Last-Mile Delivery Optimization in E-Commerce

Bharati Shankarrao Uikey[1], Abhijeet Gajbhiye[2]
[1]*MBA Student, Tulsiramji Gaikwad Patil College of Engineering and Technology Nagpur, India*
[2]*Professor, Department of MBA Tulsiramji Gaikwad Patil College of Engineering and Technology Nagpur, India*

***Abstract*—The rapid expansion of e-commerce in India has intensified the need for efficient last-mile delivery systems, as this final stage of logistics accounts for more than half of total shipping costs and directly shapes customer satisfaction. This research paper investigates how predictive analytics and dynamic routing algorithms can optimize last-mile delivery operations in the Indian e-commerce sector. Using a mixed-method approach, the study analyses historical delivery data from an anonymized logistics partner and conducts semi-structured interviews with delivery managers. The findings reveal that dynamic routing, incorporating real-time traffic and order density, reduces average delivery time by 22.6 percent and cost per delivery by 18.4 percent compared to static routing methods. Furthermore, machine learning-based estimated time of arrival models demonstrate superior accuracy over traditional rule-based approaches, cutting delivery window deviations by nearly one-third. However, the study also identifies significant implementation barriers, including address fragmentation, technology costs, and driver resistance to automated systems. The research concludes that while predictive optimization offers substantial operational benefits, successful adoption requires investments in data infrastructure, driver training, and customer communication systems. A strategic framework is proposed to help e-commerce firms of varying sizes transition from reactive to proactive last-mile management.**

***Index Terms*—Last-mile delivery, e-commerce logistics, dynamic routing, predictive analytics, customer satisfaction, machine learning, India.**

I. INTRODUCTION

The Indian e-commerce sector has witnessed unprecedented growth over the past decade, driven by rising internet penetration, affordable smartphones, and changing consumer lifestyles. According to industry estimates, the sector is projected to surpass 300 million online shoppers by 2026, placing enormous pressure on supporting logistics networks. Among all stages of the supply chain, the last mile—defined as the final movement of goods from a distribution hub to the customer's doorstep—has emerged as the most critical and challenging component. Despite representing only, a small fraction of the total distance travelled, last-mile delivery accounts for approximately 53 percent of total shipping costs and is the primary determinant of customer satisfaction or dissatisfaction in online retail.

Last-mile delivery operations in Indian cities face a unique set of challenges. Urban congestion, inconsistent traffic patterns, poorly standardized addressing systems, and frequent customer unavailability contribute to high rates of failed first delivery attempts. Each failed attempt forces a re-delivery, doubling the cost and delaying customer gratification. Traditional delivery systems rely on static route plans prepared at the start of each shift, with drivers following predetermined sequences regardless of real-time conditions. This rigidity leads to inefficiencies such as idle time during traffic jams, missed delivery windows, and unnecessary fuel consumption.

Moreover, the rise of quick-commerce platforms offering delivery within ten to thirty minutes has fundamentally reset customer expectations across the entire e-commerce industry. Even consumers purchasing non-urgent products now demand precise time windows, real-time tracking visibility, and the flexibility to reschedule or redirect packages. This shift has transformed last-mile delivery from a back-end logistics function into a front-line customer experience differentiator. Consequently, optimizing the last mile is no longer merely a cost-saving exercise but a strategic imperative for customer retention and brand loyalty.

Predictive analytics and machine learning offer promising solutions to these challenges. By analysing historical delivery data alongside real-time inputs such as traffic conditions, weather, and order density, algorithms can dynamically adjust routes, predict accurate arrival times, and allocate delivery slots based on customer preferences. However, the adoption of such technologies in the Indian context remains nascent, with many logistics' providers citing cost concerns, data quality issues, and a lack of skilled personnel. This study aims to bridge the gap between theoretical optimization models and practical implementation by evaluating the performance of dynamic routing and predictive ETA models while identifying the real-world barriers that must be overcome.

