Volume 6, Issue 1, June 2026 - ISSN 2735-5969
http://dx.doi.org/10.21622/IBL.2026.06.1.1795

Machine Learning-Driven Prediction of Cost-Efficient Shipping Lines: Evidence from Freight Forwarding Operations

Muhammad Aref Arfeen[1], Nermine Khalifa[2] and Hatem Abdelkader[3]

[1,2]Arab Academy for Science, Technology & Maritime Transport, Egypt
[3]Menoufia University, Egypt

arfeen@aast.edu, Nermine_khalifa@egypt.aast.edu, hatem6803@yahoo.com

Received 14 November 2025 **Accepted 08 January 2026** **Published 05 May 2026**

Abstract

Purpose: *This research evaluates the performance of various machine learning models in forecasting cost-efficient transportation lines from the viewpoint of freight forwarders. It addresses the requirement for data-informed decision-support tools that can enhance the precision of shipment line selection.*

Design/Methodology/Approach: A dataset comprising 983 shipment records from 37 freight-forwarding companies in Egypt was examined. Six machine learning models, Naive Bayes, K-Nearest Neighbour, Support Vector Machines, Decision Trees, Random Forests, and Neural Networks, were trained and assessed employing an 80/20 train-test partition. The performance of the model was evaluated utilizing accuracy, precision, recall, and F1-score.

Findings: Random Forests and Decision Trees achieved the finest predictive performance, with accuracy scores of 0.83 and 0.80, respectively. K-Nearest Neighbour exhibited moderate performance, whereas Naive Bayes and Neural Networks demonstrated comparatively lower predictive accuracy. Support Vector Machines demonstrated suboptimal performance across all evaluated metrics. The findings suggest that ensemble and tree-based methodologies are highly appropriate for modeling the cost efficacy of shipping lines.

Research Implications/Limitations: The research emphasizes the significance of model selection in logistics forecasting tasks and demonstrates that more complex or computationally demanding models do not necessarily surpass the performance of simplified tree-based approaches. The dataset is restricted to shipments originating from Egypt in 2022, which may impact its overall generalizability.

Practical Implications: The findings provide freight forwarders with a foundation for incorporating machine learning into their operational decision-support systems. Utilizing Random Forests or Decision Trees can improve the precision of shipping line selection, lower expenses, and facilitate data-driven logistics planning.

Originality: The study offers a comparative analysis of various machine learning methods specifically applied to the selection of transportation lines, a subject with limited prior exploration. It provides empirical evidence on the predictive models that most effectively aid in making freight-forwarding decisions related to cost efficiency.

Keywords: *Cost-effective shipping lines; Decision Trees; Freight forwarding; Machine learning; Predictive modelling; Random Forests; Shipping line selection*

Volume 6, Issue 1, June 2026 - ISSN 2735-5969
http://dx.doi.org/10.21622/IBL.2026.06.1.1795

1. Introduction

In the global logistics network, freight forwarding is essential. Establishing a connection between the shipper and the carrier guarantees that international cross-border shipments reach their destinations effectively and efficiently, that is, in a timely and economical manner (Ho *et al.*, 2017; Dzakah Fanam, Nguyen, and Cahoon, 2018). Selecting the best shipping line is a crucial strategic choice that forwarders must make for every cargo in this dynamic global environment, where there are numerous routes from point A to point B. This will guarantee high operational efficiency and a competitive edge. When making such a crucial choice, freight forwarders need to consider many aspects, including carrier reliability, transit times, and service charges (Ho *et al.*, 2017; Dzakah Fanam, Nguyen, and Cahoon, 2018; Ergin and Alkan, 2023).

The logistics and supply chain management field is now reaping the benefits of machine learning (ML), a key emerging technology with the potential to reveal complex patterns within large datasets and forecast the best logistical choices. It has a variety of models, and the field has not yet settled on which one or ones work best for predicting cost-effective shipping routes. This research examines six prominent machine learning models: Naive Bayes, K-Nearest Neighbor (KNN), Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks, and evaluates their comparative effectiveness in predicting cost-effective shipping line selection. These models were selected to represent probabilistic, distance-based, margin-based, tree-based, ensemble, and non-linear learning paradigms.

