PREDICTING AND ANALYZING SHIPPING TIME USING ENSEMBLE TREE MODELS WITH SHAPLEY ADDITIVE EXPLANATIONS

Panupong Wanjantuk[1*] and Ruth Banomyong[2]
[1] Department of Computer Engineering, Khon Kaen University, Thailand
[2]Department of International Business, Logistics and Transport, Thammasat Bussiness School, Thammasat University, Thailand

Abstract

Purpose: This study aims to develop a predictive model for shipping times using advanced machine learning techniques, specifically ensemble learning models, to enhance accuracy and reliability in supply chain management. Additionally, the study seeks to interpret the model's predictions to understand the influence of various factors on shipping durations, providing actionable insights for optimizing logistics operations.
Design/methodology/approach: The research utilizes a dataset comprising 5,114 rows of historical shipment records. Data preprocessing included one-hot encoding for categorical variables, scaling numerical features, and addressing imbalanced distributions. Seven ensemble tree models were employed. The models were evaluated using cross-validation and various evaluation metrics. SHAP (SHapley Additive exPlanations) was used to interpret the best-performing model, providing insights into feature importance and interactions.
Findings: The CatBoost model demonstrated the highest accuracy in predicting shipping times, followed by Random Forest and LightGBM. SHAP analysis revealed that normalized shipment charges, processing days, and turnaround time thresholds were the most significant features influencing shipping times. Interaction plots highlighted the complex dependencies between features.
Research Limitation: The study's limitations include the quality and completeness of the dataset, which can affect model performance. Additionally, the models' predictive power may diminish when applied to significantly different future conditions or new shipping routes not represented in the training data.
Practical Implication: The predictive models developed in this study can be integrated into logistics management systems to provide real-time shipping time estimates and insights, enhancing inventory management, reducing costs, and improving customer satisfaction. The interpretability provided by SHAP helps logistics managers optimize processes, resource allocation, and scheduling.
Originality/value: This research demonstrates the effectiveness of combining ensemble learning models with SHAP for accurate and interpretable shipping time predictions. The study provides a comprehensive approach to understanding the factors influencing shipping durations, offering valuable insights for improving supply chain efficiency.

Keywords: Shipping Time Prediction, Ensemble Learning Models, SHAP, Supply Chain Management, Predictive Analytics

Introduction

Shipping times are critical in supply chain management, significantly impacting logistics efficiency and reliability. Timely delivery is essential for maintaining supply chain continuity and meeting customer expectations [1]. Delays can disrupt inventory levels, production schedules, and customer satisfaction, especially in sectors like retail, manufacturing, and healthcare. Modern supply chains are complex, driven by globalization and the need for just-in-time delivery, necessitating accurate shipping time predictions [2]. Accurate predictions help businesses optimize inventory management, reducing costs associated with overstocking or stockouts [3]. They also enhance customer satisfaction by providing reliable delivery estimates [4]. Additionally, accurate predictions improve decision-making within the supply chain by anticipating delays and allowing for operational adjustments, thus enhancing overall resilience [5]. Advanced predictive models using machine learning and deep learning techniques offer the potential for greater precision in shipping time forecasts by leveraging large datasets and sophisticated algorithms. Integrating these techniques is essential for maintaining competitiveness and efficiency in the modern business landscape.

The primary objective of this research is to develop a predictive model for shipping times using historical shipment data. By leveraging advanced machine learning techniques, we aim to create a model that can accurately forecast the time required for shipments to reach their destinations, enabling logistics companies to better manage their operations and provide more reliable delivery estimates to their customers. Additionally, a secondary objective is to interpret the model's predictions to understand the influence of various factors on shipping times. Employing interpretability techniques, we seek to uncover key determinants impacting shipping durations, which will validate the model and offer actionable insights for optimizing shipping processes and mitigating delays.

The rest of the paper is organized as follows: Section 2 reviews existing methods and models for predicting shipping times, focusing on statistical and machine learning techniques and the importance of model interpretability. Section 3 outlines the methodology, including data collection, preprocessing, model selection, training, and evaluation. Section 4 presents the results, highlighting performance metrics and interpreting the best-performing model with SHAP Tree Explainer. Section 5 discusses the implications, limitations, and future research directions. Finally, Section 6 concludes key findings and potential applications in supply chain management.

