven decisions, ultimately improving operational efficiency and competitive advantage.

Evaluating the performance of three forecasting models ARIMA, LSTM, and GRU by analyzing their error metrics, including MSE, RMSE, and MAE. The goal was to determine which model provides the most accurate sales forecasts. The findings revealed that LSTM outperformed ARIMA and GRU, achieving the lowest error values across all metrics. GRU performed competitively but slightly worse than LSTM, while ARIMA exhibited the highest error rate, highlighting its limitations in handling complex sales patterns. These results confirm that deep learning models, particularly LSTM, are better suited for time-series forecasting compared to traditional statistical models such as ARIMA.

Future research could explore hybrid models that combine ARIMA with LSTM or GRU to leverage the strengths of both approaches. Additionally, expanding the dataset to include external variables, such as promotional events or economic factors, can improve forecast accuracy and provide a more comprehensive analysis of sales dynamics. Furthermore, applying cross-validation techniques can improve the robustness of model evaluation, ensuring consistent performance across different subsets of data.

REFERENCE

Abubaker, F., & Ala'Khalifeh. (2023). Sales' Forecasting Based on Big Data and Machine Learning Analysis. *2023 9th International Conference on Control, Decision and Information Technologies (CoDIT)*, 804–808. https://doi.org/10.1109/CoDIT58514.2023.10284159

Abumohsen, M., Owda, A. Y., Owda, M., & Abumihsan, A. (2024). Hybrid machine learning model combining of CNN-LSTM-RF for time series forecasting of Solar Power Generation. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, 9, 100636. https://doi.org/10.1016/j.prime.2024.100636

AlSalehy, A. S., & Bailey, M. (2025). Improving Time Series Data Quality: Identifying Outliers and Handling Missing Values in a Multilocation Gas and Weather Dataset. *Smart Cities*, 8(3), 82. https://doi.org/10.3390/smartcities8030082

Bilgili, M., & Pinar, E. (2023). Gross electricity consumption forecasting using LSTM and SARIMA approaches: A case study of Türkiye. *Energy*, 284, 128575. https://doi.org/10.1016/j.energy.2023.128575

Brykin, D. (2024). Sales Forecasting Models: Comparison between ARIMA, LSTM and Prophet. *Journal of Computer Science*, 20(10), 1222–1230. https://doi.org/10.3844/jcssp.2024.1222.1230

Casado-Vara, R., Martin del Rey, A., Pérez-Palau, D., de-la-Fuente-Valentín, L., & Corchado, J. M. (2021). Web Traffic Time Series Forecasting Using LSTM Neural Networks with Distributed Asynchronous Training. *Mathematics*, 9(4), 421. https://doi.org/10.3390/math9040421

Darmawan, A. K., Alfarisi, S., & Hozairi, H. (2024). Predicting Gen-Z Personality On Twitter Based On Big Five Model With KNN and SVM. *Jurnal Pilar Nusa Mandiri*, 20(2), 137–145. https://doi.org/10.33480/pilar.v20i2.5580

Jain, S., Agrawal, S., Mohapatra, E., & Srinivasan, K. (2024). A novel ensemble ARIMA-LSTM approach for evaluating COVID-19 cases and future outbreak preparedness. *Health Care Science*, 3(6), 409–425. https://doi.org/10.1002/hcs2.123

Karunasingha, D. S. K. (2022). Root mean square error or mean absolute error? Use their ratio as well. *Information Sciences*, 585, 609–629. https://doi.org/10.1016/j.ins.2021.11.036

Kurniawan, J. D., Parhusip, H. A., & Trihandaru, S. (2024). Predictive Performance Evaluation of ARIMA and Hybrid ARIMA-LSTM Models for Particulate Matter Concentration. *Jurnal Online Informatika*, 9(2), 259–268. https://doi.org/10.15575/join.v9i2.1318

Ospina, R., Gondim, J. A. M., Leiva, V., & Castro, C. (2023). An Overview of Forecast Analysis with ARIMA Models during the COVID-19 Pandemic: Methodology and Case Study in Brazil. *Mathematics*, 11(14), 3069. https://doi.org/10.3390/math11143069

Pierre, A. A., Akim, S. A., Semenyo, A. K., & Babiga, B. (2023). Peak Electrical Energy Consumption Prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU Approaches. *Energies*, 16(12), 4739. https://doi.org/10.3390/en16124739

Rusman, J. A., Chunady, K., Makmud, S. T., Setiawan, K. E., & Hasani, M. F. (2023). Crude Oil Price Forecasting: A Comparative Analysis of ARIMA, GRU, and LSTM Models. *2023 IEEE 9th International Conference on Computing, Engineering and Design (ICCED)*, 1–6. https://doi.org/10.1109/ICCED60214.2023.10425576

Si, Y., Nadarajah, S., Zhang, Z., & Xu, C. (2024). Modeling opening price spread of Shanghai Composite Index based on ARIMA-GRU/LSTM hybrid model. *PLOS ONE*, 19(3), e0299164. https://doi.org/10.1371/journal.pone.0299164

Teixeira, M., Oliveira, J. M., & Ramos, P. (2024). Enhancing Hierarchical Sales Forecasting with Promotional Data: A Comparative Study Using ARIMA and Deep Neural Networks. *Machine Learning and Knowledge Extraction*, 6(4), 2659–2687. https://doi.org/10.3390/make6040128

Wu, S. S., Du, Z. H., Zhang, F., Zhou, Y., & Liu, R. Y. (2023). Time-Series Forecasting of Chlorophyll-a in Coastal Areas Using LSTM, GRU and Attention-Based RNN Models. *Journal of Environmental Informatics*. https://doi.org/10.3808/jei.202300494

Xiang, Y. (2024). Forecasting the NASDAQ Index Based on the ARIMA Model. *Advances in Economics, Management and Political Sciences*, 141(1), 206–213. https://doi.org/10.54254/2754-1169/2024.GA18869

Yavasani, R., & Wang, H. (2023). Comparative Analysis of LSTM, GRU, and ARIMA Models for Stock Market Price Prediction. *Journal of Student Research*, 12(4). https://doi.org/10.47611/jsrhs.v12i4.5888