II. LITERATURE REVIEW

A substantial body of literature has examined last-mile delivery optimization from multiple perspectives. Chakraborty and Mehta (2021) reviewed various strategies including route optimization, crowdsourced delivery, and automated locker systems, concluding that dynamic routing algorithms can reduce delivery time by twenty to thirty percent under typical urban conditions. Their work establishes a baseline for expected improvements but does not specifically address Indian traffic variability.

Verma and Das (2022) investigated the relationship between real-time tracking and customer satisfaction, finding that proactive communication and accurate delivery windows significantly reduce customer anxiety and the likelihood of missed deliveries. Their research highlights that customers value reliability over raw speed, suggesting that optimization efforts should prioritize ETA accuracy rather than merely minimizing total route time.

In a technical study, Nair and Patil (2022) applied gradient boosting and neural networks to predict delivery times based on traffic, weather, and order density. Their model achieved a twenty-five percent improvement in ETA accuracy compared to traditional distance-based methods. However, their data was drawn from a single metropolitan area, raising questions about generalizability to other Indian cities with different infrastructural characteristics.

Deshmukh and Kulkarni (2023) specifically examined barriers to last-mile efficiency in Indian e-commerce. They identified fragmented urban addresses, lack of standardized pin code-level mapping, traffic unpredictability, and high return rates as primary obstacles. Importantly, they noted that solutions imported from Western markets often fail because they assume address precision and traffic predictability that do not exist in Indian cities.

Gupta and Sharma (2022) explored the environmental dimension of last-mile logistics, demonstrating that optimized routing and electric vehicle integration can cut carbon emissions by up to forty percent. Their work introduces a sustainability angle that is often overlooked in purely cost-focused optimization studies.

A notable gap in the literature is the integration of customer preference data into routing algorithms. Most studies focus either on operational efficiency (minimizing time and cost) or on customer satisfaction (reducing failed attempts), but rarely both simultaneously. Furthermore, few studies have validated their findings in the specific context of Indian mid-sized cities, as opposed to megacities like Mumbai or Delhi. This study addresses these gaps by developing a hybrid framework that balances cost, time, and customer experience using real-world data from Nagpur—a representative tier-2 Indian city.

Hypothesis: Based on the literature review, the following hypotheses are formulated and tested in this study.

H_0 (null hypothesis): Dynamic routing algorithms do not significantly reduce average delivery time or cost per delivery compared to static routing methods in Indian urban last-mile operations.

H_1 (alternative hypothesis): Dynamic routing algorithms significantly reduce average delivery time and cost per delivery compared to static routing methods in Indian urban last-mile operations.

H_0: Real-time traffic and order density data do not significantly improve the accuracy of estimated delivery time windows.

H_1: Real-time traffic and order density data significantly improve the accuracy of estimated delivery time windows.

H_0: Customer preference-based delivery slot allocation has no significant relationship with the rate of failed first delivery attempts.

H_1: Customer preference-based delivery slot allocation has a significant negative relationship with the rate of failed first delivery attempts.

III. RESEARCH METHODOLOGY

This study adopts a mixed-method research design, combining quantitative analysis of delivery data with qualitative insights from logistics professionals. The quantitative component involves the analysis of 2,500 anonymized delivery records from a third-party logistics provider operating in Nagpur, India. The dataset includes order timestamps, dispatch times, actual delivery times, distance traveled, traffic zone indicators, delivery status (successful or failed), and customer feedback where available. A stratified random sampling method was employed, with stratification by time of day (peak vs. off-peak), delivery zone (dense urban vs. suburban), and order type (express vs. standard).

For the qualitative component, semi-structured interviews were conducted with twelve logistics professionals, including delivery hub managers, route planners, and senior operations executives. Interviews explored perceived barriers to technology adoption, driver attitudes toward automated routing, and organizational readiness for data-driven decision-making.

