Kim, Namhun; Cha, Junhee; Jeon, Junwoo

Article

A comparative evaluation of machine learning approaches for container freight rates prediction

Asian Journal of Shipping and Logistics (AJSL)

Provided in Cooperation with:
Korean Association of Shipping and Logistics, Seoul

Suggested Citation: Kim, Namhun; Cha, Junhee; Jeon, Junwoo (2025) : A comparative evaluation of machine learning approaches for container freight rates prediction, Asian Journal of Shipping and Logistics (AJSL), ISSN 2352-4871, Elsevier, Amsterdam, Vol. 41, Iss. 2, pp. 99-109, https://doi.org/10.1016/j.ajsl.2025.05.001

The Asian Journal of Shipping and Logistics 41 (2025) 99–109

Contents lists available at ScienceDirect

The Asian Journal of Shipping and Logistics

journal homepage: www.elsevier.com/locate/ajsl

A comparative evaluation of machine learning approaches for container freight rates prediction

Namhun Kim [a], Junhee Cha [a], Junwoo Jeon [b,*]

[a] Department of Business Administration, Sungkyul University, Republic of Korea
[b] Department of Global Logistics, Sungkyul University, Republic of Korea

A R T I C L E I N F O

Keywords:
Container Freight Rates
Machine learning
Decision Tree
Random Forest
LSTM
Prophet

A B S T R A C T

This study evaluates the predictive performance of four models—Decision Tree, Random Forest, Prophet, and LSTM—in forecasting container freight rates, a key metric for strategic decision-making in the shipping industry. To address data heterogeneity, Min-Max normalization was applied, and the Johansen co-integration test confirmed long-term relationships among the variables, justifying the use of raw data in our analysis. Performance was assessed using MSE, RMSE, NMSE, MAE, MAPE and SMAPE. While both Decision Tree and Random Forest models yielded lower absolute errors compared to LSTM and Prophet, the Decision Tree model demonstrated superior relative accuracy, outperforming Random Forest by approximately 91.8 % on the USWC route, 52.1 % on USEC, 43.5 % on MED, and 22.7 % on NEUR. These findings highlight the robustness of the Decision Tree model for container freight rate forecasting under volatile market conditions.

1. Introduction

The shipping industry is a cornerstone of global trade, responsible for transporting over 80 % of the world's cargo. This vital function underpins global economic growth and supply chain stability (Wang et al., 2024). In this context, the accurate prediction of ocean freight rates is critical. Freight rates not only reflect current market conditions but also serve as a mechanism for balancing supply and demand through strategic rate adjustments (Jeon et al., 2020; Scarsi, 2007; Schramm & Munim, 2021; Wang et al., 2024).

The importance of forecasting ocean freight rates is multifaceted. First, key stakeholders—including ship owners, cargo carriers, logistics companies, and end consumers—rely on these predictions for informed decision-making across various functions such as product pricing, cost accounting, financial management, and asset allocation (Hirata & Matsuda, 2022; Jeon et al., 2021). The volatility of freight rates is a fundamental component of maritime transport costs; therefore, precise forecasting of this volatility is crucial for market participants (Naima et al., 2023). Second, the inherent periodicity of ocean freight rates arises from the lengthy interval between ship orders and deliveries, which typically exceeds two years. This temporal lag reinforces cyclical patterns in freight rate dynamics (Scarsi, 2007). Third, ship owners make routine decisions regarding ship sales and charters based on prevailing freight rate levels, further emphasizing the need for robust forecasting (Jeon et al., 2020). Fourth, for shippers, accurate rate predictions are essential as logistics costs are directly correlated with freight rates—costs tend to escalate during periods of high rates and decline when rates are low. Additionally, seasonal variations, exemplified by the implementation of Peak Season Surcharges (PSS) in months such as March or October, further complicate the forecasting landscape (Yin & Shi, 2018). Finally, external shocks, including geopolitical risks and events such as the Suez Canal blockage during the COVID-19 pandemic and the recent Red Sea crisis, have demonstrated significant adverse impacts on container freight rates. Incorporating these disruptions into forecasting models has been shown to enhance predictive accuracy (Naima et al., 2023).

Moreover, the shipping industry is notably capital-intensive, with substantial financial commitments required for ship ordering and operation (Jeon & Yeo, 2017). Overcapacity, often resulting from aggressive shipbuilding, can precipitate a decline in freight rates, thereby intensifying the need for accurate market assessments and risk evaluations. For instance, using system dynamics, Jeon et al. (2020) identified a cyclical pattern of approximately 32 months in the China Container Freight Index (CCFI), highlighting the interplay between supply-demand imbalances. Additionally, stringent environmental regulations imposed by the International Maritime Organization (IMO),

* Corresponding author.
E-mail address: jwjeon@sungkyul.ac.kr (J. Jeon).

https://doi.org/10.1016/j.ajsl.2025.05.001
Received 21 March 2025; Accepted 6 May 2025
Available online 17 May 2025

such as the Energy Efficiency Design Index (EEDI), Energy Efficiency Ship Index (EEXI), and Carbon Intensity Indicator (CII), are compelling shipping companies to invest in eco-friendly vessels. These regulatory requirements further amplify the necessity for reliable freight rate forecasts to mitigate the risks associated with new ship orders (Bayraktar & Yuksel, 2023; Lee, 2024).

In light of these considerations, this study aims to compare the predictive performance of four distinct models—Prophet, Decision Tree, Random Forest, and LSTM—in forecasting container freight rates. The objective is to identify the model that most effectively captures the complex dynamics of ocean freight rates, thereby supporting more informed decision-making by industry stakeholders.

The remainder of this paper is organized as follows. Chapter 2 reviews relevant literature on container freight rate prediction and discusses prior applications of the Prophet, Decision Tree, Random Forest, and LSTM models. Chapter 3 details the data and methodological framework employed. Chapter 4 presents empirical findings and provides a comparative evaluation of the models. Chapter 5 discusses the implications of the findings, and Chapter 6 concludes the study.

2. Literature review

2.1. Forecasting studies on container freight rates

Container freight rates are critical indicators for a multitude of stakeholders within the shipping industry, influencing decisions ranging from pricing and financial management to asset allocation. Traditionally, econometric models—such as ARIMA, VAR, VEC, and ARMA—have been extensively employed to forecast these rates (Koyuncu & Tavacıoğlu, 2021; Munim & Schramm, 2017). For example, Schramm and Munim (2021) conducted comparative analyses between ARIMA and VAR, while Munim and Schramm (2021) extended this comparison to ARIMA and VEC. In another study, Chen et al. (2021) proposed an innovative hybrid approach by integrating empirical mode decomposition (EMD) with ARMA. Luo et al. (2009) further advanced the field by employing supply and demand dynamics for forecasting container freight rates.

Notwithstanding these contributions, the inherently non-linear and complex nature of container freight rate data presents substantial challenges for traditional econometric models (Hirata & Matsuda, 2022). This limitation has catalyzed a shift towards the adoption of machine learning and deep learning techniques, which are more adept at capturing non-linear patterns. Recent studies have demonstrated that deep learning models, such as LSTM, can outperform conventional approaches in certain contexts; for instance, Hirata and Matsuda (2022) reported superior predictive performance of LSTM over ARIMA for deep-sea routes, although results were comparable for short-sea routes. Moreover, Chen et al. (2024) showed that a CNN-LSTM hybrid model could achieve an R^2 exceeding 90 %, outperforming an ARIMA-SVR framework. Similarly, machine learning models such as Random Forest have been shown to deliver prediction accuracy above 80 % (Feng, 2022; Khan & Hussain, 2022). Complementary work by Jeon et al. (2021) compared ARIMA, VEC, and System Dynamics models, finding that the latter reduced prediction error by approximately 30 % on average, thereby offering distinct advantages in capturing market dynamics.

