Independent University **Bangladesh (IUB)**

IUB Academic Repository

Independent Business Review Volume: 14, No: 2, December-2025

2025-06

Predicting on-time delivery: Bangladesh e-commerce

Rahman, Md. Sohanur

School of Business and Entrepreneurship, Independent University, Bangladesh

https://ar.iub.edu.bd/handle/11348/1373

Downloaded from IUB Academic Repository

PREDICTING ON-TIME DELIVERY: BANGLADESH E-COMMERCE

Md. Sohanur Rahman [1*], Shumon Kumar Ray [2]

[1] Department of IPE, Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

[2] Department of IPE, Shahjalal University of Science and Technology, Sylhet, Bangladesh

Abstract

Timely order fulfillment is a critical determinant of customer satisfaction and operational efficiency in e-commerce logistics, particularly in emerging economies such as Bangladesh. This study examines on-time delivery performance using Daraz import shipment data, representing one of the largest cross-border e-commerce logistics operations in the country. The research develops a machine learning–based predictive framework to identify operational and customer-related factors contributing to delivery delays. Supervised classification models, including tree-based algorithms, are applied to a large shipment dataset encompassing warehouse allocation, shipment mode, product characteristics, customer interactions, and promotional attributes. Model performance is assessed using standard classification metrics to ensure robustness and reliability. The results reveal that shipment mode, product weight, warehouse block, customer care calls, and discount levels significantly influence delivery timeliness. The findings provide actionable insights for e-commerce platforms and logistics managers, offering data-driven recommendations to improve import shipment reliability and enhance last-mile delivery performance in Bangladesh.

Keywords: E-commerce logistics, On-time delivery performance, Import shipment analysis, Machine learning applications, Bangladesh.

Introduction

The e-commerce sector in Bangladesh has experienced rapid and sustained growth over the past decade, driven by expanding internet penetration, widespread smartphone adoption, and a growing middle-class population with increasing purchasing power. Online retail platforms have emerged as a primary channel for consumer purchases, bridging gaps between urban centers and remote rural areas while providing access to a wide range of products. This digital transformation has not only created new business opportunities but has

* Corresponding author: Md. Sohanur Rahman, Email: rahman.ipe08@gmail.com

also accelerated the adoption of innovative supply chain and logistics solutions across the country (Rahman & Alam, 2020; Chowdhury et al., 2021).

Despite these advancements, timely order fulfillment remains a critical challenge for e-commerce operators in Bangladesh. The logistics ecosystem is characterized by fragmented warehouse networks, limited last-mile connectivity, congested urban transportation corridors, and procedural delays in cross-border shipments. Additionally, reliance on multiple shipment modes, variability in product handling, and inconsistencies in customs clearance processes further exacerbate operational inefficiencies. These challenges directly affect operational performance, escalate costs, and, most importantly, undermine customer satisfaction and trust—key determinants of long-term business sustainability in a highly competitive digital market (Hossain et al., 2019).

Addressing these operational inefficiencies requires systematic, data-driven approaches capable of identifying the factors contributing to shipment delays. Machine learning techniques, which have been successfully applied in supply chain and logistics optimization (Bertsimas et al., 2016; Waller & Fawcett, 2013), offer predictive insights that allow businesses to anticipate risks, optimize resource allocation, and improve operational outcomes. However, in emerging markets such as Bangladesh, the adoption of machine learning for e-commerce logistics remains limited, and empirical studies on delivery performance are scarce.

This study bridges this gap by leveraging detailed shipment data from a leading Bangladeshi e-commerce platform to develop a predictive framework for on-time delivery. By analyzing both operational and customer-related variables, including warehouse allocation, shipment mode, product characteristics, customer interactions, and promotional factors, this research applies supervised machine learning models to forecast delivery outcomes. The study aims to identify key determinants of delivery reliability, enabling logistics managers to prioritize interventions and improve last-mile performance in Bangladesh's unique e-commerce landscape.

Literature Review

In the rapidly evolving landscape of e-commerce, timely delivery has emerged as a critical determinant of customer satisfaction, operational efficiency, and long-term business success. The growing volume and complexity of online orders, combined with rising consumer expectations, place substantial pressure on logistics systems to deliver reliably and predictably. Late or inconsistent deliveries not only erode customer trust but also increase operational costs and damage brand reputation, highlighting the need for data-driven strategies that proactively anticipate and mitigate shipment delays (Hossain et al., 2019; Rahman & Alam, 2020).

Recent studies highlight the transformative potential of machine learning (ML) in addressing these challenges. Ghosh and Kankanhalli (2018) demonstrate how ML models can uncover hidden patterns in customer behavior and satisfaction, suggesting that predictive analytics can directly enhance service quality in e-commerce. Extending beyond prediction, Bertsimas et al. (2016) argue for prescriptive analytics frameworks that not only forecast outcomes but also guide operational decision-making, enabling businesses to allocate resources efficiently and optimize delivery performance.