Literature Review

The research on predicting shipping times has seen significant advancements through the application of both traditional methods and machine learning techniques. Traditional methods often relied on tank model tests and semi-empirical formulas, as noted in the comparative study by [6], which highlighted the shift towards more sophisticate machine learning (ML) algorithms to improve operational efficiency and meet regulatory requirements. Machine learning has been particularly effective in enhancing prediction accuracy and efficiency. For instance, [7] developed a predictive model using multinomial logistic regression and other classifiers to estimate ocean import shipment lead times with high accuracy. Similarly, [8] used multiple regression techniques combined with marine traffic data to predict shipment ETAs with an 89% accuracy rate. [9] demonstrated the benefits of ML in the oil and gas industry by improving job scheduling and reducing costs through precise lead time predictions. [10] successfully applied a Random Forest algorithm to predict shipping times between Southeast Asia and North America, achieving superior results compared to traditional methods. Overall, the integration of ML in shipping logistics has proven to significantly enhance visibility, predictability, and efficiency across various applications. Additionally, [11] investigates using machine learning to predict delivery times early in the order process for small batch production companies. Using data from two German manufacturers, it develops a machine learning approach to predict delivery dates upon receiving a request for an offer, incorporating the desired customer delivery date. Results demonstrate that machine learning significantly improves delivery date prediction accuracy, reduces manual effort, and enhances competitive advantage.

Model interpretability in machine learning is critical for ensuring transparency, trust, and actionable insights, especially in high-stakes domains like healthcare, finance, and logistics. Interpretability allows stakeholders to understand the reasoning behind a model's predictions, making it possible to identify and correct biases, validate model outputs, and comply with regulatory requirements. Previous work on interpretability in machine learning models for predicting shipping times and related logistics processes is relatively scant but emerging. Studies have shown the importance of using interpretability techniques to understand the influence of various factors on model predictions, offering actionable insights for optimizing shipping processes and mitigating delays. For example, [12] explored interpretability in predictive process analytics, using machine learning models trained on historical process data to make predictions about business processes, such as shipment completions. Their study emphasized the need for interpretability to ensure the reliability of high-accuracy models by providing explanations for their predictions. Similarly, [13] conducted a comparative study on interpretability techniques for models trained on time series data, including shipping logs, to predict outages and delays. They highlighted the advantages and limitations of various interpretability methods such as LIME and SHAP in providing insights into model decisions. Additionally, [14] investigated the predictability of asset prices in shipping using a novel hybrid algorithm, underscoring the role of interpretability in improving forecast accuracy and understanding model behavior.

Methodology

To begin with, the original shipment dataset is preprocessed by eliminating redundant and erroneous information. The ensemble learning models, particularly those based on boosting, are initially developed based on data partitioning into training and testing datasets. The training dataset is used to build the predictive model, and the test dataset is used to evaluate the models' performance. Once the optimal model with the best performance is identified, the SHAP approach is utilized to establish additive attributes that are then employed to determine the importance of variables for shipping times and the contributions of various factors to each predicted shipping duration.

Data Collection

The dataset used in our experiments was sourced from Kaggle, comprising 5,114 rows of historical shipment records. Each record includes various features as detailed in Table 1.

Type of Feature	Feature name	Description
Location	pick_up_point	Pick Up Point
	drop_off_point	Drop Off Point
	source_country	Country from where the goods need to be shipped
	destination_country	Country to where the goods need to be shipped
Cost and Charges	freight_cost	Cost of transportation per kilogram
	shipment_charges	Fixed cost per shipment
	total_freight_cost	Total cost of freight (freight_cost * gross_weight)
	normalized_shipment_charges	Normalized shipment charges to account for variability
Weight	gross_weight	Gross weight in kilograms of the shipment
	freight_cost_gross_weight	Interaction term combining the freight cost and gross weight
Shipment Mode	shipment_mode	Method of shipment (e.g., air, ocean)
Shipping Company	shipping_company	Candidate shipping company
	selected	Whether the company in 'shipping_company' was selected or not
Timing	day_of_week	The day the shipment was sent
	month	The month the shipment was sent
	cut_off_time_binary	Binary indicator of whether the shipment was sent within the cut-off time
	tat_binary	Binary indicator if the shipment turnaround time is within a certain threshold
	processing_days_binary	Binary indicator of whether the shipment was processed on a working day
	is_weekend	Binary indicator if the shipment was sent on a weekend
Target	shipping_time	The amount of time that it takes for goods to reach their destination (in days)

Table 1: Data Description

Data Preprocessing

Data processing included one-hot encoding for categorical variables, scaling numerical features, and handling outliers to improve data quality. The shipping time target variable showed a highly imbalanced distribution, addressed using the SMOGN (Synthetic Minority Over-sampling Technique for Regression with Gaussian Noise) technique [15].