2.2. Studies on specific forecasting models

In addition to broad econometric approaches, various forecasting models have been applied across different domains. Their relevance to container freight rate prediction is elucidated below.

2.2.1. Decision tree models

Decision tree methodologies have proven effective in numerous forecasting applications. Liu et al. (2017a) applied decision trees for copper price prediction, achieving robust performance as measured by MAPE and RMSE. Bala (2010) extended this application to demand forecasting for Indian retailers, demonstrating that decision tree-based models outperformed other methods—such as ARIMA and SARIMA—over both short and long-term horizons. Hybrid models that combine decision trees with artificial neural networks (ANN) have further enhanced prediction accuracy (Chang, 2011; Tsai & Wang, 2009), reinforcing the versatility of decision tree frameworks.

2.2.2. Random forest models

Random Forest, an ensemble extension of decision trees, mitigates the risk of overfitting and enhances predictive stability. Liu and Li (2017b) utilized Random Forest to forecast gold price fluctuations, identifying critical predictors such as the DJIA and S&P 500 indices. Kumar and Thenmozhi (2006), demonstrated that Random Forest provided competitive accuracy relative to SVM, LDA, and logistic regression for predicting stock volatility. Subsequent studies by Abraham et al. (2022) and Vairagade et al. (2019) have consistently reported that Random Forest outperforms alternative methods in various forecasting contexts. Huertas Tato and Centeno Brito (2018), further validated these findings through applications in solar energy production forecasting, while Xue et al. (2021) compared multi-objective Random Forest variants, confirming its robustness in handling complex prediction tasks.

2.2.3. LSTM models

Long Short-Term Memory (LSTM) networks, a subset of recurrent neural networks (RNN), are particularly well-suited for time-series forecasting due to their ability to capture long-term dependencies. Bhandari et al. (2022) found that single-layer LSTM models provided superior predictive accuracy for S&P 500 closing prices when evaluated against RMSE, MAPE, and R^2 metrics. Conversely, Abbasimehr et al. (2020) reported that multi-layer LSTM models achieved even greater performance, outperforming conventional models such as ARIMA, ANN, and SVM. Sagheer and Kotb (2019) extended the LSTM architecture (DLSTM) for petroleum production forecasting, demonstrating notable improvements in RMSE and MAPE, while Siami-Namini et al. (2018) documented significant error reductions of 84–87 % in comparison to ARIMA models.

2.2.4. Prophet models

Prophet, developed by Taylor and Letham (2018), is designed for robust time-series forecasting by accommodating trends, seasonality, and external events. Empirical comparisons by Jha and Pande (2021) revealed that Prophet achieved lower RMSE and MAPE values compared to ARIMA in the context of supermarket sales forecasting. Similarly, Yenidoğan et al. (2018) applied both models to Bitcoin price prediction, reporting an R^2 of 0.94 for Prophet versus 0.68 for ARIMA. Kaninde et al. (2022) further demonstrated the utility of Prophet in volatile stock market forecasting, particularly due to its ability to integrate holiday effects and other exogenous factors. Recent external shocks, including the Suez Canal blockage during COVID-19 and the Red Sea crisis, underscore the importance of incorporating exogenous variables into forecasting models to enhance predictive accuracy (Naima et al., 2023; Wang et al., 2024).

2.2.5. Contributions

Despite the extensive body of work utilizing econometric models for container freight rate prediction, there remains a significant gap in the application of advanced machine learning and deep learning techniques to this domain. In particular, the relative underutilization of models such as Prophet, Random Forest, and LSTM—compared to traditional approaches—suggests a promising avenue for further research. Additionally, decision tree-based models have received limited attention in the context of container freight rate forecasting, despite demonstrated success in other fields. Therefore, the present study seeks to address these gaps by systematically comparing the performance of Prophet, Decision Tree, Random Forest, and LSTM models. This comparative analysis aims to elucidate the strengths and limitations of each approach and to identify the most effective methodology for capturing the complex, non-linear dynamics inherent in container freight rate data.

3. Methodology

3.1. Data

Ocean freight rates are influenced by various factors, including supply and demand dynamics, cargo weight and volume, and the distance to destination (Khan & Hussain, 2022). Recent geopolitical developments—such as the Red Sea crisis, which prompted container ships to bypass the Suez Canal in favor of the South African Cape of Good Hope—have further impacted these rates. Historical events, including the global financial crisis, the COVID-19 pandemic, and episodes of overcapacity, have also demonstrated significant effects on ocean freight rates (Naima et al., 2023; Wang et al., 2024).

To capture these dynamics, this study utilizes variables representing container shipping volume, container capacity, and key economic indicators. Specifically, Asia-Europe capacity and Asia-North America capacity are employed to quantify shipping capacity, while Asia-Europe shipping volume and Asia-North America shipping volume measure shipping volume. Ocean freight rates are examined along the North American, European, and Mediterranean Sea routes, as derived from the Shanghai Containerized Freight Index (SCFI). The economic indicators incorporated into the analysis include Global Economic Policy Uncertainty (EPU), the G20 Composite Leading Indicator (CLI), and Global Geopolitical Risk (GPR). EPU is calculated based on the frequency of terms such as "economy" and "policy" in media articles (Baker et al., 2016), CLI reflects expectations of future economic activity (OECD), and GPR quantifies geopolitical tensions based on news frequency (Caldara & Iacoviello, 2022).

This study utilizes monthly data spanning from January 2014 to June 2024. Due to the heterogeneous units of measurement (e.g., TEU versus USD/TEU), all features are normalized using a Min-Max scaling technique, which scales values to the [0,1] range. Subsequently, the Johansen co-integration test is performed to ascertain the existence of long-term relationships among the variables. The presence of co-integration justifies the use of the raw data in the forecasting models. Table 1 provides a summary of the descriptive statistics for the data used in this study.

3.2. Forecasting models

3.2.1. Prophet model

The Prophet model, introduced by Taylor and Letham (2018), is specifically designed for robust time-series forecasting. Its principal strength lies in its capacity to model seasonal patterns, long-term trends, and holiday effects independently, allowing for easy customization to specific business contexts. This flexibility enables Prophet to effectively capture the non-linear characteristics and periodic fluctuations inherent in ocean freight rate data.

3.2.2. Decision tree and random forest models

The Decision Tree model is a well-established method in predictive analytics, originally conceptualized by Belson (1959) and further developed through the CART methodology by Breiman et al. (1986). Although decision trees are intuitive and easy to interpret, they are susceptible to overfitting. To address this limitation, Random Forest—an ensemble learning technique proposed by Breiman (2001)—aggregates the predictions of multiple decision trees through a voting mechanism. This ensemble approach mitigates overfitting and enhances the model's performance, particularly when dealing with high-dimensional data.

3.2.3. LSTM model

Long Short-Term Memory (LSTM) networks, introduced by Hochreiter (1997), are a specialized type of recurrent neural network designed to overcome the limitations of traditional RNNs in capturing long-term dependencies. LSTM networks incorporate gating mechanisms—namely, the input, forget, and output gates—along with a memory cell that preserves and updates information as needed. These features address the vanishing gradient problem, thereby enabling the efficient learning of temporal dependencies and making LSTM particularly well-suited for time-series forecasting tasks such as predicting ocean freight rates.