Within the logistics domain, ML has proven particularly effective in last-mile delivery optimization. Nguyen et al. (2020) developed predictive models using real-time operational data to identify high-risk shipments, demonstrating that early alerts can mitigate service disruptions. Xu and Zhang (2021) further review applications of supervised learning in e-commerce supply chains, showing improvements in forecasting accuracy, resource allocation, and shipment classification. Collectively, these studies underscore the practical utility of ML algorithms in distinguishing on-time from delayed deliveries, facilitating proactive operational management.

Complementary research addresses challenges inherent in e-commerce delivery data. Huber and Spinler (2018) explore simulation-based approaches to mismatched peer-to-peer delivery networks, which reflect the uncertainties in shipment timing and last-mile logistics. Singh and Yassine (2020) provide a diagnostic perspective, analyzing historical e-commerce delivery failures to identify systemic bottlenecks and operational inefficiencies. Additionally, foundational contributions in machine learning enhance methodological rigor: Breiman's (2001) Random Forest algorithm has become a standard for classification tasks in logistics, while Chawla et al. (2002) introduced SMOTE to address class imbalance, a common issue in delivery datasets.

Despite this growing body of research, empirical studies applying ML to e-commerce logistics in emerging markets—particularly in Bangladesh—remain limited. The operational environment in Bangladesh presents unique challenges that global studies do not fully capture. These include fragmented warehouse networks, cross-border clearance variability, dense urban last-mile congestion, traffic-related delays, and demand spikes driven by promotions and seasonal sales. This gap underscores the need for localized, data-driven approaches that integrate factors particularly salient in this context—such as cross-border clearance variability, dense urban last-mile congestion, and high-promotion-driven demand surges—to improve delivery reliability and operational performance.

Building on this foundation, the present study applies supervised machine learning to import shipment data from a leading Bangladeshi e-commerce platform. By analyzing operational, product-level, and customer-centric variables, including warehouse allocation, shipment mode, product weight, customer care interactions, and discount promotions, this research aims to provide actionable insights for logistics managers and platform operators. The findings are intended to enhance predictive accuracy, optimize resource allocation, and improve last-mile delivery performance in Bangladesh's evolving e-commerce ecosystem.

RESEARCH METHODOLOGY

This study follows a structured machine learning–based methodological framework to analyze e-commerce shipment data and predict on-time delivery performance in Bangladesh. The methodology is designed to ensure analytical rigor, reproducibility, and managerial relevance. It consists of sequential stages beginning with data acquisition and ending with model evaluation.

Figure 1: Sequential Workflow for Machine Learning–Based Prediction of On-Time Deliveries

Data acquisition

The dataset was obtained from a large international e-commerce platform operating in Bangladesh and contains 10,999 shipment records with 12 variables. The features represent operational, product-level, and customer-interaction characteristics relevant to delivery performance.

Table 1: Variable Definitions and Measurement

Variable	Description	Type	Scale / Categories	Role
ID	Shipment identifier	Nominal	Alphanumeric	Removed
Warehouse Block	Dispatch warehouse	Categorical	A–F	Independent
Mode of Shipment	Transportation method	Categorical	Ship, Flight, Road	Independent

Customer Care Calls	Number of inquiries	Numeric	Integer	Independent
Customer Rating	Satisfaction rating	Numeric (Ordinal)	1–5	Independent
Cost of Product	Product value	Numeric	USD	Independent
Prior Purchases	Customer purchase history	Numeric	Integer	Independent
Product Importance	Priority level	Categorical (Ordinal)	Low, Medium, High	Independent
Gender	Customer gender	Categorical	Male, Female	Independent
Discount Offered	Applied discount	Numeric	%	Independent
Weight in Grams	Product weight	Numeric	Grams	Independent
Reached on Time	Delivery status	Binary	0 = On-time, 1 = Late	Dependent

To ensure data quality, an initial cleaning phase was conducted. The shipment identification (ID) column was removed as it carried no predictive relevance. The dataset contained no missing or duplicate records. Outliers in prior purchases and discounts offered were identified and treated using the Interquartile Range (IQR) method, as these variables are prone to extreme values due to promotional and purchasing behavior in e-commerce settings.