Model Selection

In this study, we selected seven ensemble tree models for regression due to their effectiveness and versatility in handling various regression tasks. These models—Random Forest, Extra Trees, Gradient Boosting Machine (GBM), XGBoost, LightGBM, CatBoost, and NGBoost—were chosen for their unique strengths, such as robustness to overfitting, handling of categorical data, and providing uncertainty estimates. Random Forest combines multiple decision trees, improving accuracy and reducing overfitting through bootstrapping and feature randomness [16]. Extra Trees incorporate additional randomness in tree construction, which helps reduce variance and is suitable for complex data [17]. GBM builds sequential trees, correcting errors with each iteration using gradient descent [18]. XGBoost enhances gradient boosting with regularization, efficient handling of sparse data, and improved scalability [19]. LightGBM uses a histogram-based algorithm for faster training and better accuracy, especially with large datasets [20]. CatBoost efficiently handles categorical features and reduces the need for extensive preprocessing [21]. Finally, NGBoost employs a probabilistic framework, providing uncertainty estimates useful for decision-making processes [22]. By leveraging the complementary strengths of these models, we aim to achieve highly accurate and reliable predictions for shipping times in supply chain management.

SHAP (SHapley Additive exPlanations)

1.1 SHAP is a powerful interpretability technique in machine learning that provides consistent and locally accurate explanations for model predictions. By assigning each feature an importance value for a particular prediction, SHAP helps to understand the contribution of each feature to the final output. This method is based on cooperative game theory and provides both global and local interpretability. SHAP values are particularly useful for identifying the influence of features in complex models, making them essential for transparent and explainable AI in fields such as logistics, healthcare, and finance [23].

Model Training and Evaluation

To ensure the robustness and reliability of our predictive model, we employed a 5-fold cross-validation approach during the training phase. This method involves partitioning the dataset into five equally sized folds. In each iteration, one fold is used as the validation set while the remaining four folds are combined to form the training set. This process is repeated five times, with each fold serving as the validation set exactly once. The model's performance metrics are then averaged across all five iterations, providing a more comprehensive evaluation of its generalizability and stability. The metrics used include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2).

<u>Results</u>

Model Performance

The performance metrics for the seven ensemble tree models in Table 2 indicate that CatBoost achieves the highest accuracy in predicting shipping times, with an MAE of 3.86 and an R-squared of 0.7659. Random Forest and LightGBM also perform well, with MAE values of 3.61 and 3.83 and R-squared values of 0.7616 and 0.7573, respectively. Extra Trees shows strong performance with the lowest MAE of 3.42 but slightly lower R-squared of 0.7589. Gradient Boosting, XGBoost, and NGBoost exhibit higher MAE and lower R-squared values compared to the top performers, indicating moderate prediction accuracy. Overall, CatBoost stands out as the most effective model for this application, closely followed by Random Forest and LightGBM.

Model	MAE	MSE	RMSE	MAPE	R-squared
CatBoost	3.860030	35.166733	5.923949	25.570718	0.765891
Random Forest	3.613944	35.767870	5.975536	21.440336	0.761555

LightGBM	3.836864	36.428714	6.030282	24.273327	0.757307
Extra Trees	3.417776	36.177137	6.008218	20.757644	0.758860
Gradient Boosting	4.093316	38.590532	6.207841	24.930634	0.742974
XGBoost	4.015028	39.003233	6.239473	27.141350	0.740433
NGBoost	4.115320	39.136199	6.252095	24.784378	0.739384

Table 2: Model Performance

Interpretation of Results

The analysis utilizes SHAP (SHapley Additive exPlanations) to interpret and visualize the impact of features on the predictions made by a CatBoost model. The SHAP TreeExplainer is employed with the trained model to compute SHAP values for the training dataset. Two types of SHAP summary plots are generated: a bar plot (shown in Figure 1) that highlights the overall importance of each feature, and a detailed feature importance plot that illustrates the distribution of SHAP values for each feature across all samples. These plots help in understanding which features have the most significant impact on the model’s predictions.