The Egyptian freight market in 2022 was very unstable on a macroeconomic level, with currency devaluations, changing freight rates, and limits on foreign currency. This made it impossible to choose a shipping line based on traditional heuristics, so machine learning-based decision-support systems had to be put in place. While evidence exists regarding the effectiveness of machine learning in various logistics applications, the existing literature lacks comprehensive research assessing the predictive capabilities of ML algorithms for real-time, cost-efficient shipping line selection from the perspective of freight forwarders. This study addresses the gap by employing diverse machine learning models to identify the principal cost factors influencing freight forwarders' shipping line selection and to precisely forecast optimal decisions, thereby providing the data-driven framework essential for practitioners to make systematic choices in unstable contexts.

The goal of this study is to determine how well various machine learning models perform in terms of prediction. From the perspective of a freight forwarder, these forecasts relate to integrating economical shipping line options into a logistics decision-support system.

The following research question serves as the study's guide: Which machine learning models are the most effective at predicting cost-effective shipping lines?

2. Related Work

Related works on this topic have two research streams. The following sections review the related works on freight forwarding. The next stream investigates the research on using machine learning in freight forwarding. Finally, a shipping line selection from the perspective of a freight forwarder.

2.1 Freight Forwarding

Freight forwarding services play a vital role in the international trade and commerce setup, allowing for the movement of goods from one place to another with precision, efficiency, and cost-effectiveness. Freight forwarding services. According to the International Federation of Freight Forwarders Associations (FIATA):

"Freight Forwarding Services means services of any kind relating to the carriage, consolidation, storage, handling, packing or distribution of the Goods as well as ancillary and advisory services in connection therewith, including but not limited to customs and fiscal matters, declaring the Goods for official purposes, procuring insurance of the Goods and collecting or procuring payment or documents relating to the Goods." (FIATA, 2019)

The services of a freight forwarder include the following: the management of shipping routes, customs operations, documentation, storage, and distribution of goods. These intermediaries facilitate the link between shippers and carriers, promoting seamless and timely movement of products across international borders (Ding *et al.*, 2017; Huang, Bulut, and Duru, 2019).

Freight forwarding requires a deep understanding of global trade complexities, including navigating various international regulations and procedures. Research indicates that freight forwarders significantly contribute

Volume 6, Issue 1, June 2026 - ISSN 2735-5969
http://dx.doi.org/10.21622/IBL.2026.06.1.1795

to optimizing supply chains by selecting the most efficient shipping routes and facilitating smooth cross-border trade (Rajasekar and Prabhakar, 2015).

Their role is important in the liner shipping industry because they represent the most frequent customer of the shipping lines. Hence, their criteria for carrier selection are paramount in defining winners and losers in the marketplace. Typically, forwarders would pay special attention to various elements when picking shipping lines: freight rates, adherence to schedules, reliability in service, door-to-door services, and environmental performance (Dzakah Fanam, Nguyen, and Cahoon, 2018).

2.2 Machine Learning in Freight Forwarding

Machine learning enables systems to improve performance by learning from past data. This added sophistication in data analytics is through the kind of algorithms that learn from patterns, adjust with new data, and keep perfecting their predictions. This is beneficial for handling large datasets or situations in which the data is too complex for conventional statistical methods (Rao, Chandra, and Kumar, 2017).

The use of machine learning in logistics and freight forwarding has greatly transformed these sectors. Machine learning empowers systems to learn from data, thus improving predictive accuracy over time. This combination is highly effective at solving complex logistic problems; among them, there are predicted demand (Ermagun *et al.*, 2020), optimizing shipping routes (Cao, 2022), and reducing operational costs (Gkerekos, Lazakis, and Theotokatos, 2019).

The studies (Cao, 2022) and (Sert *et al.*, 2020) also found that the combination of data analytics and machine learning in logistics for task automation, even in such simple tasks as tracking of shipments and route planning. This not only increases efficiency but also raises the satisfaction level among the clients. It was found that companies implementing machine learning models can reduce delivery time and operational costs by real-time insights into automation of decision-making and freight forwarder performance.