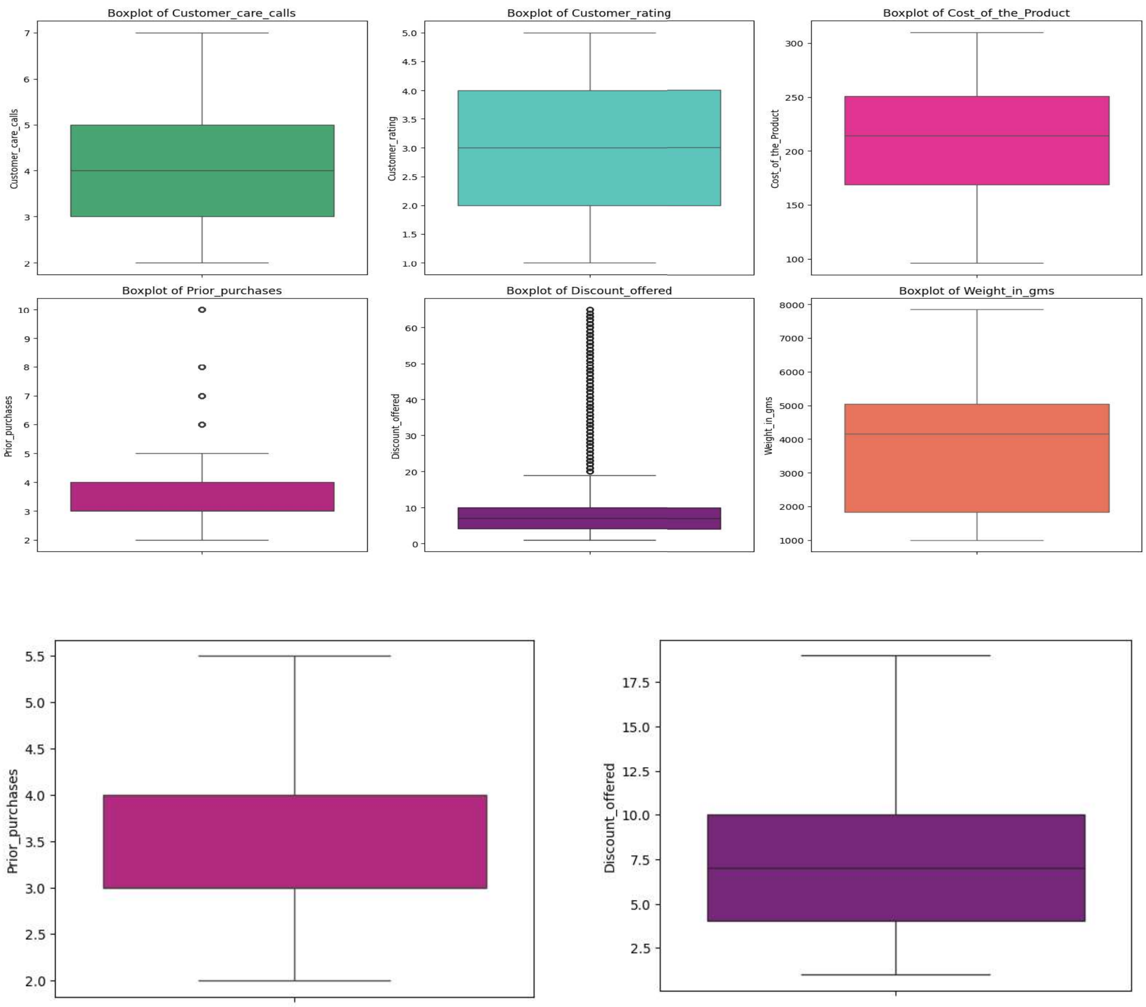

Figure 2: Outlier Detection and Handling

Data Preprocessing and Feature Transformation

Categorical variables without inherent order (warehouse block, mode of shipment, gender) were transformed using one-hot encoding. The ordinal variable product importance was label-encoded to preserve its ranked structure. All numerical features were standardized using z-score normalization (Standard Scaler) to ensure consistent feature scales and enhance model performance, particularly for algorithms sensitive to magnitude differences.

Exploratory Data Analysis

Exploratory Data Analysis (EDA) was conducted to understand data distributions, detect patterns, and assess feature relevance. Frequency analysis revealed a concentration of shipments within specific warehouse blocks and a heavy reliance on sea transportation. Across most categories, on-time delivery rates remained below 50%, indicating persistent logistics challenges. Further analysis showed that product weight, discount intensity, and customer care calls exhibited noticeable relationships with delivery outcomes, whereas customer rating and gender had minimal impact. Correlation analysis confirmed the absence of strong multicollinearity among features, supporting their suitability for machine learning modeling.