3.2.4. Model settings

Random Search was employed to identify the optimal hyperparameter configurations for the Decision Tree, Random Forest, and Prophet models. In contrast, the LSTM model was trained using a consistent hyperparameter configuration across all sections, a design choice made to balance performance and computational demands. Table 2 presents the section-specific settings used for hyperparameter optimization.

3.2.5. Model validation

To assess the forecasting performance of the models, this study employs a comprehensive set of evaluation metrics: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Normalized Mean Squared Error (NMSE) and Symmetric Mean Absolute Percentage Error (SMAPE). MAE provides a direct measure of the average absolute error between the predicted and actual values. MAPE expresses this error as a

Table 1
Data descriptive statistics.

	Count	Mean	Std	Min	Max	50 %
EPU*	126	208.86	72.57	86.63	431.73	208.31
GPR* *	126	105.78	33.75	58.42	318.95	102.19
CLI* **	126	99.73	1.45	89.48	101.41	99.99
Capacity_EUR* ** *	126	1858,579	153,128.7	1559,476	2142,365	1860,088
Capacity_NA* ** *	126	2015,801	424,777.4	1462,072	2912,504	1876,666
Volume_EUR* ** *	126	1330,787	153,564.1	67,620	1616,700	1364,350
Volume_NA* ** *	126	1615,153	284,394.8	81,270	2119,400	1582,550
Med* ** **	126	2019.58	2064.47	220.50	7522.75	984.71
NEUR* ** **	126	1847.21	2055.13	223.50	7784.25	912.63
USWC* ** **	126	2587.77	1895.83	796.50	8079.00	1800.13
USEC* ** **	126	4091.98	2631.86	1589.50	2962.75	11,778.50

* Baker, S. R., Bloom, N., & Davis, S. J. (2016). "Measuring Economic Policy Uncertainty." Available at: Policy Uncertainty. / Unitless Index
* * Caldara, D and Iacoviello, M. (2022). "Global Policy Uncertainty." Available at: Matteo Iacoviello's Website. / Unitless Index
* **OECD. "Composite Leading Indicator (CLI) - G20." Available at: OECD. / Unit: Long-term average = 100
* ** *Bloomberg L.P. "Shipping Capacity and Shipping Volume Data." Bloomberg Terminal. / Unit: TEU
* ** **Shanghai Shipping Exchange. "Shanghai Containerized Freight Index." Available at: Shanghai Shipping Exchange. / Unit: USD/TEU, Unit: USD/FEU

Table 2
Section settings for hyperparameter optimization.

Decision Tree	Random Forest	Prophet	LSTM
Max Depth: 1–20 Min Samples Split: 10–100 Min Samples Leaf: 1–20 Max Features: 1–15 Max Leaf Nodes: 1–100 Min Weight Fraction Leaf: 0–1 Criterion: squared_error, friedman_mse, absolute_error, poisson	N Estimators: 1–100 Max Depth: 1–20 Min Samples Split: 10–100 Min Samples Leaf: 1–20 Max Features: 1–15 Max Leaf Nodes: 1–100 Min Weight Fraction Leaf: 0–1	Growth: Linear, Logistic Daily Seasonality: True, False Weekly Seasonality: True, False Yearly Seasonality: True, False Changepoint Prior Scale: 0.001–100(Log Scale) Seasonality Prior Scale: 0.001–100(Log Scale)	Sequence Length: 3 Dropout: 0.5 Units: 128 Epochs: 100 Batch Size: 64 Optimizer: Adam Learning Rate: 0.001 Loss: MSE

percentage, facilitating comparisons across models with different scales. MSE and RMSE, by emphasizing larger errors through squaring, offer insights into the models' sensitivity to significant deviations. NMSE normalizes the error relative to the variance of the data, allowing for a relative assessment of predictive performance. MAPE has a problem that the error rises infinitely when the actual value is close to 0, and SMAPE is used to compensate for this. SMAPE reflects the symmetry between the predicted and actual values and works stably near zero. Lower values across these metrics indicate better model performance, guiding the selection of the most effective forecasting approach for container freight rates.

4. Results

4.1. Johansen test

The Johansen co-integration test was conducted to ascertain the existence of a long-term equilibrium relationship among the variables. In this test, the null hypothesis posits that the number of co-integrating vectors is less than a specified rank ($n < r$), whereas the alternative hypothesis asserts that r is less than or equal to n. If the test statistic exceeds the critical value, the null hypothesis is not rejected; conversely, it is rejected when the test statistic falls below the critical value.

As summarized in Table 3, the test results confirm the presence of co-integration across all routes examined. Consequently, despite any non-stationarity in the individual series, the existence of a long-term relationship justifies the application of the raw data in subsequent modeling. Specifically, the number of co-integrations was determined to lie within range $2 < r \leq 3$ for both the USEC and USWC routes, and within $1 < r \leq 2$ for the MED and NEUR routes.

4.2. Model setting

All models utilizing normalized data were trained using 80 % of the dataset, with the remaining 20 % reserved for testing. The performance evaluations presented in Table 5 are based on models trained using the hyperparameter configurations detailed in Table 4. Specifically, Table 4

Table 3
Johansen test.

Route	r_0	r_1	Test Statistic	Critical Value (%)
USEC	0	6	125.2	95.75
	1	6	84.97	69.82
	2	6	51.54	47.85
	3	6	28.44	29.80
USWC	0	6	134.0	95.75
	1	6	87.06	69.82
	2	6	53.06	47.85
	3	6	27.28	29.80
MED	0	6	125.5	95.75
	1	6	75.16	69.82
	2	6	44.02	47.85
NEUR	0	6	129.0	95.75
	1	6	76.92	69.85
	2	6	45.93	47.85

Table 4
Settings of hyperparameter.