The table delineates global machine learning applications in freight forwarding. It highlights both established research domains and notable knowledge deficiencies. Studies originate from developed economies like the USA, the Netherlands, and Portugal, with scant contributions from emerging markets, including Indonesia and Poland. Notably, there exists a substantial gap in research on critical logistics hubs and developing economies, such as China, India, North Africa, and the Middle East. In Table 1, different machine learning techniques are presented, encompassing options like Naive Bayes and linear regression, together with Bayesian Nonparametric structures. The research covers various aspects of freight forwarding activities, including risk assessment, fraud detection, personnel performance forecasting, traffic management, and cost and time evaluations.

Table 1. Machine Learning Applications in Freight Forwarding

#	Paper	Application Country	ML Technique	Prediction Area
1	(Shang, Dunson, and Song, 2017)	Global	Bayesian Nonparametric (Probit Stick-Breaking Process mixture model)	Risk Assessment in Air Cargo Transport
2	(Cyperski, Okulewicz and Domański, 2024)	Poland	Hybrid Approach: DBSCAN, k-NN, XGBoost	Cost Estimation for Full Truckload (FTL) Shipping
3	(Triepels, Feelders and Daniels, 2015)	Netherlands	Probabilistic Classification: Naive Bayes (NB), Tree-Augmented Naive Bayes (TAN)	Document Fraud Detection in Maritime Freight
4	(Calixto and Ferreira, 2020)	Portugal	Naive Bayes Classifier	Salespeople Performance
5	(Hathikal, Chung, and Karczewski, 2020)	USA	Multinomial Logistic Regression, Decision Tree, KNN, SVM	Shipment Lead Time for Ocean Freight
6	(Wahyudi and Septya Arroufu, 2022)	Indonesia	Linear Regression	Delivery Time Prediction
7	(Bridgelall, 2024)	USA	Data Mining and GIS-based Analysis	Freight Traffic Reduction
8	(Birkel, Kopyto, and Lutz, 2020)	Germany	NA	Capacity forecasting, resource allocation
9	(Jang, Chang, and Kim, 2023)	Republic of Korea	LightGBM, XGBoost, DNN, MLR	Cost prediction for freight shipping

Taken together, Table 1 highlights a pronounced empirical gap in machine learning applications for freight forwarding within developing economies, particularly in volatile markets such as Egypt, a gap that the present study directly addresses by providing evidence from a real-world 2022 Egyptian freight forwarding dataset.

2.3 Shipping Line Selection

To achieve the best possible results, freight forwarders must perform a careful evaluation of the shipping lines they utilize. They consider many aspects of a shipping company over which that company has direct control. First, and most important, is the aspect of price. The freight rates a shipping line offers directly impact both the forwarder and the forwarder's customer. The second major aspect, service reliability, is equally crucial. Price advantages can mean little to a forwarder and its customer if the service that the forwarder is paying for is unreliable. Service reliability encompasses both the ship's performance (arriving on time and in good condition) and the performance of the shipping line's staff at both ends of the journey. After price and service reliability comes the matter of transit time. These factors significantly affect forwarder efficiency and client satisfaction (Lukinskiy and Lukinskiy, 2015; Ho *et al.*, 2017; Dzakah Fanam, Nguyen, and Cahoon, 2018; Ergin and Alkan, 2023).

3. Methods

3.1 Data Description

The dataset used in this study includes 983 shipment records collected from 37 freight forwarding businesses operating in Egypt, encompassing a wide range of shipment characteristics such as service cost, cargo type, port of loading, port of discharge, importer country, and shipping line. The data covers the year 2022 and includes significant elements crucial to the decision-making process for choosing shipping lines.

Table 2 lists these features and describes them.

Table 2. Summary of Dataset Features and Descriptions

Feature	Description	Type
Cargo	Type of cargo being transported	Qualitative
Industry	Industry associated with the cargo	Qualitative
Port of Loading	Departure port for the cargo	Qualitative
Port of Discharging	Arrival port for the cargo	Qualitative
Importers Countries	Countries where the importers are located	Qualitative
Service Cost (USD)	Cost of service per tonnage in USD	Quantitative
Shipping Line	Name of the shipping company	Qualitative

The data cleaning process involved handling missing values, removing outliers, and standardizing the service cost to a common currency using the Central Bank of Egypt's exchange rates. Additionally, new features such as Service Cost per Tonnage were engineered to improve model accuracy.