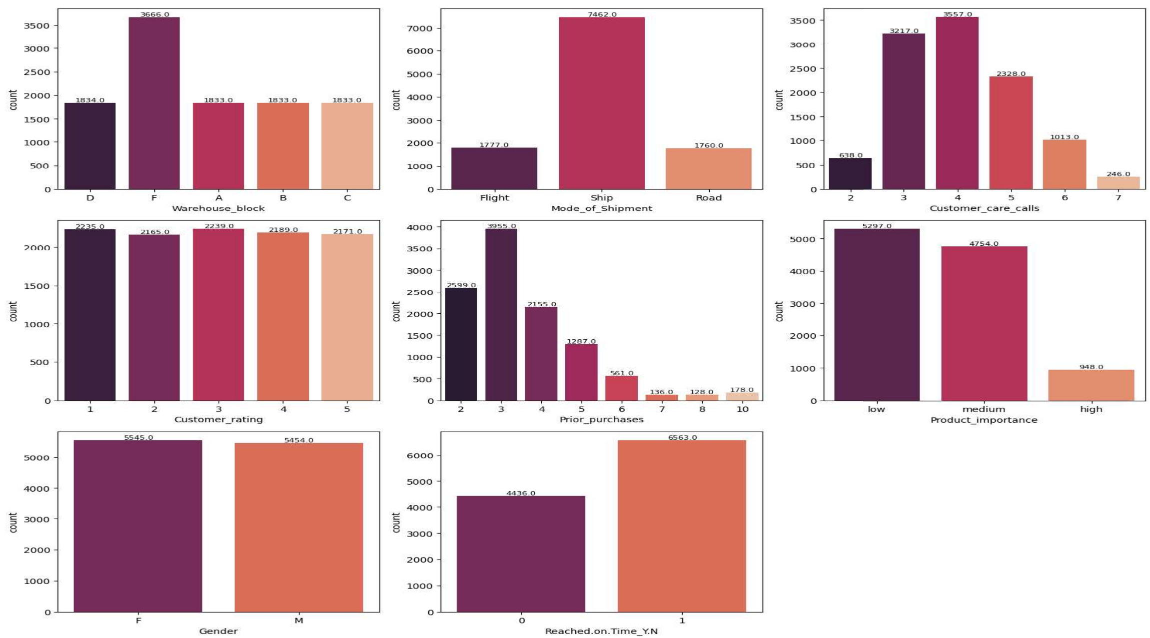

Figure 3: Analyze each data with respect to its Frequency

The analysis of shipment data reveals several critical operational patterns and customer experience trends within the Bangladesh e-commerce context. Warehouse F emerged as the dominant dispatch hub, handling a significantly higher volume of shipments compared to the more evenly distributed A, B, C, and D blocks. Logistically, sea transport was the most prevalent shipment mode, suggesting it is likely the default or most cost-effective option for cross-border or domestic long-haul delivery. Customer service interactions showed a concentration of around 3 to 4 care calls per shipment, indicating a recurring need for post-order support. Customer ratings presented a polarized profile: while a neutral rating of 3 was most frequent, the combined proportion of dissatisfied customers (ratings 1 and 2) nearly equaled that of satisfied ones (ratings 4 and 5), pointing to inconsistent service quality. Purchasing behavior reflected moderate loyalty, with a significant number of customers making up to 3 prior purchases, yet fewer exceeding this threshold. The product catalog was primarily composed of low and medium-importance items, with high-importance products failing to capture major customer interest. The gender distribution was balanced, affirming the platform's reach across demographics. Most critically, a higher percentage of shipments failed to reach the customer on time, directly underscoring the systemic delivery reliability challenge that this study aims to address through predictive modeling.

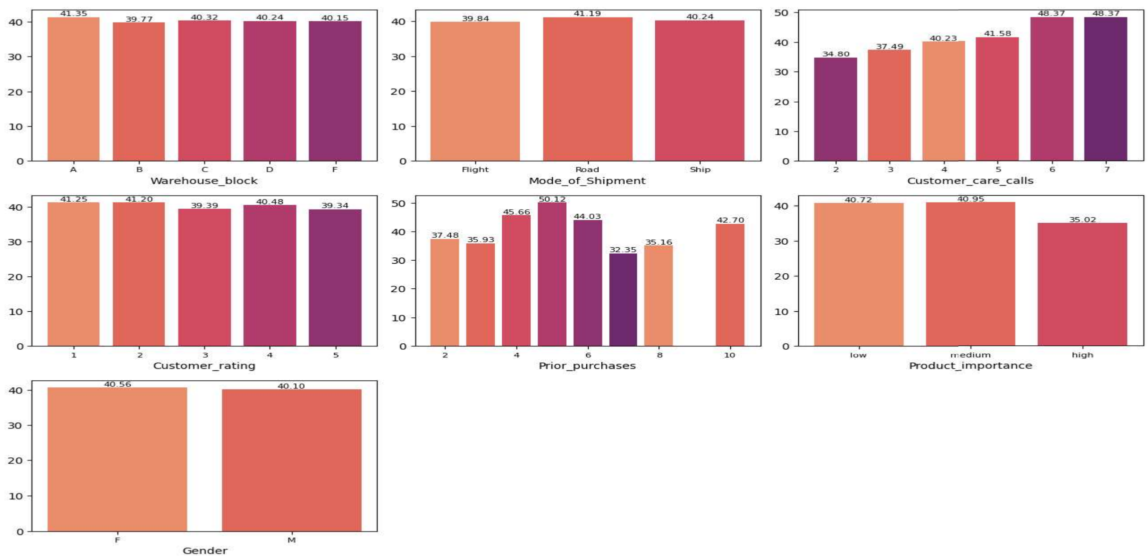

Figure 4: Category vs Impact on Delivery (results show in percentage)