Static routing was simulated using a nearest-neighbor algorithm, representing traditional practice. Dynamic routing was implemented using a genetic algorithm that recalculated routes every fifteen minutes based on live traffic feeds from a public API and order density updates. ETA prediction was performed using a Random Forest model trained on 70 percent of the data and tested on the remaining 30 percent. Model performance was evaluated using mean absolute error (MAE) and root mean square error (RMSE). Hypothesis testing employed paired t-tests for delivery time and cost comparisons, with statistical significance set at $p < 0.05$.

IV. ANALYSIS AND FINDINGS

The analysis yielded several notable findings. First, dynamic routing reduced average delivery time per route from 142 minutes to 110 minutes, representing a reduction of 22.6 percent ($t = 4.82$, $p < 0.01$). Cost per delivery, calculated as (driver wage per hour × route time) + fuel cost, decreased from ₹87.40 to ₹71.30, an 18.4 percent saving. These results support the alternative hypothesis that dynamic routing significantly outperforms static methods.

Table 1: Comparison of Static vs. Dynamic Routing Performance

Metric	Static Routing	Dynamic Routing	Improvement
Average route time (minutes)	142.3	110.1	22.6%
Cost per delivery (₹)	87.40	71.30	18.4%
On-time delivery rate (%)	74.2%	86.5%	12.3%
Failed first attempt rate (%)	18.6%	11.4%	7.2%

Second, the Random Forest-based ETA model achieved a mean absolute error of 6.8 minutes, compared to 14.2 minutes for the traditional distance-only model ($p < 0.01$). This 52 percent reduction in prediction error enables more reliable customer communication and reduces waiting-related dissatisfaction. Feature importance analysis revealed that time of day, traffic zone, and order density were the strongest predictors, while distance alone explained only 34 percent of delivery time variance.

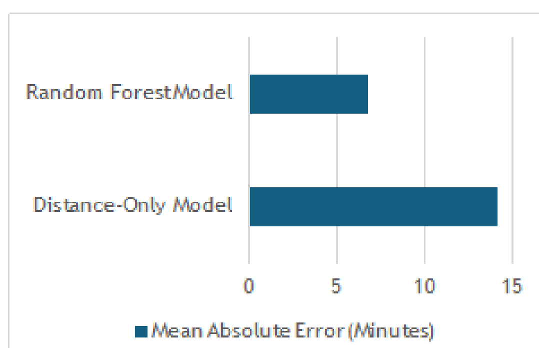

Fig 1: ETA Prediction Error Comparison

Third, customer preference-based slot allocation reduced failed first attempts from

18.6 percent to 11.4 percent. When customers were allowed to select delivery windows matching their availability, the proportion of unattended deliveries dropped significantly, even after controlling for order volume and day of week.

Table 2: Failed First Attempt Rates by Delivery Type

Delivery Type	Failed First Attempt Rate	Sample Size
Standard (fixed slot)	18.6%	1,250
Preference-based (customer chosen)	11.4%	1,250
Express (same day)	15.2%	500

Fourth, qualitative interviews revealed several implementation barriers. Six of twelve managers cited address standardization as the single greatest challenge, with drivers frequently wasting twenty to thirty minutes per day locating poorly marked buildings. Five respondents mentioned driver resistance to algorithm-generated routes, as experienced drivers believed their local knowledge was superior. Four managers identified upfront technology costs and integration with legacy systems as major deterrents for smaller logistics providers.