Route	Model	Hyperparameter
USWC	Decision Tree	random_state= 42, criterion= 'squared_error', max_depth= 19, max_features= 2, max_leaf_nodes= 68, min_samples_leaf= 15, min_samples_split= 2, min_weight_fraction_leaf= 0
	Random Forest	random_state= 42, max_depth= 6, max_features= 7, max_leaf_nodes= 81, min_samples_leaf= 1, min_samples_split= 16, min_weight_fraction_leaf= 0, n_estimators= 21, bootstrap=True, oob_score=False, n_jobs=None
	LSTM	Sequence Length= 3, Dropout= 0.5, Units= 128, Epochs= 100, Batch Size= 64, Optimizer=Adam, Learning Rate= 0.001, Loss=MSE
	Prophet	growth= 'logistic', changepoint_prior_scale= 0.0081, seasonality_prior_scale= 572.237, yearly_seasonality=True, daily_seasonality=False, weekly_seasonality=False
USEC	Decision Tree	random_state= 42, criterion= 'squared_error', max_depth= 15, max_features= 13, max_leaf_nodes= 27, min_samples_leaf= 13, min_samples_split= 52, min_weight_fraction_leaf= 0
	Random Forest	random_state= 42, max_depth= 6, max_features= 3, max_leaf_nodes= 96, min_samples_leaf= 11, min_samples_split= 21, min_weight_fraction_leaf= 0, n_estimators= 26, bootstrap=True, oob_score=False, n_jobs=None
	LSTM	Sequence Length= 3, Dropout= 0.5, Units= 128, Epochs= 100, Batch Size= 64, Optimizer=Adam, Learning Rate= 0.001, Loss=MSE
	Prophet	growth= 'logistic', changepoint_prior_scale= 0.0035, seasonality_prior_scale= 0.0013, yearly_seasonality=True, daily_seasonality=False, weekly_seasonality=False
MED	Decision Tree	random_state= 42, criterion= 'squared_error', max_depth= 19, max_features= 9, max_leaf_nodes= 74, min_samples_leaf= 6, min_samples_split= 16, min_weight_fraction_leaf= 0
	Random Forest	random_state= 42, max_depth= 7, max_features= 14, max_leaf_nodes= 38, min_samples_leaf= 10, min_samples_split= 18, min_weight_fraction_leaf= 0, n_estimators= 39, bootstrap=True, oob_score=False, n_jobs=None
	LSTM	Sequence Length= 3, Dropout= 0.5, Units= 128, Epochs= 100, Batch Size= 64, Optimizer=Adam, Learning Rate= 0.001, Loss=MSE
	Prophet	growth= 'logistic', changepoint_prior_scale= 0.0081, seasonality_prior_scale= 0.6136, yearly_seasonality=True, daily_seasonality=False, weekly_seasonality=False
NEUR	Decision Tree	random_state= 42, criterion= 'squared_error', max_depth= 19, max_features= 3, max_leaf_nodes= 70, min_samples_leaf= 3, min_samples_split= 27, min_weight_fraction_leaf= 0
	Random Forest	random_state= 42, max_depth= 13, max_features= 12, max_leaf_nodes= 76, min_samples_leaf= 2, min_samples_split= 40, min_weight_fraction_leaf= 0, n_estimators= 57, bootstrap=True, oob_score=False, n_jobs=None
	LSTM	Sequence Length= 3, Dropout= 0.5, Units= 128, Epochs= 100, Batch Size= 64, Optimizer=Adam, Learning Rate= 0.001, Loss=MSE
	Prophet	growth= 'logistic', changepoint_prior_scale= 0.0023, seasonality_prior_scale= 0.0327, yearly_seasonality=True, daily_seasonality=False, weekly_seasonality=False

Table 5
Evaluation indicators of models.

Route	indicator	Decision Tree	Random Forest	LSTM	Prophet
USWC	MSE	0.0778	0.0322	3788.1917	0.2171
	RMSE	0.2790	0.1796	61.5482	0.4659
	NMSE	0.8638	0.5853	75024.7661	1.7127
	MAE	0.1615	0.1007	54.7361	0.4194
	MAPE	36.9699	448.6831	39944.7295	330.4461
	SMAPE	50.0792	47.3774	83.4179	73.5523
USEC	MSE	0.0249	0.0310	3873.1899	0.2420
	RMSE	0.1579	0.1761	62.2349	0.4920
	NMSE	0.3618	0.4559	56941.6229	2.3690
	MAE	0.1007	0.1146	55.1988	0.4235
	MAPE	53.8369	112.3587	87888.3997	365.6392
	SMAPE	42.1499	55.9611	76.8371	88.9091
MED	MSE	0.0427	0.0347	4025.4914	0.1472
	RMSE	0.2068	0.1863	63.4467	0.3837
	NMSE	0.5136	0.4359	136494.9406	1.6701
	MAE	0.1181	0.1265	56.1486	0.3487
	MAPE	45.8080	81.1313	23215.9150	152.6911
	SMAPE	43.2302	55.3836	41.5510	71.0931
NEUR	MSE	0.0632	0.0447	3363.9074	0.1842
	RMSE	0.2514	0.2115	57.9992	0.4292
	NMSE	0.9562	0.5434	40844.9406	1.9608
	MAE	0.1559	0.1297	52.3133	0.3877
	MAPE	81.4221	105.2359	38170.2764	342.0722
	SMAPE	66.1932	60.7961	68.7988	93.0079

outlines the optimized hyperparameter settings for each of the four models, while Table 5 compares the predictive performance of these models following training with the specified configurations.

4.3. Model comparison

The predictive performance of the four forecasting models—Decision Tree, Random Forest, LSTM, and Prophet—was rigorously evaluated using six metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Normalized Mean Squared Error (NMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Symmetric Mean Absolute Percentage Error (SMAPE). Table 5 presents a detailed summary of these evaluation indicators for each shipping route.

Overall, the Decision Tree and Random Forest models demonstrated superior performance compared to the LSTM and Prophet models when assessed using absolute error metrics (MSE, RMSE, NMSE, and MAE). This observation suggests that tree-based methods are more effective in capturing the underlying non-linear dynamics of container freight rate data. The lower error values achieved by these models indicate their enhanced capability to model the complex relationships present in the dataset.

For the USWC route, the Random Forest model achieved the lowest absolute error values, recording an MSE of 0.0322, an RMSE of 0.1796, an NMSE of 0.5853, and an MAE of 0.1007. Specifically, the MSE obtained by the Random Forest model was approximately 58.6 % lower than that of the Decision Tree, 85.2 % lower than that of the Prophet model, and 99.15 % lower than that of the LSTM model. These improvements in error metrics underscore the robust performance of the Random Forest approach in capturing the dynamics of the USWC route.

In contrast, for the USEC route, the Decision Tree model exhibited the most favorable performance. The Decision Tree recorded an MSE of 0.0310, an NMSE of 0.4559, an RMSE of 0.1579, and an MAE of 0.1007, outperforming the Random Forest model by a modest margin. Specifically, its MSE was about 2.9 % lower than that of the Random Forest, approximately 87.2 % lower than that of the Prophet model, and nearly 99.2 % lower than that of the LSTM model. This result indicates that, for the USEC route, the Decision Tree model more effectively captures the underlying data structure.

For the MED route, while the Random Forest model generally achieved lower error metrics—with an MSE of 0.0347, an RMSE of 0.1863, and an NMSE of 0.4359—the Decision Tree model yielded the lowest MAE at 0.1181. The performance differences between these models, although small, suggest that each model may have strengths in capturing different aspects of the data variability for the MED route.

On the NEUR route, the Random Forest model again recorded the lowest absolute error values, with an MSE of 0.0447, an RMSE of 0.2115, an NMSE of 0.5434, and an MAE of 0.1297. The Random Forest's performance on this route was approximately 29.3 % better in terms of MSE, 15.9 % better in RMSE, and 30.5 % better in NMSE compared to the Decision Tree model. Nonetheless, the differences in performance metrics between the two models on the NEUR route were relatively marginal.

A comparison between MAPE and SMAPE reveals that SMAPE consistently demonstrates better performance across all routes. This suggests that many of the actual values used in the MAPE denominator are close to zero, thereby inflating the MAPE error. To address this issue, both MAPE and SMAPE were jointly examined, and models exhibiting relatively stable values across both metrics were selected. Based on this comprehensive evaluation, the Decision Tree and Random Forest models generally outperformed the LSTM and Prophet models across most routes.