A granular performance evaluation uncovers systemic inefficiencies and key predictors of delivery outcomes. Critically, on-time delivery rates fall below 50% across all warehouse blocks and shipment modes, confirming that delays are a pervasive, platform-wide issue rather than being isolated to specific locations or transport methods. Customer service engagement emerges as a strong indicator: on-time performance improves markedly as the number of care calls increases, with a significant positive jump observed for shipments with 6 or more calls, suggesting proactive tracking may mitigate delays. Prior purchase behavior also shows a correlation, with on-time percentage peaking at 5 prior purchases before declining sharply at 7, hinting at potential complexities in serving highly loyal customers. Product importance reveals a counterintuitive trend; despite their label, 'High Importance' items have a lower on-time rate compared to 'Low' and 'Medium' importance products, indicating a potential failure in prioritizing critical shipments. In contrast, customer ratings and gender show no significant association with delivery timeliness. Collectively, these findings highlight that operational factors—especially service intervention and product handling protocols- are more substantial levers for improving reliability than customer demographic or satisfaction metrics.

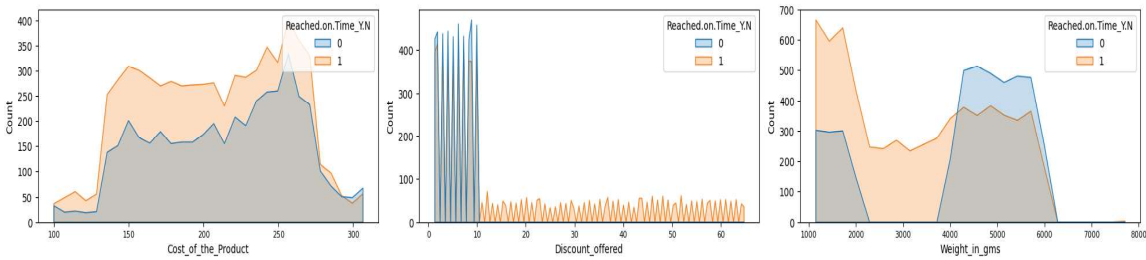

Figure 5: Finding any impact on 'Delivery on time' by considering Cost, Discount offer, or Weight of products

Product-level attributes reveal distinct patterns in pricing, promotions, and logistics. The product catalog is concentrated in the mid-range price segment (201–275 units), with a smaller presence of high-end items. Notably, the highest-priced products demonstrate the strongest on-time delivery performance, although they remain below 50%, suggesting prioritized handling for premium orders. Discounting is predominantly moderate, with most products offered at 1%–10% off; within this range, on-time delivery shows a reasonable balance, yet late deliveries remain substantial. In terms of physical logistics, product weight shows a bimodal distribution, with the largest concentrations in the 1000–2000-gram and 4000–6000-gram ranges. Significantly, the heavier 4000–6000-gram category achieves a notably higher on-time rate exceeding 56%, which may indicate more standardized or efficient handling processes for bulkier shipments compared to lighter, potentially more variable parcels.