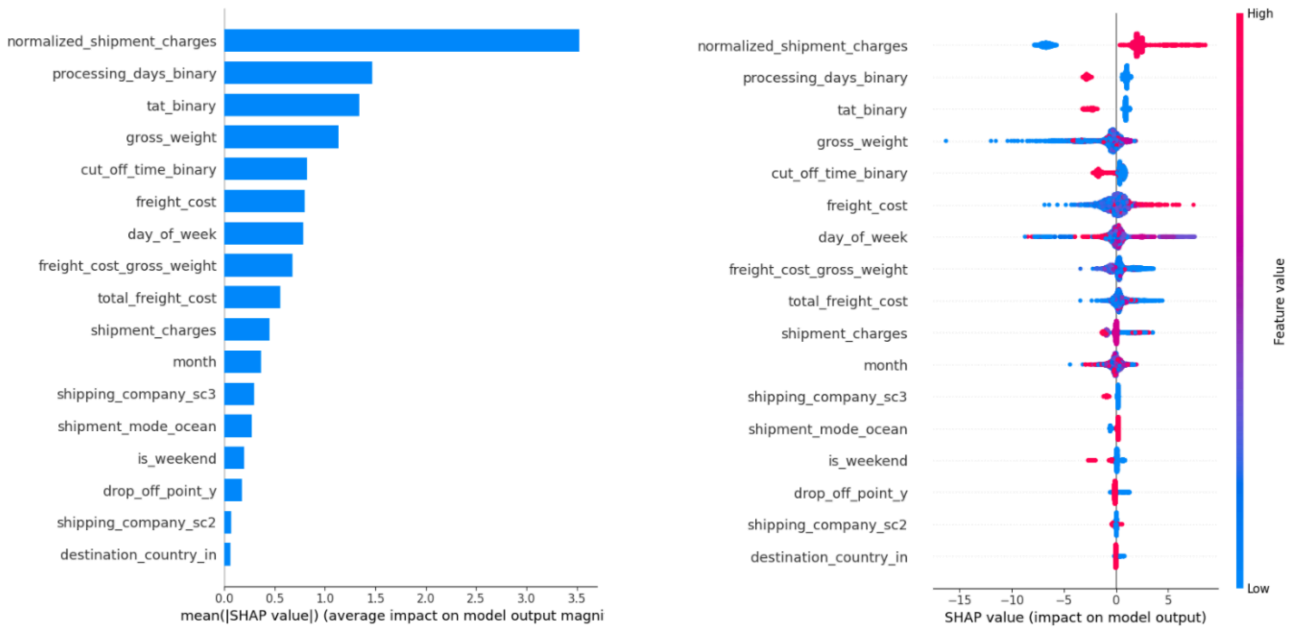
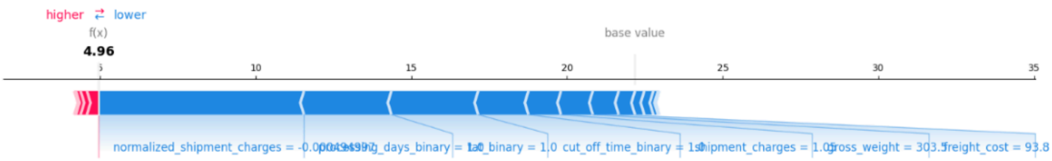