Despite the favorable outcomes for both the Decision Tree and Random Forest models with respect to absolute error metrics, the relative error measure—MAPE and SMAPE—revealed considerably higher prediction error ratios across all models. This elevated MAPE and SMAPE is largely attributable to the increased volatility in container freight rates following disruptions such as the COVID-19 pandemic and the Red Sea crisis. Elevated MAPE and SMAPE values across all models are largely attributable to the increased volatility in container freight rates following disruptions such as the onset of the COVID-19 pandemic and the Red Sea crisis, which amplified global supply chain disturbances and resulted in larger percentage errors, as illustrated in Fig. 1. Comparing SMAPE, it was analyzed that USEC and MED have good decision tree performance, and USWC and NEUR have good Random Forest performance. Notably, a comparison of the MAPE values across the routes indicates that the Decision Tree model consistently exhibits lower relative error ratios than the Random Forest model. Specifically, the Decision Tree model achieved MAPE improvements of approximately 91.8 % for the USWC route, 52.1 % for the USEC route, 43.5 % for the MED route, and 22.7 % for the NEUR route when compared to the Random Forest model. Therefore, both MAPE and SMAPE adopted a relatively stable Decision Tree.

Finally, using the Decision Tree model, the influence of various features on container freight rates was visualized across all routes. SHAP (SHapley Additive Explanations) was employed to quantitatively assess the impact of these features on the machine learning prediction outcomes. The graphical representations of feature influence for the Random Forest, Prophet, and LSTM model predictions are provided in the Appendix. Specifically, both the Decision Tree and Random Forest models utilized SHAP for feature impact visualization, the Prophet model evaluated feature influence through regression coefficients, and the LSTM model illustrated influence based on the average weights of its feature set.

5. Discussion

The empirical analysis presented in Figs. 2–5 elucidates the multifaceted determinants of container freight rates, emphasizing the critical roles of both supply-side and demand-side factors. In particular, capacity and volume exert substantial influences on freight rate fluctuations across all examined routes. These findings reinforce the theoretical framework advanced in earlier studies (Jeon et al., 2020; Scarsi, 2007), which posit that shipping capacity and shipping volume are primary drivers of freight rate dynamics.

Beyond these fundamental supply-demand variables, our analysis highlights the pivotal role of the G20 CLI on routes other than USEC. As a

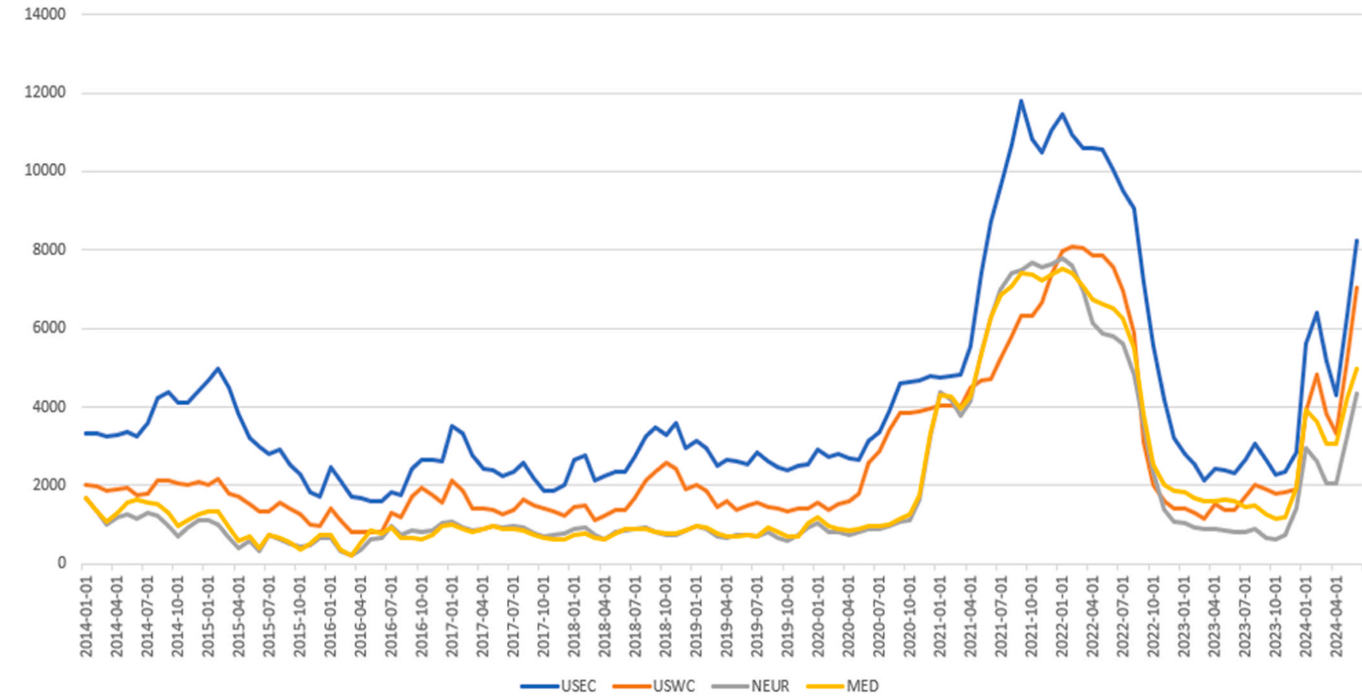

Fig. 1. SCFI.

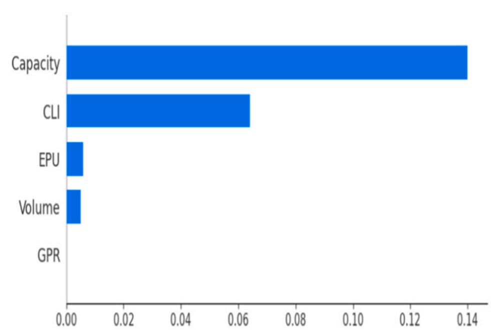

Fig. 2. The influence of feature (USWC).

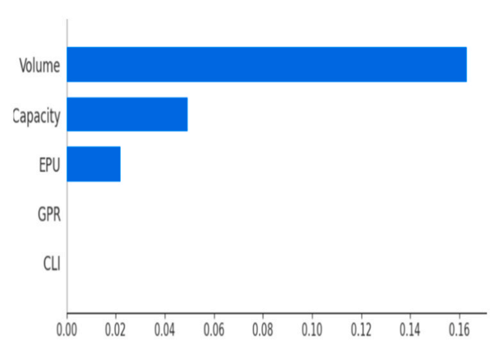

Fig. 3. The influence of feature (USEC).

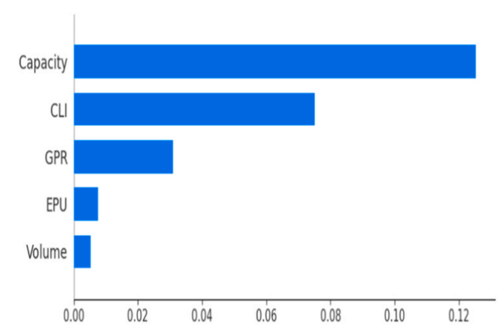

Fig. 4. The influence of feature (MED).

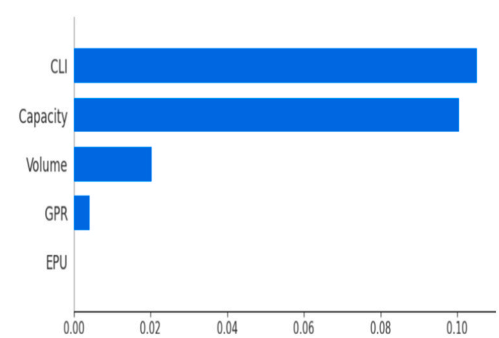

Fig. 5. The influence of feature (NEUR).

comprehensive measure that encapsulates key facets of economic activity—such as manufacturing performance, consumer confidence, and new order levels—the CLI serves as a robust proxy for assessing the overall economic environment. High CLI values typically signal an expanding economy, prompting shipping companies to increase capacity in anticipation of rising demand. Conversely, low CLI readings suggest economic stagnation or contraction, thereby incentivizing strategies such as capacity reduction or blank sailing. The strong correlation observed between CLI fluctuations and freight rate movements underscores its utility as a leading indicator for maritime transport planning.