3.2 Model Training and Evaluation

Six machine learning models, Naive Bayes, K-Nearest Neighbor, Support Vector Machines, Decision Trees, Random Forests, and Neural Networks, were trained and evaluated using an 80/20 train–test split. The following evaluation metrics were used to compare model performance:

- Accuracy: The expression of the ratio of correctly predicted observations to total observations (Tohka and van Gils, 2021).
- Precision: The ratio between true positive predictions and the total amount of positive predictions (Hicks *et al.*, 2022).
- Recall: The ratio between correctly classified positive samples and all samples assigned to the positive class (Tohka and van Gils, 2021; Hicks *et al.*, 2022).
- F1-Score: The harmonic mean of precision and recall, providing a balanced measure of model performance (Tohka and van Gils, 2021; Hicks *et al.*, 2022).

4. Results

This study's main goal is to assess how well various machine learning models perform when it comes to predicting shipping selections from a cost perspective for a freight forwarder. The models used were Naive Bayes, K-Nearest Neighbour, Support Vector Machines, Decision Trees, Random Forests, and Neural Networks. They were trained on a dataset of 983 shipment records, and then their performances were compared using four key metrics: accuracy, precision, recall, and F1-score.

Volume 6, Issue 1, June 2026 - ISSN 2735-5969
http://dx.doi.org/10.21622/IBL.2026.06.1.1795

The figure below shows the performance metrics of each model:

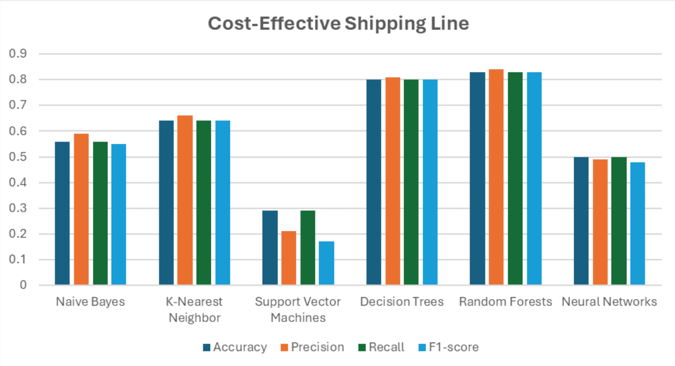

Figure 1: Performance metrics for Cost-Effective Shipping Line Prediction

4.1 Naive Bayes Performance

The Naive Bayes model has a consistent but moderate performance across all metrics in predicting shipping lines that are cost-effective. Its accuracy is just 0.56, which amounts to a correct prediction in only 56% of cases, indicating only a fair ability to make accurate predictions. Its precision score of 0.59 indicates that it's not the class of cost-effective shipping lines that has been identified as cost-effective; rather, it's nearly 41% of the time that we've encountered false positives. Finally, when we look at recall, we can capture only a moderate proportion of the actual cost-effective shipping lines. Combined with its F1-score of 0.55, which represents a balance of precision and recall, the model provides an average result. In summary, Naive Bayes ranks as one of the least effective models within this analysis. Its inability to deliver high accuracy and reliable predictions makes it unsuitable for supporting high-quality logistics decision-making, particularly in contexts where minimizing errors is critical.

4.2 Support Vector Machines Performance

The Support Vector Machines (SVM) Method proved less efficient with values oscillating from 0.17 to 0.29. Its accuracy of 0.29 means that we got the predictions correct only in 29% of the cases. This came with a precision score of 0.21, implying that the false positive rate is 79%, an alarming figure, along with an abysmal F1-score of 0.17, which again consolidates its ineffectiveness overall. This analysis indicates that performance is extremely bad with the SVM model; all metrics, particularly accuracy, precision, recall, and F1-score, are significantly less than 0.5, which is a clear indication of very poor performance. It indicates that SVM struggles very much to distinguish low-cost shipping lines, so it misclassifies them a lot. These low metrics indicate that SVM is not appropriate for this logistics application because its predictions misclassify too many shipping line costs.