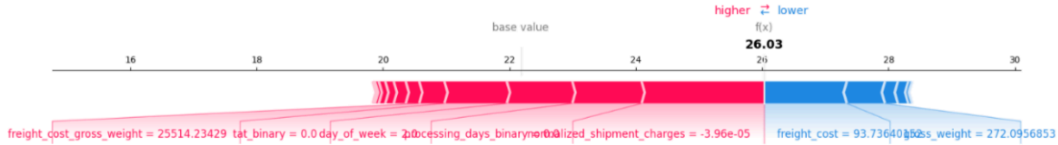
Figure 1: SHAP summary plots

The SHAP summary plot shows the impact of various features on the model's output for shipping time predictions. The most influential feature is 'normalized_shipment_charges' (Normalized shipment charges), indicating a significant effect on shipping times. 'processing_days_binary' (Binary indicator of whether the shipment was processed on a working day) and 'tat_binary' (Binary indicator if the shipment turnaround time is within a certain threshold) are also crucial, highlighting the importance of operational timing in meeting shipping deadlines. 'gross_weight' (Gross weight in kilograms of the shipment) and 'freight_cost' (Cost of transportation per kilogram) further emphasize the role of shipment weight and transportation cost in predicting shipping durations. Other important features include 'cut_off_time_binary' (Binary indicator of whether the shipment was sent within the cut-off time), 'day_of_week' (The day the shipment was sent), and 'total_freight_cost' (Total cost of freight). Less impactful features, such as specific shipping companies ('shipping_company_sc3', 'shipping_company_sc2'), shipment mode ('shipment_mode_ocean'), and destination country ('destination_country_in'), still contribute to the model but to a lesser extent. This analysis underscores the significance of cost-related and operational timing factors in determining shipping times.

The SHAP force plot in Figure 2 (a) details how each feature influences a specific shipping time prediction of 4.96 days, with the base value representing the average model prediction. Key operational efficiency factors—such as normalized shipment charges, processing on working days, meeting TAT thresholds, and sending shipments within cut-off times—significantly reduce shipping times. Conversely, the gross weight of the shipment increases the predicted time. These insights can help optimize logistics operations for more accurate and efficient shipping time predictions.



(a) plot for an instance value less than the base value

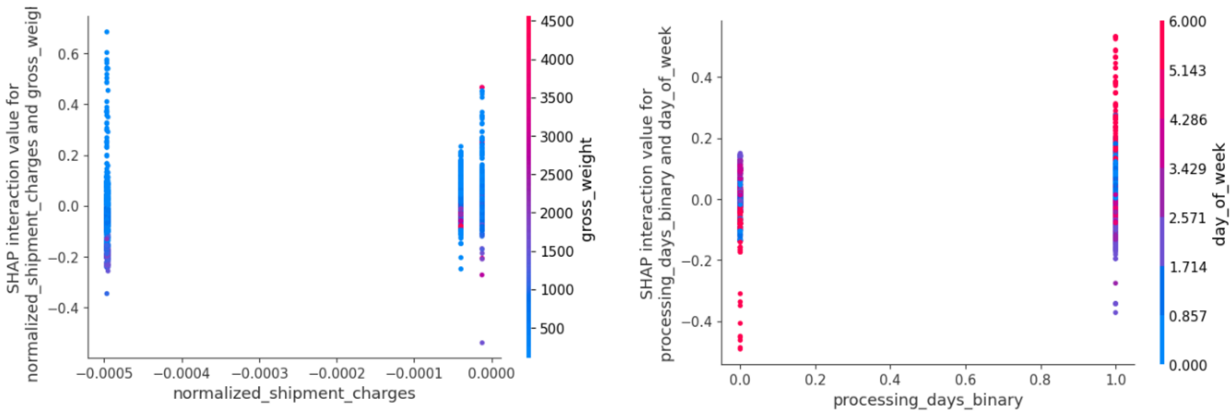


(b) plot for an instance value greater than the base value.

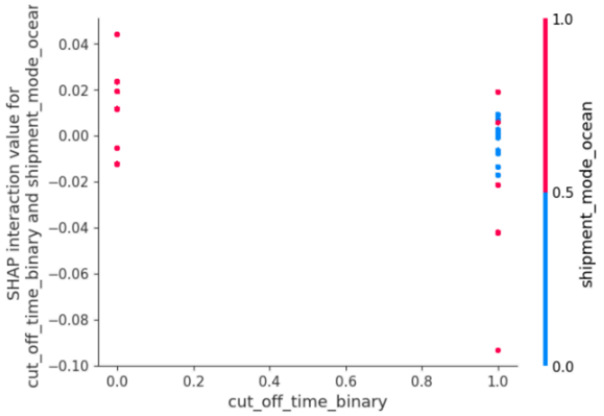
Figure 2: SHAP explanatory force plots

The SHAP force plot in Figure 2 (a) shows the contribution of each feature to a specific shipping time prediction of 4.96 days, with the base value representing the average model prediction. Key factors like normalized shipment charges, processing on working days, meeting TAT thresholds, and sending shipments within cut-off times significantly reduce shipping times, while the gross weight of the shipment increases it. These insights can help optimize logistics operations for more accurate and efficient shipping time predictions.

