Machine Learning-Based Classification-Regression Model for Home Appliance Logistics Delivery Time Prediction

Yixuan Sun, Xin Wang*, Yi Li

Chongqing University of Posts and Telecommunications, Chongqing 400065, China

*s242431019@stu.cqupt.edu.cn

Abstract. Accurate prediction of home appliance delivery time is crucial for enhancing customer satisfaction, yet existing research fails to account for appliance characteristics and ignores variations in delivery time windows. Based on 3.491 million home appliance delivery samples from Ririshun Logistics, this study employs cumulative interval experiments to identify ≤96 hours as the core prediction interval. A two-stage classification-regression model is designed (first assigning five labels, then customizing a regression model). Experiments demonstrate that compared to a single regression model, this approach reduces Root Mean Square Error (RMSE) by 6.4% and Mean Absolute Error (MAE) by 25.2%, filling a gap in home appliance logistics delivery time prediction and supporting scenario-specific operations.

Keywords: Home Appliance Logistics, Classification-Regression, Delivery Time Prediction, Machine Learning.

1 Introduction

Delivery time prediction is a critical challenge for intelligent logistics systems,as accurate forecasting positively impacts both businesses and customers [1]. Against the backdrop of e-commerce expansion, delivery time directly correlates with customer satisfaction and retailer performance outcomes. Providing precise delivery times effectively manages customer expectations [2], while delayed deliveries trigger consumer dissatisfaction and may reduce repurchase rates [3]. Both overly conservative and insufficient delivery time commitments cause issues, making accurate prediction central to balancing these conflicts [4].

Extensive research has explored logistics delivery time prediction for general merchandise, covering factors such as temporal periodicity, transport distance, and order complexity [1,2,4]. However, home appliance logistics involve large, heavy products requiring specialized equipment and installation processes [5], resulting in significantly different delivery time-influencing factors compared to general goods. Meanwhile, although Zhang and Smutkupt [6] validated the critical role of home appliance delivery

Y. Xu et al. (eds.), *Proceedings of the 2026 5th International Conference on Engineering Management and Information Science (EMIS 2026)*, Advances in Computer Science Research 130,
https://doi.org/10.2991/978-94-6239-652-4_8

time in customer satisfaction, no subsequent studies have delved into the specific factors affecting delivery time or developed targeted prediction models for this scenario.

Building on these observations, three shortcomings of current research need to be addressed: First, studies on general logistics [1-2] ignore appliance-specific characteristics, leading to inadequate factor coverage for home appliance delivery. Second, prior prediction models [7,9] treat delivery time as a homogeneous whole, overlooking variability across different time intervals and amplifying prediction errors. Third, despite the proven importance of home appliance delivery time [6], there is a lack of scenario-specific prediction frameworks, with most existing methods fail to adapt to the unique operational needs of home appliance logistics.

This study utilizes the specialized home appliance delivery dataset from Ririshun Logistics [5], which covers the entire Chinese home appliance supply chain. Delivery time (delivery_hour) is defined as the difference between the customer's actual receipt time (QS node) and the time the order was dispatched to the fulfillment center (XF node), providing a clear target variable for model construction.

This study contributes to the field by focusing on the home appliance-specific logistics scenario to analyze the unique factors influencing delivery time, filling the gap in scenario-specific prediction for home appliance; identifying the optimal core prediction interval through cumulative interval experiments to overcome the neglect of interval variability in existing research; and developing a two-stage classification-regression model that classifies orders into sub-intervals before training customized regression models.

2 Literature Review

To systematically contextualize this study, the literature review is structured around three core themes: factors influencing logistics delivery time, home appliance-specific logistics research, and delivery time prediction methods.

Logistics delivery time in general scenarios is influenced by multidimensional factors: temporally, periodicity and delivery time slots exert effects [1,2,7]; spatially, transport distance and spatio-temporal characteristics of delivery routes are critical[1,2]; at the environmental and operational level, factors such as weather, road congestion, and warehouse efficiency play a role [1,2,4,8]. Regarding order and parcel attributes, order complexity and parcel weight reflect demand variations [4,7,9]. However, these studies have almost exclusively focused on ordinary commodities, failing to account for the unique attributes of home appliances.

In contrast, existing research on home appliance logistics has primarily focused on service quality and customer satisfaction, with insufficient attention paid to time prediction.For instance, Wang and Song [10] validated the positive impact of service quality on customer loyalty, while Zhang and Smutkupt [6] constructed an online logistics service quality evaluation framework to demonstrate the importance of time. However, neither study delved into the factors influencing time or predictive methodologies.

In terms of prediction methods, approaches have evolved from traditional machine learning to deep learning [1,2,11]. However, single models tend to reduce the accuracy of core order predictions when handling data heterogeneity. The advantages of the two-stage classification-regression approach have been validated in high-volatility time series forecasting [12], but this method remains unexplored in home appliance logistics.