4.3 K-Nearest Neighbor (KNN) Performance

In our assessment of economic shipping lines, the K-Nearest Neighbor (KNN) algorithm surpassed Support Vector Machines. With an accuracy of 64%, KNN uses a distance-based approach that is more suitable for this dataset, but it also shows potential for enhancements. KNN's precision is 0.66, which means that we can have a certain level of trust in about two-thirds of the identified cost-effective shipping lines, certainly an improvement over Support Vector Machines, but still risky in terms of false positives. KNN's recall is 0.64; it recognized two-thirds of the true cost-effective shipping lines, which I suspect also represents a slight improvement over the recall of our previous model.

The KNN algorithm is good at identifying the correct class for positive instances, albeit it does include some false positives in the result. Its F1-score of 0.64 indicates that the KNN algorithm has a moderate balance of precision and recall. It is a good algorithm to use in contexts where we are slightly okay with identifying some negative instances as positive (false positives) and identifying some positive instances as negative (false negatives). It has an intermediate performance across all the metrics we considered in our analysis, making it a reasonable choice for solving uncomplicated, low to medium complexity logistics decisions. On the other hand, KNN is sensitive to data scaling and not the best algorithm if you put priority on computational requirements.

4.4 Decision Trees Performance

The Decision Tree model demonstrated strong predictive capability by achieving an accuracy of 80%. Its strength lies in its transparent, rule-based structure, which enables the model to capture non-linear relationships between shipment attributes such as cargo type, ports, and service cost.

With a precision of 0.81, the model reliably distinguished cost-effective shipping lines while maintaining a manageable false-positive rate. The recall score of 0.80 further indicates that the Decision Tree successfully identified most genuinely cost-efficient alternatives.

A key advantage of the Decision Tree model is interpretability. Unlike more complex models, their decision paths can be directly examined and translated into operational decision rules, making them particularly suitable for freight forwarders seeking both accuracy and managerial transparency.

4.5 Random Forests Performance

Random Forests outperformed all other models, achieving the highest accuracy of 83%. By aggregating predictions from multiple decision trees trained on randomized feature subsets, the model mitigates overfitting and improves generalization performance.

The model's precision (0.84) and recall (0.83) indicate superior stability and robustness compared to a single Decision Tree, particularly in handling noisy and heterogeneous freight forwarding data.

While Random Forests sacrifice some interpretability due to their ensemble structure, this limitation is offset by their enhanced predictive reliability. For operational decision-support systems prioritizing accuracy over explainability, Random Forests represent the most effective modeling choice in this study.

4.6 Neural Networks Performance

In our assessment of the most cost-effective shipping lines, the Neural Networks model demonstrated intermediate performance across all evaluated metrics. The accuracy of Neural Networks is 0.5, which indicates that the model correctly identified half of the cost-effective shipping lines. This is a significant drop compared to the leading models, such as Decision Trees and Random Forests.

The precision of Neural Networks is about 0.49, which means that there is a 49% chance that the model will correctly identify cost-effective shipping lines among its positive predictions. This, however, really means that there is a huge risk of false positives, and while attending to the predictions made by this model, it is still a very cautious and restrained approach.

As for the recall, Neural Networks scored 0.5, meaning that the model could classify correctly only fifty percent of actual cost-efficient shipping lines. This result indicates the need for improvement, especially if it is to reduce missed opportunities in logistic decision-making.

An F1-score of 0.48 suggests that Neural Networks hardly achieve an optimal balance between precision and recall, as the interaction between both is only moderate. It also means that while Neural Networks can handle some logistic predictions, they are not the most reliable choice for the presented dataset without further tuning.

In general, Neural Networks show a performance that is relatively moderate and less impressive compared to alternatives such as Decision Trees or Random Forests. While the adaptability and capacity of the model for improvement are considerable, it is currently more suitable for scenarios where high computational costs are tolerable and where a moderate level of accuracy is acceptable. To obtain better results for logistics decisions, Neural Networks might require more training, optimized parameters, and data scaling.