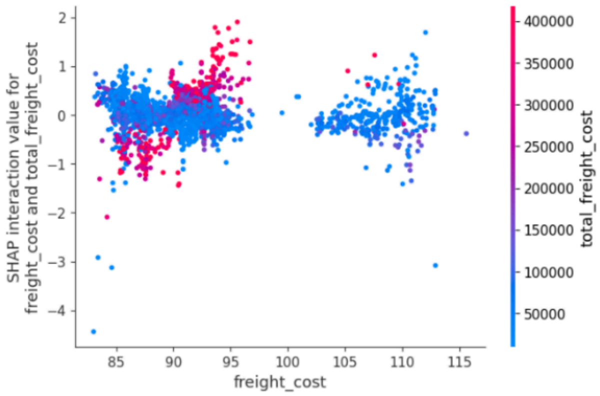
The SHAP interaction plot provides a detailed analysis of how pairs of features jointly influence the predictions made by the CatBoost model. The SHAP TreeExplainer computes SHAP interaction values for the training dataset, visualizing the combined effects of two features on the model's output. The interaction plot (shown in Figure 3) highlights how the interplay between specific features.



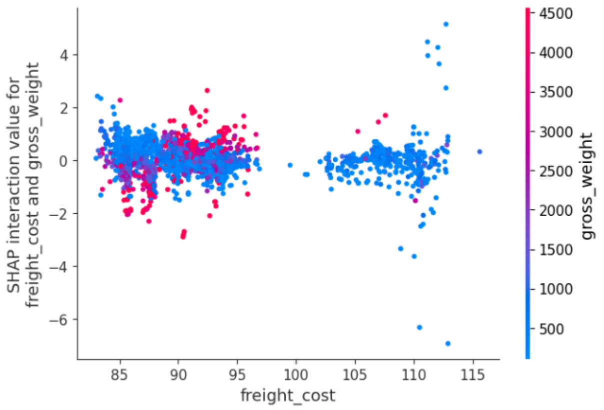
(a) (b)



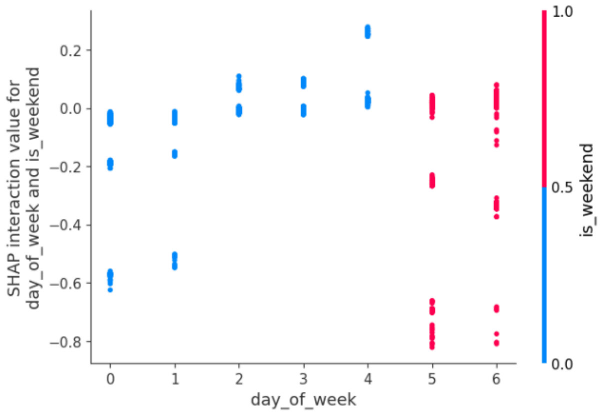
(c)



(d)



(e)



(f)

Figure 3: The SHAP interaction plots

The SHAP interaction plots reveal critical insights into how various factors jointly influence shipping time predictions. Higher normalized shipment charges, particularly for heavier shipments, lead to more consistent SHAP interaction values, highlighting the importance of cost efficiency and weight (Figure 3 (a)). The interaction between processing days and day of the week (Figure 3 (b)) shows that non-working days, especially near weekends, cause greater variability and longer shipping times, emphasizing the need to optimize non-working day processes. Cut-off times are crucial for consistent shipping times, with ocean shipments inherently slower, stressing the importance of careful planning (Figure 3 (c)). The relationship between freight cost and total freight cost (Figure 3 (d)) indicates that lower per kilogram costs can lead to higher total costs and varied shipping times, while higher per kilogram costs result in more consistent shipping times. Similarly, the interplay between freight cost and gross weight (Figure 3 (e)) shows that lower freight costs increase variability, especially for heavier shipments, whereas higher costs lead to more predictable times. Finally, the difference between weekday and weekend operations (Figure 3 (f)) underscores the need for efficient weekend processes to reduce delays and improve performance.