Synthesizing these findings, three gaps emerge: first, general logistics overlooks home appliance-specific attributes, while home appliance logistics ignores time prediction; second, existing studies treat delivery time as homogeneous, ignoring interval volatility and amplifying core order prediction errors; third, single models fail to adapt to heterogeneous delivery data, and the classification-regression approach is unapplied. This study addresses these gaps to enhance prediction accuracy.

3 Data Processing and Feature Engineering

After preprocessing—including time filtering and multi-table joins—the dataset was organized into a one-order-per-row format. Through business logic anomaly cleaning and statistical outlier handling, 3.491 million valid samples were retained. Descriptive statistics verified variable distributions. Core factors were identified through correlation analysis, and 10 key variables were retained after removing highly collinear ones. Delivery time exhibited a pronounced right-skewed distribution (see Fig. 1), with significant variations in order proportions and volatility across cumulative time intervals.

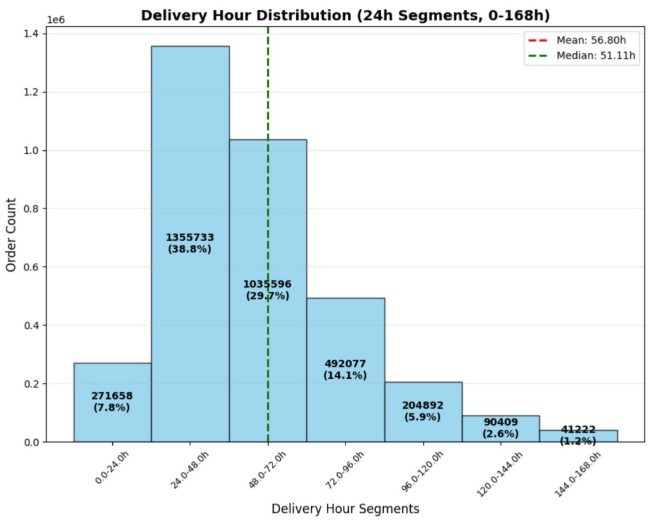

Fig. 1. Distribution of Delivery Time.

Based on core variables, features were expanded across four dimensions: order volume, time derivation, route operations, and interaction combinations. After secondary correlation verification and redundancy removal, 16 input features were finalized (see Table 1), comprehensively covering key influencing factors throughout the entire delivery chain. Among these, the xf_time timestamp was retained to partition the dataset by time series.

Table 1. Model Input Feature Details.

Feature Category	Variables	Definition
Time Features	time_period_detailed	6 time periods within 1 day
	is_evening_peak	Evening peak: 5-7 PM
	is_morning_peak	Morning peak: 7-9AM
	week_type	Weekends/Weekdays
	month	Order month
	xf_time	Order issued and delivered to the Origin Center
Product Features	max_length	Widest dimension of goods within the order
	total_amt	Total number of goods in the order
	total_volume	Total volume of order
Delivery Features	is_same_city	Same-city/Cross-city
	total_distance	Total delivery distance (kilometers)
	distc_dest_org	Direct travel distance from the Destination Center to the last-mile hub (kilometers)
	distance_bin	Five-tier distance distribution label
	route_ODH	(Origin-Destination-Hub) Path Combination
Order Features	daily_order_volume	Daily Order Volume
	route_order_frequency	Orders per Route
	distance_frequency_interaction	Total Distance×Route Order Frequency

To align with the practical scenario of predicting future orders using historical data in logistics operations, the time series segmentation method was employed to divide the dataset into training and testing sets. Using xf_time as the reference, the earlier 80% of samples were allocated to the training set and the later 20% to the testing set, maintaining an 8:2 ratio.

4 Model Construction and Validation

Based on the data analysis conclusions from the preceding sections, this chapter focuses on selecting the core interval for home appliance logistics delivery time prediction and optimizing the model. Through cumulative interval experiments, the optimal core interval is identified. Subsequently, a classification-regression architecture is employed to enhance prediction accuracy in specific scenarios, ultimately forming a modeling system tailored to the characteristics of home appliance logistics.

4.1 XGBoost Algorithm

XGBoost (Extreme Gradient Boosting) is an optimized gradient boosting decision tree algorithm [13]. Its core advantages lie in efficient handling of high-dimensional data and strong anti-overfitting ability, both highly aligned with this study: First, home appliance logistics data includes multi-dimensional features such as time, product, and delivery attributes, and XGBoost can effectively capture the complex relationships between these features and delivery time; Second, there are a small number of abnormal records in the dataset, and XGBoost's anti-overfitting ability prevents the model from being disturbed by such noise.