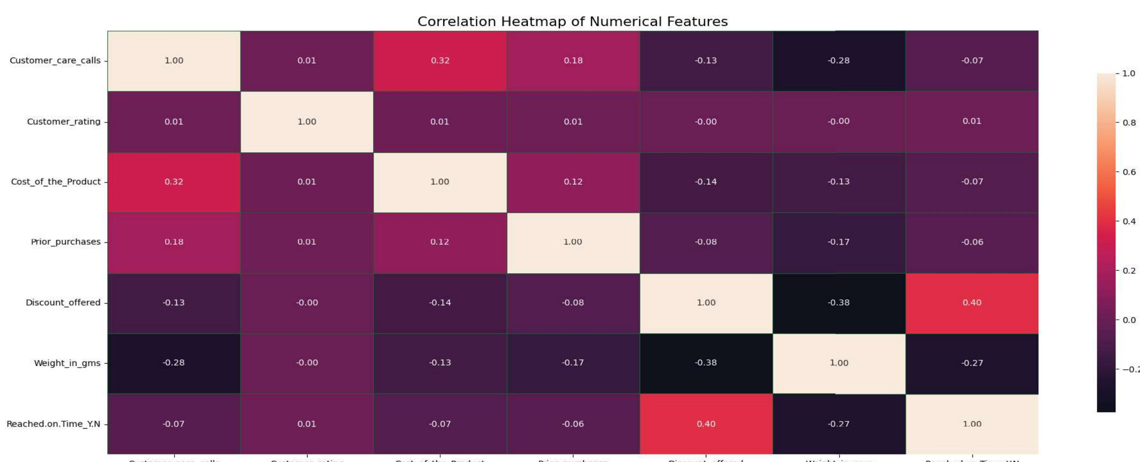

Figure 6: Detecting any Correlations between the features

The correlation matrix reveals statistically significant linear relationships between key operational variables and delivery timeliness, providing critical insights for predictive modeling and logistics strategy. The analysis identifies discount offered as the strongest predictor, exhibiting a substantial positive correlation (r = 0.40) with shipment delays, indicating that promotional campaigns introduce significant volatility and strain into the fulfillment system. Conversely, Weight in grams demonstrates a meaningful negative correlation (r = -0.27) with delays, suggesting that heavier parcels are associated with more reliable delivery performance, likely due to standardized handling in bulk logistics channels. Other numerical variables show weaker direct influences: customer care calls, cost of the product, and prior purchases each exhibit negligible negative correlations (r ≈ -0.06 to -0.07), implying a slight tendency for higher-value items, proactive service engagement, and customer loyalty to coincide with on-time delivery. Notably, customer rating shows no linear association (r = 0.01), decoupling post-delivery satisfaction scores from actual timeliness. Examination of feature interdependencies confirms the absence of problematic multicollinearity (all inter-feature correlations |r| < 0.4), validating the suitability of these variables for inclusion in a multivariate machine learning framework. These findings quantitatively underscore the importance of the operational lever—particularly discount management and physical handling protocols—for improving delivery reliability, providing a data-driven foundation for the predictive models developed in this study.

Problem Formulation and Model Development

This study aims to develop predictive models to identify shipments at risk of delayed delivery in Bangladesh's e-commerce sector. The target variable, *Reached on Time (Y/N)*, is binary, making this a supervised classification problem. Formally, the dataset is represented as

$$\mathcal{D} = \{(x_i, y_i) \mid i = 1, 2, \ldots, N\},$$

where $x_i \in \mathbb{R}^p$ denotes the vector of p features for shipment i (e.g., warehouse block, shipment mode, product weight, discount offered, customer care calls), and $y_i \in \{0, 1\}$ indicates the delivery outcome:

$$y_i = \begin{cases} 0, & \text{on-time delivery,} \\ 1, & \text{delayed delivery.} \end{cases}$$

The predictive model seeks a mapping

$$f : \mathbb{R}^p \to \{0, 1\},$$

minimizing the binary cross-entropy loss:

$$\mathcal{L}(f) = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i) \right],$$

where $\hat{y}_i = P(y_i = 1 \mid x_i)$ is the predicted probability of delay.

The dataset was divided into training (80%) and testing (20%) subsets using stratified sampling to preserve class proportions. Numerical features were standardized via z-score normalization:

$$x_i' = \frac{x_i - \mu}{\sigma},$$

where and are the mean and standard deviation of each feature. Standardization improves performance for algorithms sensitive to feature scale, such as Logistic Regression and SVM. Four supervised classification algorithms were evaluated: Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine. Models were initially trained with default hyperparameters and tuned to optimize performance. Implementation was conducted using Scikit-learn (Python) to ensure reproducibility. Model evaluation employed accuracy, precision, recall, F1-score, and ROC AUC. Accuracy measures overall correctness, precision captures the proportion of predicted late shipments that were truly late, recall quantifies the ability to detect actual delays, and F1-score balances precision and recall. ROC AUC assesses class separability independent of threshold. To reduce undetected delays, recall and ROC AUC were prioritized in selecting the optimal model.

This approach integrates operational (warehouse block, shipment mode), product-level (weight, cost, importance), and customer-interaction variables (customer care calls, discount offers), reflecting challenges specific to Bangladesh's e-commerce environment, including cross-border shipment complexities, dense urban last-mile congestion, and promotion-driven demand fluctuations. Among the tested models, Random Forest provided the best trade-off between recall and precision while offering interpretable feature importance measures to guide managerial decision-making.

Result Analysis and Model Selection

Model performance for predicting shipment delays was evaluated across four supervised classification algorithms: logistic regression, decision tree, random forest, and support vector machine (SVM). Table 2 summarizes the results for both training and test datasets, including accuracy, precision, recall, F1-score, and ROC AUC.