Discussion

Implications of Findings

The findings of this study have significant implications for improving logistics operations. By leveraging ensemble tree models combined with SHAP (SHapley Additive exPlanations), logistics companies can achieve more accurate and interpretable shipping time predictions. The ability to predict shipping times with high precision allows businesses to optimize their inventory management, reducing costs associated with overstocking or stockouts. Furthermore, reliable delivery estimates enhance customer satisfaction and trust, which are critical in sectors such as retail, manufacturing, and healthcare.

The interpretability provided by SHAP helps logistics managers understand the impact of various features on shipping times, enabling targeted interventions to address bottlenecks and inefficiencies.

For instance, insights into the effects of normalized shipment charges, processing days, and turnaround time thresholds can inform decisions about resource allocation, scheduling, and pricing strategies, ultimately leading to more resilient and efficient supply chains.

Limitations

While the study demonstrates the effectiveness of ensemble tree models and SHAP in predicting and analyzing shipping times, there are several limitations to consider. First, the quality and completeness of the dataset can significantly influence model performance. The dataset used in this study, although comprehensive, may still contain inaccuracies or missing values that could affect the results. Additionally, the models were trained on historical shipment data, and their predictive power may diminish when applied to significantly different future conditions or new shipping routes not represented in the training data. Another limitation is the computational complexity of ensemble models, which can be resource-intensive to train and deploy. This might limit their practical application in scenarios where computational resources are constrained. Furthermore, while SHAP provides valuable interpretability, it does not account for potential interactions between unobserved variables, which could lead to oversights in understanding the full context of shipping time predictions.

Future Work:

Future research should focus on addressing the limitations identified in this study to further enhance the model and its applications. One area for further investigation is the incorporation of additional data sources, such as real-time traffic information, weather conditions, and port congestion data, to improve the accuracy and robustness of shipping time predictions. Integrating these dynamic factors can help create a more comprehensive model that better reflects the complexities of real-world logistics operations. Additionally, exploring the application of deep learning techniques, such as recurrent neural networks (RNNs) or transformers, could potentially enhance the predictive performance, especially in capturing temporal dependencies in shipment data. Another promising direction is the development of hybrid models that combine the strengths of ensemble tree models and other machine learning approaches to achieve better predictive accuracy and interpretability. Finally, conducting case studies or pilot implementations with industry partners could provide valuable insights into the practical challenges and benefits of deploying these predictive models in real-world logistics environments, facilitating further refinement and adoption of the technology.

Conclusion

This study aimed to develop and interpret predictive models for shipping times using ensemble tree methods enhanced with SHAP (SHapley Additive exPlanations). The analysis demonstrated that ensemble tree models, particularly CatBoost, effectively predict shipping times with high accuracy. The models' interpretability, provided by SHAP, revealed critical insights into the impact of various features on shipping durations. Key determinants such as normalized shipment charges, processing days, and turnaround time thresholds were identified as significant influencers of shipping times. These findings underscore the robustness and practicality of using advanced machine learning techniques in logistics.

The predictive models developed in this study have substantial practical applications in the field of logistics. By implementing these models, logistics companies can enhance their shipping time estimates, leading to improved inventory management, cost reductions, and heightened customer satisfaction through reliable delivery promises. The interpretability provided by SHAP allows managers to pinpoint specific areas for operational improvements, such as optimizing processing schedules and adjusting pricing strategies based on shipment weights and costs. Furthermore, the models can be integrated into existing logistics management systems to provide real-time predictions and insights, enabling proactive decision-making and enhanced overall supply chain resilience.

The broader implications of this research highlight the transformative potential of integrating advanced machine learning models with interpretability techniques in logistics. As global supply chains become increasingly complex, the ability to predict and analyze shipping times accurately is more critical than ever. This study not only demonstrates the efficacy of ensemble learning models but also emphasizes the importance of understanding the underlying factors influencing predictions. Future research and practical implementations should continue to build on these findings, exploring new data sources and hybrid models to further enhance prediction accuracy and operational efficiency. Ultimately, the adoption of these advanced predictive tools can lead to more resilient, efficient, and customer-centric logistics operations, driving competitive advantage in the industry.