The analysis also reveals that EPU significantly impacts container freight rates on most routes, with the notable exception of the NEUR route. In recent years, ongoing tariff disputes and anti-dumping investigations—particularly between Europe and the United States

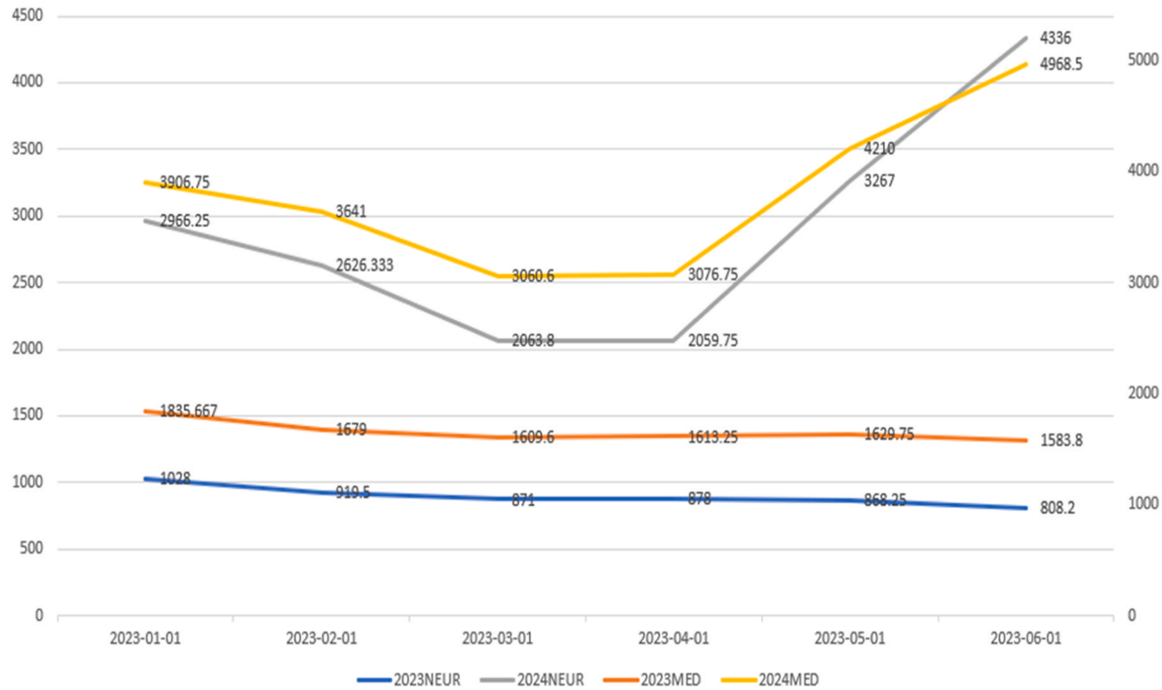

Fig. 6. MED SCFI and NEUR SCFI (2023.01 – 2023.06 / 2024.01 – 2024.06).

concerning imports from China—have elevated EPU levels. These conditions have led shippers to increase import volumes preemptively, which, in turn, have driven shipping companies to expedite the implementation of Peak Season Surcharges (PSS). A case in point is Maersk's decision to enforce PSS earlier than usual for shipments from the Asia Pacific region to North America and Canada, a response attributed to surging import volumes observed in early July 2024. The muted influence of EPU on the NEUR route is likely due to the stabilizing effect of the European Union's integrated economic framework, which buffers against policy-induced volatility.

Furthermore, the GPR indicator has demonstrated a pronounced impact on the MED and NEUR routes. The recent Red Sea crisis, characterized by the occupation of the Red Sea by Yemeni Houthi rebels starting in late 2023, led to significant disruptions in Suez Canal operations. This geopolitical instability forced shipping companies to reroute vessels via the longer Cape of Good Hope, thereby increasing transit distances and operational costs. Quantitative analysis, as depicted in Fig. 6, indicates that during the first half of 2024, container freight rates on the MED route increased by approximately 130 % relative to the same period in 2023, while the NEUR route experienced an increase of about 226 %. These findings highlight the critical impact of exogenous geopolitical shocks on freight rate volatility.

The discussion underscores that container freight rate determination is a complex interplay of inherent supply-demand dynamics and exogenous factors, including macroeconomic indicators and geopolitical risks. The integration of these diverse variables into our forecasting models not only enhances predictive performance but also offers significant insights for strategic decision-making in the maritime transport sector. Future research should further explore these interdependencies, particularly under conditions of heightened market volatility and geopolitical uncertainty, to develop more resilient and adaptive forecasting frameworks.

6. Conclusion

Container freight rates play a pivotal role in logistics cost management and strategic decision-making within the shipping industry. The cyclical nature of ship ordering—where high freight rates prompt new orders that are delivered two to three years later, potentially leading to overcapacity and subsequent rate declines—exacerbates the inherent market volatility. Moreover, the increasing regulatory pressures to reduce carbon emissions have driven shipping companies to invest in eco-friendly vessels. Consequently, accurate forecasting of ocean freight rates becomes essential for effective risk management in an industry characterized by long-term planning and significant capital investments.

This study undertook a comparative analysis of several forecasting models, including Decision Tree, Random Forest, LSTM, and Prophet, to determine their relative performance in predicting container freight rates. Our empirical results indicate that, in terms of absolute error metrics, both the Decision Tree and Random Forest models exhibit strong predictive capabilities. In this study, the model demonstrating the best overall performance was selected based on two evaluation approaches: (1) comparison of MSE, RMSE, NMSE, and MAE, and (2) comparison of MAPE and SMAPE. In the first evaluation, both the Random Forest and Decision Tree models exhibited strong performance, with Random Forest outperforming in some metrics. However, in the second evaluation using MAPE and SMAPE, the Decision Tree model was ultimately selected due to its relatively stable and superior performance across both indicators. These results highlight that the optimal model may vary depending on the chosen evaluation metric. Notably, the Decision Tree model emerged as the most effective when assessed using the relative error metric, MAPE and SMAPE, despite overall MAPE and SMAPE performance remaining suboptimal across all models. The elevated MAPE and SMAPE values can be attributed to the amplified impact of external shocks, particularly EPU and GPR, which have increasingly influenced market dynamics. This confirmation of the impact of EPU and GPR is expected to help the shipping industry stakeholders make decisions.

The findings of this study underscore the critical influence of both traditional supply-demand factors and external economic and geopolitical shocks on freight rate forecasting. However, the persistence of relatively high prediction error ratios highlights a significant limitation of the current modeling approaches. This suggests the need for the development of more sophisticated hybrid models—potentially integrating approaches such as ANN and RNN—that are better equipped to capture the complex, non-linear interactions between exogenous factors and freight rate dynamics.

Future research should focus on enhancing predictive accuracy by incorporating additional external variables and exploring ensemble methods that synergistically combine the strengths of various machine learning and deep learning models. Such advancements will be crucial for developing robust forecasting tools capable of supporting strategic decision-making in an increasingly volatile global shipping environment.