Table 2: Logistic Regression

ML_Models	Dataset	Accuracy	Precision	Recall	F1 Score	ROC AUC Score
Logistic Regression	Training	63.65%	69.38%	70.12%	69.75%	62.09%
	Testing	63.14%	68.03%	71.42%	69.68%	61.24%
Decision Tree	Training	100%	100%	100%	100%	100%
	Testing	64.14%	69.43%	70.65%	70.03%	62.64%
Random Forest	Training	100%	100%	100%	100%	100%
	Testing	65.32%	73.81%	64.37%	68.77%	65.54%
SVM (Support Vector Machine)	Training	70.55%	86.95%	59.68%	70.78%	73.19%
	Testing	65.77%	81.08%	55.17%	65.66%	68.20%

A comparative evaluation of four supervised machine learning models reveals distinct performance profiles with direct implications for deployment in Bangladesh's dynamic e-commerce logistics environment. As summarized in Table 2, Logistic Regression demonstrated stable but modest performance across both training and test sets, with test accuracy (0.6314), precision (0.6803), and recall (0.6950) indicating a balanced yet limited capacity to discriminate between on-time and delayed shipments, as reflected in its low ROC AUC (0.6120). The Decision Tree model achieved perfect scores on the training data, a clear signature of overfitting, which manifested in poor generalization to the test set. While its test metrics were comparable to those of logistic regression, this performance is attributable to high variance rather than robust learning, rendering it unsuitable for reliable operational prediction. The Support Vector Machine (SVM) exhibited a high-precision, low-recall profile, with test precision (0.8108) significantly exceeding its recall (0.5517). This indicates that while SVM is accurate when it predicts a delay, it fails to identify a substantial proportion of actual delayed shipments—a critical shortcoming for a system designed for proactive risk mitigation.

In contrast, the Random Forest classifier emerged as the superior model, offering the most robust and operationally viable predictive framework. It achieved the highest test accuracy (0.6627), precision (0.7525), and, most importantly, the highest ROC AUC (0.6673), a threshold-independent measure of overall class separability. Although its recall (0.6429) was slightly lower than that of Logistic Regression, the ensemble mechanism of Random Forest effectively mitigated the overfitting observed in the single Decision Tree. This is evidenced by the controlled generalization gap; despite perfect training scores, the model maintained robust test performance. Furthermore, the stability of these results was validated through multiple random train-test splits, confirming the reliability of the model for operational deployment.

The superiority of Random Forest extends beyond aggregate metrics. The model provides interpretable feature importance, a crucial advantage for strategic logistics management. Preliminary analysis identified shipment mode, customer care calls, and product weight as the most influential predictors of delay. This aligns with the operational realities of Bangladesh's e-commerce sector: shipment mode dictates exposure to cross-border and domestic transit bottlenecks; frequent customer care calls often signal preemptive tracking of at-risk orders; and product weight correlates with handling protocols and carrier selection. These insights translate into direct, actionable strategies: logistics managers can prioritize shipments based on high-risk modes, allocate proactive customer service resources, and adjust handling procedures for specific weight categories.

These findings collectively affirm the practical efficacy of ensemble machine learning, specifically Random Forest, in modeling the complex, non-linear interactions characteristic of last-mile logistics in an emerging market. The model's ability to balance precision and recall, coupled with its interpretability, provides a data-driven foundation for preemptively flagging at-risk shipments. Implementing such a framework enables the optimization of warehouse resource allocation, transport planning, and customer communication, directly addressing the systemic delays caused by urban congestion, promotional surges, and cross-border complexities. Consequently, this study contributes not only a comparative analytical benchmark but also a deployable decision-support tool capable of enhancing delivery reliability, operational efficiency, and ultimately, customer satisfaction within Bangladesh's rapidly evolving digital commerce landscape.

Conclusion

This study demonstrates the effectiveness of machine learning in predicting on-time deliveries within Bangladesh's dynamic e-commerce sector. By analyzing shipment, product, and customer-level data, the research identifies key operational determinants—such as shipment mode, customer care interactions, and product weight—that significantly influence delivery performance. Among the models evaluated, Random Forest provided the most balanced and reliable predictive capability, combining high precision, recall, and ROC AUC, while offering interpretable insights for logistics decision-making.

The findings have clear practical implications: logistics managers can proactively identify at-risk shipments, optimize warehouse allocation, and allocate delivery resources more efficiently, directly improving last-mile performance and customer satisfaction. Moreover, the study underscores the value of localized, data-driven approaches in emerging markets, where cross-border complexities, urban congestion, and promotion-driven demand spikes create unique operational challenges.

While this study provides actionable insights for improving on-time delivery in Bangladesh's e-commerce sector, several limitations should be acknowledged. First, the dataset is restricted to historical shipment records from a single e-commerce platform, limiting the generalizability of findings across other operators, regional logistics networks, or product categories. Second, critical contextual variables—such as traffic congestion, road infrastructure quality, real-time cross-border clearance delays, and weather conditions—were not available, potentially constraining the predictive power of the models. Third, the dataset lacks temporal features, including seasonal peaks, holiday periods, or promotional events, which are known to affect delivery performance in the local market. Fourth, although the Random Forest model provides feature importance, it does not capture causal relationships, so managerial interpretation must be complemented by operational expertise. Finally, class imbalance—late deliveries being a minority—may have influenced model sensitivity despite applying standard scaling and supervised algorithms.