In this study, XGBoost is used in two stages: it serves as a multi-classifier to assign orders to 5 time sub-intervals, and simultaneously acts as a regression model for each sub-interval to predict specific delivery hours.

4.2 Core Interval Selection Experiment

This section employs the XGBoost regression model as the core algorithm based on the 16 previously identified key features. Validation was conducted five cumulative time intervals: ≤24h, ≤48h, ≤72h, ≤96h, and ≤168h, using Root Mean Square Error(RMSE) and order coverage rate as primary metrics.

Experimental results are shown in Table 2: Although the ≤24h and ≤48h intervals exhibit lower RMSE, their order coverage ratios are only 7.79% and 46.62%, respectively, making them unsuitable for large-scale operational requirements; The ≤96h interval achieved 90.36% order coverage with an RMSE of 12.13h, representing the optimal balance between accuracy and coverage. This interval was therefore identified as the core optimization range.

Table 2. Model Performance and Order Coverage Across Different Cumulative Time Intervals.

Time Interval	RMSE(h)	MAE(h)	Coverage Ratio
≤24	2.71	1.98	7.79%
≤48	6.10	4.13	46.62%
≤72	8.97	6.13	76.27%
≤96	12.13	8.41	**90.36**%
≤total	14.74	11.33	100%

4.3 Design and Results Comparison of Classification-Regression Models

For the core ≤96h interval, leveraging the dense dual-cluster distribution characteristic of 24-48h orders, we departed from traditional equidistant segmentation. Instead, we subdivided it into five sub-intervals: 0-24h, 24-36h, 36-48h, 48-72h, and 72-96h. This enabled a two-stage architecture: first classifying to define scenarios, then applying regression for precise prediction. The classifier pinpoints the specific time-window segment for each order, eliminating cross-interval feature interference. Dedicated regression models are then trained for each segment.

The classification stage employs an XGBoost multi-class model using 16 input features to predict the sub-interval of each order, with the core objective of minimizing cross-interval feature interference. The classifier achieves an overall accuracy of

73.5%, with F1-scores for each sub-interval ranging from 0.63 to 0.77 (see Table 3), validating the rationality of the interval segmentation.

Table 3. Classification Performance Metrics of the Classifier.

Time Window(h)	Accuracy	Recall	F1-score	Sample Proportion
0-24	0.64	0.62	0.63	8.6%
24-36	0.71	0.73	0.72	19.9%
36-48	0.70	0.75	0.72	23.1%
48-72	0.77	0.77	0.77	32.8%
72-96	0.80	0.70	0.75	15.6%

During the regression phase, customized regression models were developed for each of the five subintervals. Hyperparameters were optimized based on the sample size and distribution characteristics of each interval.

4.4 Model Performance Validation and Comparison

Using a single regression model with consistent input features and no interval classification as the baseline, the effectiveness of the fusion model was validated by comparing end-to-end overall accuracy and interval-level accuracy. The results are shown in Table 4:

Table 4. Performance Comparison Between Fusion Model and Regression Model.

Evaluation Dimension	Metric (h)	Regression	**Classification-Regression**	Improvement
End-to-End Performance	RMSE	12.13	11.35	**6.4%**
	MAE	8.53	6.38	**25.2%**
Interval Performance	0-24 RMSE	2.27	2.09	7.9%
	24-36 RMSE	2.05	2.18	-6.3%
	36-48 RMSE	2.16	2.12	1.9%
	48-72 RMSE	4.05	3.66	9.6%
	72-96 RMSE	4.28	3.93	8.1%

The performance advantages of the classification-regression fusion model are fully validated at both global and local dimensions: its end-to-end total RMSE decreases to 11.35 hours, representing a 6.4% reduction compared to pure regression models, while total Mean Absolute Error (MAE) decreases by 25.2%, demonstrating the effectiveness of the classification-regression architecture. At the sub-interval level, while the 24-36h interval shows slight RMSE fluctuations, accuracy improves across the 0-24h, 48-72h, and 72-96h intervals. This stems from the classifier's sample refinement, which eliminates ambiguous boundary samples that disrupt interval pattern learning in pure regression models, enabling more precise regression fitting.

In summary, the complementary validation of global and local accuracy demonstrates that the classification-regression model maintains overall performance advantages while achieving precise fitting across sub-intervals. This fully meets the time-sensitive prediction requirements for the core ≤96h interval in home appliance logistics.

4.5 Model Application Scenarios and Business Value

Pure regression models offer high computational efficiency, suitable for scenarios requiring broad coverage over precision, such as initial order time screening and comprehensive overviews. The classification-regression model delivers high accuracy, aligning with precision-critical needs like core order forecasting and scenario-specific operations. In terms of business value, it optimizes warehouse sorting efficiency within the 24-48 hour window, supplements transit capacity in the 72-96 hour window, eliminates interference from non-standard orders exceeding 96 hours, reduces model maintenance costs, minimizes delivery commitment deviations, and enhances customer trust.