CRediT authorship contribution statement

Cha Junhee: Data curation, Formal analysis. **Kim Namhun:** Conceptualization, Data curation. **Jeon Junwoo:** Conceptualization, Formal analysis, Supervision.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Appendix

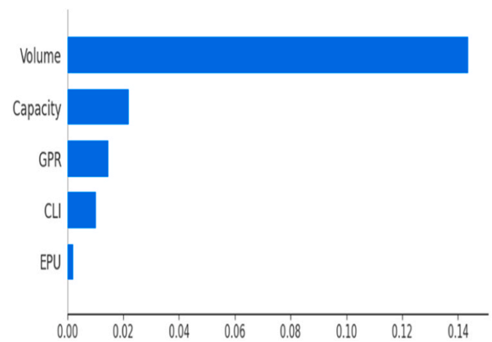

Fig. 7. The Influence of Feature (USWC), Random Forest

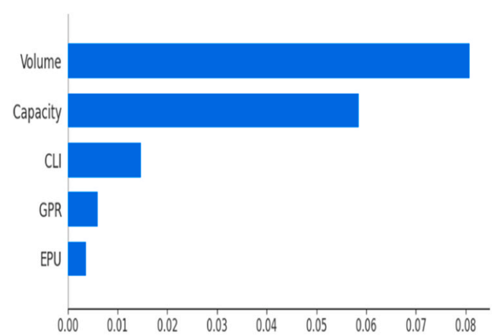

Fig. 8. The Influence of Feature (USEC), Random Forest

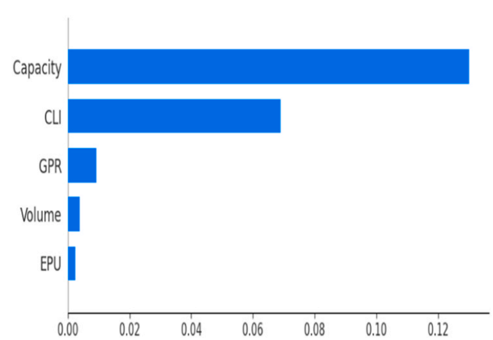

Fig. 9. The Influence of Feature (MED), Random Forest

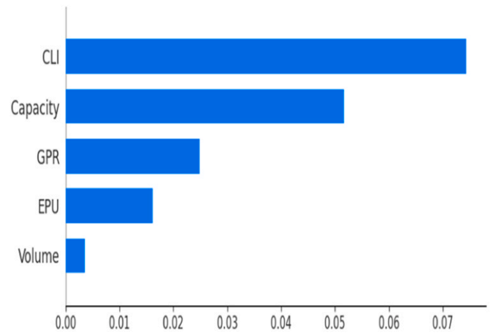

Fig. 10. The Influence of Feature (NEUR), Random Forest

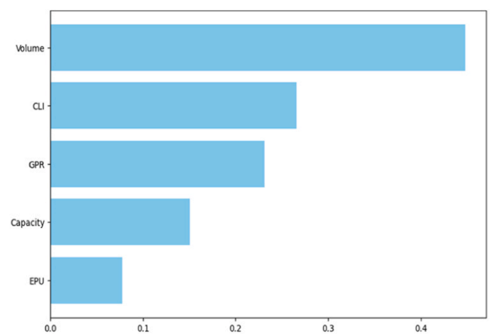

Fig. 11. The Influence of Feature (USWC), Prophet

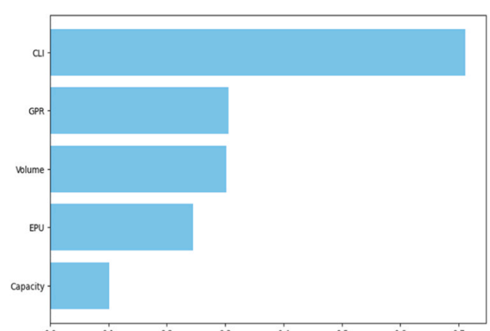

Fig. 12. The Influence of Feature (USEC), Prophet

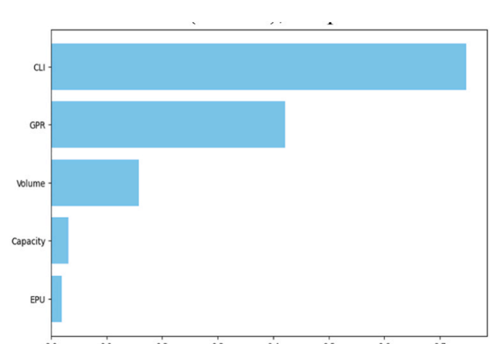

Fig. 13. The Influence of Feature (MED), Prophet

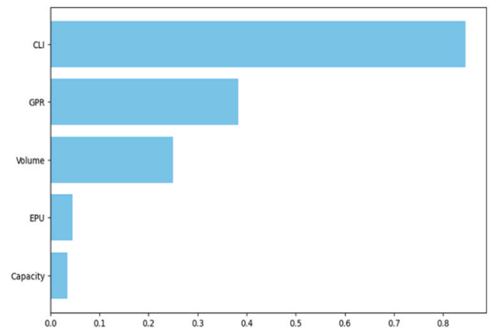

Fig. 14. The Influence of Feature (NEUR), Prophet

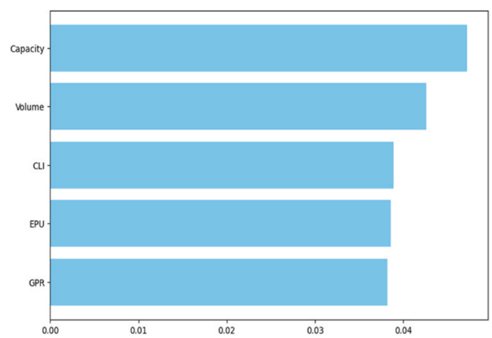

Fig. 15. The Influence of Feature (USWC), LSTM

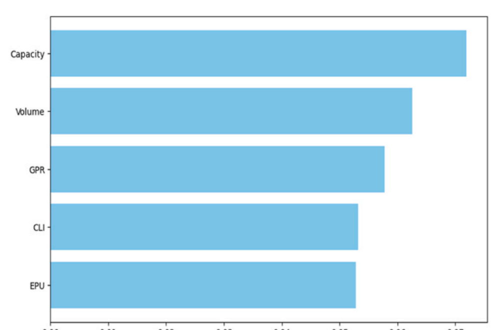

Fig. 16. The Influence of Feature (USEC), LSTM

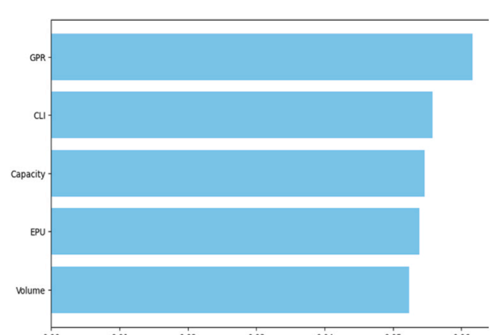

Fig. 17. The Influence of Feature (MED), LSTM

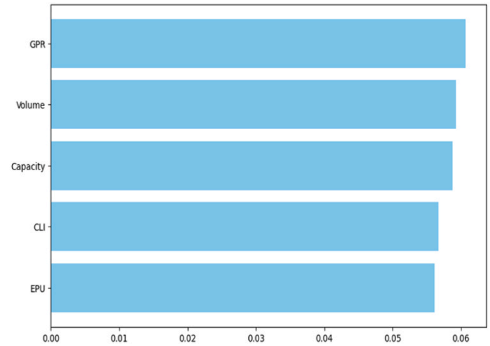

Fig. 18. The Influence of Feature (NEUR), LSTM

References

Abbasimehr, H., Shabani, M., & Yousefi, M. (2020). An optimized model using LSTM network for demand forecasting. *Computers Industrial Engineering, 143*, Article 106435.