5. Discussion

A general look at the effectiveness of different predictive models towards cost-effective shipping line forecasts finds Random Forests and Decision Trees the most effective and beneficial predictive models. Such pronounced, clear superiority in their performance on a wide range of evaluation metrics outlines their status as reliable tools for freight forwarding companies' decision-makers. The great capability of these models to handle complex interactions with several variables, and their built-in robustness when dealing with highly diversified data sets, especially equips them for characteristics in freight operations that are always evolving (Gkerekos, Lazakis, and Theotokatos, 2019).

The superior performance of tree-based models can be attributed to their ability to natively handle categorical variables and complex feature interactions without requiring extensive preprocessing. Freight forwarding datasets typically include high-cardinality categorical attributes such as ports of loading, ports of discharge, cargo types, and shipping lines. Tree-based methods efficiently partition such feature spaces, whereas distance-based models like KNN and margin-based models such as SVM struggle with mixed data types and complex distance relationships.

5.1 Interpretation of Results

From the machine learning analysis done on the adoption of various models, such as Random Forests, Decision Trees, and K-Nearest Neighbor (KNN), on the historical shipment dataset, the shipping line choice was extrapolated. Key findings are given below:

- The best models for predicting cost-effective shipping lines were Random Forests and Decision Trees. These models had remarkably high accuracy and recall scores, with Random Forest achieving an accuracy of 0.83, while Decision Trees scored 0.8 for cost-effective shipping line predictions.

- K-Nearest Neighbor ranks third, with moderate performance and metrics around 0.63-0.65; this is a much bigger drop from the tree-based methods, but continues in terms of consistency of performance overall metrics.

- Naive Bayes comes fourth in the ranking, with a lower performance score but consistent on most measures (at around 0.55), indicative of slightly dependable predictive capabilities.

- Neural Networks come in fifth, with a performance that is rather modest, at about 0.50, which is surprising and disappointing given their usual prowess in many other areas.

- Interestingly, the performance of Support Vector Machines (SVM) in general average was underperformed, in predicting cost-effective shipping lines, suggesting that models like random forests, decision trees, and, to a lesser extent, the K-nearest neighbor model were best suited to dealing with the complexities of shipping data.

Machine Learning enables predictive analysis that helps a logistics manager anticipate cost efficiency across different shipping lines. These findings underscore the value of integrating data-driven decision-making into logistics management for optimizing both operational efficiency and competitiveness.

5.2 Practical Implications for Freight Forwarders

The findings from this study provide valuable insights for freight forwarders seeking to optimize their shipping line selection processes:

- Understanding how to use and integrate Random Forests and Decision Tree models into their decision-support systems can benefit freight forwarders. Doing so will help them predict with high accuracy the most cost-effective shipping lines and, in turn, will help them create additional decision rules that make digesting their complex data into complex yet simple enough insights that are actionable.

- Using machine learning models enables freight forwarders to make important decisions based on more than just tradition and experience. They allow for the analysis of vast quantities of data and can improve decision-making. They provide for the kind of operational excellence that can lead to cost savings and improved efficiency.

5.3 Comparison with Existing Literature

The current research investigates the application of machine learning in the decision-making process, particularly concerning selecting a shipping line that is cost-effective. It offers a summary of how freight forwarders can arrive at palatable choices, affirming some theories while introducing new concepts to the field. The work is based on prior studies regarding data analytics and decision-making within the realm of logistics.

The results from the machine learning analysis confirm that cost-effective shipping line predictions can be made quite accurately using Random Forest and Decision Tree. This is consistent with a handful of studies (Ermagun *et al.*, 2020; Zhou *et al.*, 2023) that have found Random Forest and Decision Tree to be a good bet for predictions involving logistics outcomes. K-Nearest Neighbor demonstrates moderate performance, typically scoring between 0.5 and 0.7, indicating their potential as a cost-effective option (Kotenko, 2022). Support Vector Machines (SVM) show the lowest performance for all parameters (Uddin, Anowar, and Eluru, 2021). Nonetheless, the study's findings do suggest that if one is going to predict something related to logistics, one should opt for Random Forest and Decision Tree rather than either of the two algorithms mentioned above.