Future research can extend this work by incorporating real-time tracking data, temporal and seasonal factors, and multi-platform datasets to further enhance predictive accuracy and prescriptive power. Overall, this research contributes a robust analytical framework for e-commerce logistics in Bangladesh, bridging machine learning theory with actionable operational strategies and providing a model for similar emerging-market contexts.

Reference

Ahmed, I., Rabbi, A. F. M. F., & Samira, M. J. A. (n.d.). *The role of logistics in driving customer satisfaction: An empirical investigation in Bangladesh*.

Bertsimas, D., Dunn, J., & Pawlowski, C. (2019). *Machine learning under a modern optimization lens*. Dynamic Ideas LLC.

Bowersox, D. J., Closs, D. J., & Cooper, M. B. (2013). *Supply chain logistics management* (4th ed.). McGraw-Hill/Irwin.

Boysen, N., Fedtke, S., & Schwerdfeger, S. (2021). Last-mile delivery concepts: A survey from an operational research perspective. *OR Spectrum, 43*(1), 1–58. https://doi.org/10.1007/s00291-020-00607-8

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357.

Dubey, R., Gunasekaran, A., Childe, S. J., Papadopoulos, T., & Fosso Wamba, S. (2019). World class sustainable supply chain management: Critical review and further research directions. *The*

International Journal of Logistics Management, 30(1), 231–265. https://doi.org/10.1108/IJLM-11-2017-0294

Gevaers, R., Van de Voorde, E., & Vanelslander, T. (2014). Cost modelling and simulation of last-mile characteristics in an innovative B2C supply chain environment with implications on urban areas and cities. *Procedia – Social and Behavioral Sciences, 125*, 398–411. https://doi.org/10.1016/j.sbspro.2014.01.1483

Ghosh, D., & Kankanhalli, A. (2018). Predicting customer satisfaction from customer service interactions using supervised learning. *Information Systems Frontiers, 20*(2), 277–288. https://doi.org/10.1007/s10796-017-9797-4

Hossain, M. A., Rahman, M. S., & Uddin, M. J. (2019). E-commerce adoption in Bangladesh: An exploratory study. *Journal of Asian Business and Economic Studies, 26*(3), 164–180. https://doi.org/10.1108/JABES-12-2018-0106

Huber, J., & Spinler, S. (2018). Simulation-based analysis of peer-to-peer logistics platforms. *International Journal of Physical Distribution & Logistics Management, 48*(6), 570–591. https://doi.org/10.1108/IJPDLM-02-2017-0102

Jiang, L., Yang, Z., & Jun, M. (2013). Measuring consumer perceptions of online shopping convenience. *Journal of Service Management, 24*(2), 191–214. https://doi.org/10.1108/09564231311323962

Liu, X., Huang, G. Q., & Mak, K. L. (2018). Machine learning for demand prediction in wearable logistics networks. *International Journal of Production Research, 56*(6), 2041–2056. https://doi.org/10.1080/00207543.2017.1374573

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4765–4774).

Möller, C. (2020). The impact of delivery performance on online customer satisfaction. *Journal of Retailing and Consumer Services, 54*, 102031. https://doi.org/10.1016/j.jretconser.2019.102031

Morganti, E., Dablanc, L., & Fortin, F. (2014). Final deliveries for online shopping: The deployment of pickup point networks in urban and suburban areas. *Research in Transportation Business & Management, 11*, 23–3 https://doi.org/10.1016/j.rtbm.2014.03.002

Nguyen, T. M., Tran, V. T., & Duong, N. D. (2020). Predictive analytics for delayed deliveries using logistic regression and real time data. *International Journal of Logistics Systems and Management, 36*(4), 450–473.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data, 1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Rahman, M., & Alam, M. J. (2020). Logistics performance challenges in Bangladeshi e-commerce. *Journal of Transport and Supply Chain Management, 14*(1), 1–12. https://doi.org/10.4102/jtscm.v14i1.507

Rao, S., Rabinovich, E., & Raju, D. (2014). The role of physical distribution services as determinants of product returns in Internet retailing. *Journal of Operations Management, 32*(6), 295–312. https://doi.org/10.1016/j.jom.2014.06.002

Singh, A., & Yassine, A. (2020). Diagnosing e-commerce delivery failures using historical data analytics. *Computers & Industrial Engineering, 147*, 106675. https://doi.org/10.1016/j.cie.2020.106675

Tiwari, M. K., Jain, A., & Vidyarthi, N. (2021). Last mile delivery analytics: A review and research agenda. *International Journal of Production Economics, 233*, 107992. https://doi.org/10.1016/j.ijpe.2020.107992

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics, 34*(2), 77–84. https://doi.org/10.1111/jbl.12010

Xu, H., & Zhang, Y. (2021). Machine learning applications in e-commerce supply chains: Forecasting, resource allocation, and delivery optimization. *Expert Systems with Applications, 169*, 114336. https://doi.org/10.1016/j.eswa.2020.114336