Abraham, R., Samad, M. E., Bakhach, A. M., El-Chaarani, H., Sardouk, A., Nemar, S. E., & Jaber, D. (2022). Forecasting a stock trend using genetic algorithm and random forest. *Journal of Risk and Financial Management, 15*(5), 188.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics, 131*(4), 1593–1636.

Bala, P. K. (2010). Decision tree based demand forecasts for improving inventory performance (December). *2010 IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 1926–1930). IEEE (December).

Bayraktar, M., & Yuksel, O. (2023). A scenario-based assessment of the energy efficiency existing ship index (EEXI) and carbon intensity indicator (CII) regulations. *Ocean Engineering, 278*, Article 114295.

Belson, W. A. (1959). Matching and prediction on the principle of biological classification. *Journal of the Royal Statistical Society: Series C (Applied Statistics, 8*(2), 65–75.

Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R., & Khatri, R. K. (2022). Predicting stock market index using LSTM. *Machine Learning with Applications, 9*, Article 100320.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1986). Classification and Regression Trees.

Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review, 112*(4), 1194–1225.

Chang, T. S. (2011). A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction. *Expert Systems with Applications, 38* (12), 14846–14851.

Chen, L., Liu, J., Pei, R., Su, Z., & Liu, Z. (2024). Shanghai containerised freight index forecasting based on deep learning. *Methods: Evidence from Chinese Futures Markets East Asian Economic Review, 28*(3), 359–388.

Chen, Y., Liu, B., & Wang, T. (2021). Analysing and forecasting China containerized freight index with a hybrid decomposition–ensemble method based on EMD, grey wave and ARMA. *Grey Systems: Theory and Application, 11*(3), 358–371.

Feng, Y. (2022). Container freight rate level forecasting with machine learning methods. *International Journal of Business Performance and Supply Chain Modelling, 13*(3), 245–263.

Hirata, E., & Matsuda, T. (2022). Forecasting shanghai container freight index: A deep-learning-based model experiment. *Journal of Marine Science and Engineering, 10*(5), 593.

Hochreiter, S. (1997). Long short-term memory. *Neural Computation Mitosz-Press*.

Huertas Tato, J., & Centeno Brito, M. (2018). Using smart persistence and random forests to predict photovoltaic energy production. *Energies, 12*(1), 100.

Jeon, J. W., Duru, O., Munim, Z. H., & Saeed, N. (2021). System dynamics in the predictive analytics of container freight rates. *Transportation Science, 55*(4), 946–967.

Jeon, J. W., Duru, O., & Yeo, G. T. (2020). Modelling cyclic container freight index using system dynamics. *Maritime Policy Management, 47*(3), 287–303.

Jeon, J. W., & Yeo, G. T. (2017). Study of the optimal timing of container ship orders considering the uncertain shipping environment. *The Asian Journal of Shipping and Logistics, 33*(2), 85–93.

Jha, B.K., & Pande, S. (2021). Time series forecasting model for supermarket sales using FB-prophet. In 2021 5th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 547-554). IEEE.

Kaninde, S., Mahajan, M., Janghale, A., & Joshi, B. (2022). Stock price prediction using facebook prophet. In *ITM Web of Conferences, 44* p. 03060). EDP Sciences.

Khan, I. A., & Hussain, F. K. (2022). Regression Analysis Using Machine Learning Approaches for Predicting Container Shipping Rates (March). *International Conference on Advanced Information Networking and Applications* (pp. 269–280). Cham: Springer International Publishing (March).

Koyuncu, K., & Tavacıoğlu, L. (2021). Forecasting Shanghai containerized freight index by using time series models. *Marine Science and Technology Bulletin, 10*(4), 426–434.

Kumar, M., & Thenmozhi, M. (2006). Forecasting stock index movement: A comparison of support vector machines and random forest. In Indian institute of capital markets 9th capital markets conference paper.

Lee, S. S. (2024). Analysis of the effects of EEDI and EEXI implementation on CO2 emissions reduction in ships. *Ocean Engineering, 295*, Article 116877.

Liu, C., Hu, Z., Li, Y., & Liu, S. (2017a). Forecasting copper prices by decision tree learning. *Resources Policy, 52*, 427–434.

Liu, D., & Li, Z. (2017b). Gold price forecasting and related influence factors analysis based on random forest. *Proceedings of the tenth international conference on management science and engineering management* (pp. 711–723). Singapore: Springer.

Luo, M., Fan, L., & Liu, L. (2009). An econometric analysis for container shipping market. *Maritime Policy Management, 36*(6), 507–523.

Munim, Z. H., & Schramm, H. J. (2017). Forecasting container shipping freight rates for the Far East–Northern Europe trade lane. *Maritime Economics Logistics, 19*, 106–125.

Munim, Z. H., & Schramm, H. J. (2021). Forecasting container freight rates for major trade routes: a comparison of artificial neural networks and conventional models. *Maritime Economics Logistics, 23*, 310–327.

Naima, S., Su, N., Kevin, C., Victor, G., & Prem, C. (2023). Forecasting container freight rates using the Prophet forecasting method. *Transport Policy*, 133.

Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing, 323*, 203–213.

Scarsi, R. (2007). The bulk shipping business: market cycles and shipowners' biases. *Maritime Policy Management, 34*(6), 577–590.

Schramm, H. J., & Munim, Z. H. (2021). Container freight rate forecasting with improved accuracy by integrating soft facts from practitioners. *Research in Transportation Business Management, 41*, Article 100662.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series (December). *2018 17th IEEE international conference on machine learning and applications (ICMLA)* (pp. 1394–1401). Ieee (December).

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician, 72*(1), 37–45.

Tsai, C.F., & Wang, S.P. (2009). Stock price forecasting by hybrid machine learning techniques. In Proceedings of the international multiconference of engineers and computer scientists (Vol. 1, No. 755, p. 60).

Vairagade, N., Logofatu, D., Leon, F., & Muharemi, F. (2019). Demand forecasting using random forest and artificial neural network for supply chain management. *Computational Collective Intelligence: 11th International Conference, ICCCI 2019, Hendaye, France, September 4–6, 2019, Proceedings, Part I 11* (pp. 328–339). Springer International Publishing.

Wang, W., He, N., Chen, M., & Jia, P. (2024). Freight rate index forecasting with Prophet model based on multi-dimensional significant events. *Expert Systems with Applications, 249*, Article 123451.

Xue, L., Liu, Y., Xiong, Y., Liu, Y., Cui, X., & Lei, G. (2021). A data-driven shale gas production forecasting method based on the multi-objective random forest regression. *Journal of Petroleum Science and Engineering, 196*, Article 107801.

Yenidoğan, I., Çayir, A., Kozan, O., Dağ, T., & Arslan, Ç. (2018). Bitcoin forecasting using ARIMA and PROPHET (September). *2018 3rd international conference on computer science and engineering (UBMK)* (pp. 621–624). IEEE (September).

Yin, J., & Shi, J. (2018). Seasonality patterns in the container shipping freight rate market. *Maritime Policy Management, 45*(2), 159–173.