6. Conclusion

The shipping line selection problem of freight forwarders has a solid solution pathway through the usage of machine learning models. This study's primary focus was on determining the models' capabilities of identifying a cost-efficient shipping line choice. From a pool of four model candidates, the work uncovered the Random Forest and Decision Tree as the best performers and most reliable solution predictors. With more than an 80% accuracy score, the models cleanly pass through a freight forwarder's dataset of 983 shipment records.

To enhance the model's generalizability, the researchers recommend expanding the dataset to encompass a broader array of geographical regions and time periods. They also suggest using different machine learning techniques for yield improvement. This indeed is an area wide open to undertaking vast research and investigation into a very diversified array of alternative machine learning methodologies/techniques that shall, through future work research, greatly improve operational efficiency for the freight forwarding industry and its intricate decision-making processes in search of predictions which are not only more efficient but also, across a variety of shipping lines and routes, more accurate.

Given the operational disruptions observed in the Egyptian market during 2022, future research should extend the modeling framework to include transit time as a secondary target variable, as cost alone does not fully capture freight forwarders' decision priorities under conditions of congestion, currency volatility, and service unreliability.

References

Alireza Ermagun, Aymeric Punel and Stathopoulos, A. (2020). Shipment status prediction in online crowd-sourced shipping platforms. *Sustainable Cities and Society*, 53. doi:https://doi.org/10.1016/j.scs.2019.101950.

Ayfer Ergin and Alkan, G. (2023). The selection of ocean container carrier: An analytic network process (ANP) approach. *Journal of ETA Maritime Science*, 11(3), pp.159–167. doi:https://doi.org/10.4274/jems.2023.96977.

Birkel, H., Matthias Kopyto and Lutz, C. (2020). Challenges of applying predictive analytics in transport logistics. ACM, pp.144–151. doi:https://doi.org/10.1145/3378539.3393864.

Calixto, N. and Ferreira, J. (2020). Salespeople performance evaluation with predictive analytics in B2B. *Applied Sciences*, 10(11), p.4036. doi:https://doi.org/10.3390/app10114036.

Cao, K. (2022). A machine learning-based approach to railway logistics transport path optimization. *Mathematical Problems in Engineering*, 2022, pp.1–11. doi:https://doi.org/10.1155/2022/1691215.

Christos Gkerekos, Iraklis Lazakis and Gerasimos Theotokatos (2019). Machine learning models for predicting ship main engine Fuel Oil Consumption: A comparative study. *Ocean Engineering*, 188, p.106282. doi:https://doi.org/10.1016/j.oceaneng.2019.106282.

D Rajasekar and Prabhakar, T.Sandeep. (2015). Importance of freight forwarding and work activities of freight forwarders-an empirical analysis. *250 X Indian Journal of Applied Research*, 5(1).

Ding, J.-F., Dye, C.-Y., Liang, S.-Y., Chou, C.-C., Tseng, W.-J. and Shyu, W.-H. (2017). Evaluating key factors influencing the development of multi-country consolidation for ocean freight forwarders in Taiwan. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 231(1), pp.342–352. doi:https://doi.org/10.1177/1475090215621577.

Egemen Sert, Hedayatifar, L., Rigg, R.A., Akhavan, A., Buchel, O., Saadi, D.E., Kar, A.A., Morales, A.J. and Yaneer Bar-Yam (2020). Freight time and cost optimization in complex logistics networks. *Complexity*, 2020, pp.1–11. doi:https://doi.org/10.1155/2020/2189275.

Fanam, P.D., Nguyen, H.O. and Cahoon, S. (2018). An empirical analysis of the critical selection criteria of liner operators: The perspective of freight forwarders. *International Journal of Shipping and Transport Logistics*, 10(5-6). doi:https://doi.org/10.1504/ijstl.2018.095222.

FIATA (2019). *FIATA model rules for freight forwardering services. Glattbrugg*.