5 Conclusion

This study focuses on home appliance logistics delivery time prediction. First, leveraging the right-skewed distribution of delivery times, it overcomes the limitations of existing subjective interval segmentation. Through cumulative interval experiments, it balances coverage and accuracy, establishing a ≤96h core interval tailored for home appliance logistics. Second, addressing pattern variations among sub-intervals within the core interval, a two-stage classification-regression model is designed, realize scenario-specific modeling.

While the study achieves core order prediction via the above methods, there is room for optimization. The core interval is divided into five sub-intervals based on delivery time distribution and basic business features, but lacks refinement by product attributes, delivery route types and other dimensions. The classification-regression model relies on independent classifiers and sub-interval regression. Features are limited to static data and lack real-time operational data, leading to insufficient generalization and dynamic adaptability.

Aligned with the digitalization trend of home appliance logistics, future improvements will focus on three aspects: 1) Integrate real-time order data and operational iterations to design dynamic threshold adjustment algorithms for automatic optimization of core and sub-interval division; 2) Connect delivery time prediction with end-to-end operational data for overall operational optimization; 3) Incorporate road congestion data into dynamic features, and build online learning models to update delivery time in real-time and improve commitment accuracy.

References

1. Zhu, D., Han, Z., Du, X., Zuo, D., Cai, L., Xue, C.: Hybrid Model Integrating Fuzzy Systems and Convolutional Factorization Machine for Delivery Time Prediction in Intelligent Logistics. IEEE TRANSACTIONS ON FUZZY SYSTEMS 33(1), 406-417 (2025).

2. Araujo, A.C. de, Etemad, A.: End-to-End Prediction of Parcel Delivery Time with Deep Learning for Smart-City Applications. IEEE Internet of Things Journal 8(23), 17043-17056 (2021).
3. Chan, T., Liu, Z., Zhang, W." Delivery service, customer satisfaction and repurchase: Evidence from an online retail platform. Customer satisfaction and repurchase: Evidence from an online retail platform (2018).
4. Salari, N., Liu, S., Shen, Z. J. M.: Real-time delivery time forecasting and promising in online retailing: When will your package arrive?. Manufacturing & Service Operations Management 24(3), 1421-1436 (2022).
5. Guo, X., Yu, Y., Allon, G., Wang, M., Zhang, Z.: RiRiShun logistics: Home appliance delivery data for the 2021 manufacturing & service operations management data-driven research challenge. Manufacturing & Service Operations Management 26(4), 1358-1371 (2024).
6. Zhang, Y., Smutkupt, S.: Investigating key antecedents of logistics services quality towards customer satisfaction and loyalty among ecommerce: a case study of home appliance retailing delivery in China. Journal of Supply Chain Management: Research & Practice 15(1), 16-26 (2021).
7. Wu, F., Wu, L.: DeepETA: a spatial-temporal sequential neural network model for estimating time of arrival in package delivery system. Proceedings of the AAAI conference on artificial intelligence 33(1), 774-781(2019).
8. Tsolaki, K., Vafeiadis, T., Nizamis, A., Ioannidis, D., Tzovaras, D.: Utilizing machine learning on freight transportation and logistics applications: A review. ICT Express 9(3), 284-295 (2023).
9. Leung, K.H., Choy, K.L., Ho, G.T.S., Lee, C.K.M., Lam, H.Y., Luk, C.C.: Prediction of B2C e-commerce order arrival using hybrid autoregressive-adaptive neuro-fuzzy inference system (AR-ANFIS) for managing fluctuation of throughput in e-fulfilment centres. Expert Systems with Applications 134, 304-324 (2019).
10. Wang, S., Yang, Q., Ruan, S., Long, C., Yuan, Y., Li, Q., Yuan, Z., Bao, J., Zheng, Y.: Spatial Meta Learning With Comprehensive Prior Knowledge Injection for Service Time Prediction. IEEE Transactions on Knowledge and Data Engineering 37(2), 936-950 (2025).
11. Issaoui, Y., Khiat, A., Haricha, K., Bahnasse, A., Ouajji, H.: An Advanced System to Enhance and Optimize Delivery Operations in a Smart Logistics Environment. IEEE Access 10, 6175-6193 (2022).
12. Iqbal, M., Iqbal, A., Alshammari, A., Ali, I., Maghrabi, L.A., Usman, N.: Sell or HODL Cryptos: Cryptocurrency Short-to-Long Term Projection Using Simultaneous Classification-Regression Deep Learning Framework. IEEE Access 12, 118169-118184 (2024).
13. Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794 (2016).