References

- [1] M. Christopher, *Logistics and Supply Chain Management: Logistics & Supply Chain Management*. Pearson UK, 2016.
- [2] J. T. Mentzer *et al.*, “Defining supply chain management,” *Journal of Business logistics*, vol. 22, no. 2, pp. 1–25, 2001.
- [3] S. Chopra and P. Meindl, *Supply chain management: strategy, planning, and operation*, Fifth edition, Global edition. in Pearson international edition. Boston: Pearson, 2013.
- [4] A. Harrison and R. I. van Hoek, *Logistics management and strategy: competing through the supply chain*, 3rd ed. Harlow, England ; New York: Prentice Hall Financial Times, 2008.
- [5] D. Ivanov and A. Dolgui, “OR-methods for coping with the ripple effect in supply chains during COVID-19 pandemic: Managerial insights and research implications,” *International journal of production economics*, vol. 232, p. 107921, 2021.
- [6] K. Alexiou, E. G. Pariotis, H. C. Leligou, and T. C. Zannis, “Towards Data-Driven Models in the Prediction of Ship Performance (Speed—Power) in Actual Seas: A Comparative Study between Modern Approaches,” *Energies*, vol. 15, no. 16, p. 6094, Aug. 2022, doi: 10.3390/en15166094.
- [7] S. Hathikal, S. H. Chung, and M. Karczewski, “Prediction of ocean import shipment lead time using machine learning methods,” *SN Appl. Sci.*, vol. 2, no. 7, p. 1272, Jul. 2020, doi: 10.1007/s42452-020-2951-5.
- [8] P. C. Mahajan, A. W. Kiwelekar, L. D. Netak, and A. B. Ghodake, “Predicting Expected Time of Arrival of Shipments Through Multiple Linear Regression,” in *ICDSMLA 2020*, vol. 783, A. Kumar, S. Senatore, and V. K. Gunjan, Eds., in Lecture Notes in Electrical Engineering, vol. 783. , Singapore: Springer Singapore, 2022, pp. 343–350. doi: 10.1007/978-981-16-3690-5_30.
- [9] F. Sethi, “Using Machine Learning Methods to Predict Order Lead Times,” *International Journal of Sciences: Basic and Applied Research*, vol. 54, no. 3, pp. 87–96, 2020.
- [10] A. Jonquais and F. Krempl, “Predicting shipping time with machine learning,” 2019.
- [11] A. Rokoss, M. Syberg, L. Tomidei, C. Hülsing, J. Deuse, and M. Schmidt, “Case study on delivery time determination using a machine learning approach in small batch production companies,” *J Intell Manuf*, Jan. 2024, doi: 10.1007/s10845-023-02290-2.
- [12] R. Sindhgatta, C. Ouyang, and C. Moreira, “Exploring Interpretability for Predictive Process Analytics,” in *Service-Oriented Computing*, vol. 12571, E. Kafeza, B. Benatallah, F. Martinelli, H. Hacid, A. Bouguettaya, and H. Motahari, Eds., in Lecture Notes in Computer Science, vol. 12571. , Cham: Springer International Publishing, 2020, pp. 439–447. doi: 10.1007/978-3-030-65310-1_31.
- [13] A. Jalali, A. Schindler, B. Haslhofer, and A. Rauber, “Machine Learning Interpretability Techniques for Outage Prediction: A Comparative Study,” *PHME_CONF*, vol. 5, no. 1, p. 10, Jul. 2020, doi: 10.36001/phme.2020.v5i1.1244.
- [14] O. Duru, E. Gulay, and S. C. Girgin, “Predictability of second-hand bulk carriers with a novel hybrid algorithm,” *The Asian Journal of Shipping and Logistics*, vol. 37, no. 4, pp. 291–300, Dec. 2021, doi: 10.1016/j.ajsl.2021.07.002.
- [15] N. Kunz, “SMOGN: Synthetic Minority Over-Sampling Technique for Regression with Gaussian Noise.” PyPI, 2020. [Online]. Available: https://pypi.org/project/smogn/
- [16] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [17] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely randomized trees,” *Machine learning*, vol. 63, pp. 3–42, 2006.
- [18] J. H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Annals of statistics*, pp. 1189–1232, 2001.
- [19] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [20] G. Ke *et al.*, “Lightgbm: A highly efficient gradient boosting decision tree,” *Advances in neural information processing systems*, vol. 30, 2017.
- [21] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” *Advances in neural information processing systems*, vol. 31, 2018.
- [22] T. Duan *et al.*, “Ngboost: Natural gradient boosting for probabilistic prediction,” in *International conference on machine learning*, PMLR, 2020, pp. 2690–2700.
- [23] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, in NIPS’17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 4768–4777.