Hicks, S.A., Strümke, I., Vajira Thambawita, Malek Hammou, Riegler, M.A., Halvorsen, P. and Sravanthi Parasa (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, 12(1), p.5979. doi:https://doi.org/10.1038/s41598-022-09954-8.

Ho, T.C., Chiu, R.H., Chung, C.C. and Lee, H.S (2017). Key influence factors for ocean freight forwarders selecting container shipping lines using the revised dematel approach. *Journal of Marine Science and Technology (Taiwan)*, 25(3). doi:https://doi.org/10.6119/JMST-016-1227-1.

Huang, S.T., Emrah Bulut and Duru, O. (2019). Service quality evaluation of international freight forwarders: an empirical research in East Asia. *Journal of Shipping and Trade*, 4(1), p.14. doi:https://doi.org/10.1186/s41072-019-0053-6.

Jang, H.-S., Chang, T.-W. and Kim, S.-H. (2023).

Prediction of shipping cost on freight brokerage platform using machine learning. *Sustainability*, 15(2), p.1122. doi:https://doi.org/10.3390/su15021122.

Jussi Tohka and Gils, M. van (2021). Evaluation of machine learning algorithms for health and wellness applications: A tutorial. *Computers in Biology and Medicine*, 132, p.104324. doi:https://doi.org/10.1016/j.compbiomed.2021.104324.

Kotenko, V. (2022). Application of algorithmic models of machine learning to the freight transportation process. *Transport technologies*, 2022(2), pp.10–21. doi:https://doi.org/10.23939/tt2022.02.010.

Lukinskiy, V. and Vladislav Lukinskiy (2015). Analysis of the logistics intermediaries choice methods in the supply chains. *Transport and Telecommunication Journal*, 16(4), pp.288–295. doi:https://doi.org/10.1515/ttj-2015-0026.

Raj Bridgelall (2024). Locating electrified aircraft service to reduce urban congestion. *Information*, 15(4), p.186. doi:https://doi.org/10.3390/info15040186.

Rao, D., Chandra, D. and Kumar, D. (2017). *A survey on machine learning: Concept, algorithms and applications'. International Conference on Innovative Research in Computer and Communication Engineering*.

Ron, Feelders, A.J. and Hennie (2015). Uncovering document fraud in maritime freight transport based on probabilistic classification. In: K. Saeed and Wladyslaw Homenda, eds. Springer International Publishing. doi:https://doi.org/10.1007/978-3-319-24369-6.

Saraswathi Hathikal, Chung, S.H. and Karczewski, M. (2020). Prediction of ocean import shipment lead time using machine learning methods. *SN Applied Sciences*, 2(7), p.1272. doi:https://doi.org/10.1007/s42452-020-2951-5.

Shang, Y., Dunson, D. and Song, J.-S. (2017). Exploiting big data in logistics risk assessment via bayesian nonparametrics. *Operations Research*, 65(6), pp.1574–1588. doi:https://doi.org/10.1287/opre.2017.1612.

Szymon Cyperski, Michał Okulewicz and Domański, P.D. (2024). On cost estimation of the full truckload contracts. *Journal of Advances in Information Technology*, 15(5), pp.614–621. doi:https://doi.org/10.12720/jait.15.5.614-621.

Tri Wahyudi and Dava Septya Arroufu (2022). IMPLEMENTATION OF DATA MINING PREDICTION DELIVERY TIME USING LINEAR REGRESSION ALGORITHM. *Journal of Applied Engineering and Technological Science*, 4(1). doi:https://doi.org/10.37385/jaets.v4i1.918.

Uddin, M., Sabreena Anowar and Naveen Eluru (2021). Modeling freight mode choice using machine learning classifiers: a comparative study using Commodity Flow Survey (CFS) data. *Transportation Planning and Technology*, 44(5), pp.543–559. doi:https://doi.org/10.1080/03081060.2021.1927306.

Zhou, Y., Chen, W., Sun, X. and Yang, D. (2023). Early warning of telecom enterprise customer churn based on ensemble learning. *PLOS ONE*, 18(10), p.e0292466. doi:https://doi.org/10.1371/journal.pone.0292466.