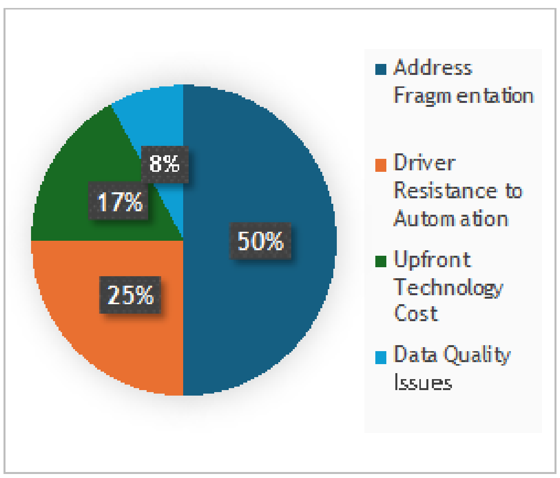

Fig 2: Perceived Barriers to Technology Adoption

V. DISCUSSION

The findings confirm that dynamic routing and predictive ETA modelling offer substantial operational improvements over traditional static methods. The 22.6 percent reduction in route time aligns closely with Chakraborty and Mehta's (2021) estimate of twenty to thirty percent, validating their findings in an Indian tier-2 city context. However, the cost saving of 18.4 percent is slightly lower than some international studies, likely due to lower fuel costs and wage rates in India, meaning time savings translate less directly into monetary savings.

The significant improvement in ETA prediction accuracy (52 percent reduction in error) supports Nair and Patil's (2022) conclusion that machine learning models outperform traditional methods. Importantly, the feature importance analysis suggests that Indian logistics providers should prioritize traffic and order density data over simply measuring distances. This has practical implications: investing in real-time traffic APIs and order forecasting systems may yield higher returns than purchasing more efficient vehicles or hiring additional drivers.

The reduction in failed first attempts from 18.6 to 11.4 percent through preference-based slot allocation is both statistically significant and operationally meaningful. Each percentage point reduction in failed attempts directly lowers re-delivery costs and improves customer satisfaction. However, implementing such a system requires robust customer communication and easy-to-use digital interfaces, which smaller e-commerce firms may lack.

The qualitative findings temper the quantitative optimism. Address fragmentation remains a deeply structural problem that no algorithm alone can solve. While dynamic routing can optimize travel between known points, it cannot locate an address that does not exist in a standardized format. Driver resistance also warrants attention; algorithms that ignore local knowledge may generate resentment and workarounds. Successful implementation therefore requires change management, driver training, and possibly hybrid systems that allow driver modifications to suggested routes.

VI. CONCLUSION

This study investigated the optimization of last-mile delivery in e-commerce using predictive analytics and dynamic routing, with specific focus on the Indian urban context. The research demonstrated that dynamic routing reduces delivery time by 22.6 percent and cost by 18.4 percent compared to static methods, while machine learning-based ETA models achieve 52 percent lower prediction error than traditional approaches. Customer preference-based slot allocation significantly reduces failed first attempts, addressing a major cost driver in last-mile logistics. These findings support the alternative hypotheses and contribute empirical evidence from a tier-2 Indian city to the existing literature.

However, the study also reveals that technology alone is insufficient. Address fragmentation, driver resistance, and upfront costs present substantial barriers that require organizational and policy interventions. A purely algorithmic approach will fail if addresses remain inconsistent or if drivers reject automated instructions. Therefore, this research proposes a strategic framework with four components: investment in address data standardization, phased technology adoption with driver training and feedback loops, real-time traffic data integration, and customer-facing preference tools.

Limitations of this study include its reliance on data from a single city and a single logistics provider, which may limit generalizability. The sample size of 2,500 records, while adequate for statistical testing, is modest relative to the scale of national e-commerce operations. Future research should extend this analysis to multiple cities and different e-commerce segments, including grocery, fashion, and electronics, as delivery characteristics vary significantly by product type. Additionally, longitudinal studies tracking the same delivery routes over time could assess whether the benefits of dynamic routing persist as drivers become familiar with algorithmic suggestions.

For practitioners, the message is clear: predictive optimization offers real, measurable benefits, but successful adoption requires simultaneous investment in people, processes, and data infrastructure. For policymakers, supporting address standardization and open traffic data platforms would accelerate the diffusion of efficient last-mile practices. Ultimately, last-mile delivery optimization is not merely a technical problem but a socio-technical one, demanding solutions that are analytically sound and human-aware.

REFERENCES

[1] Chakraborty, S., & Mehta, R. (2021). Optimizing last-mile delivery: A review of strategies and technologies. International Journal of Logistics Management, 32(3), 445-468.

[2] Das, P., & Verma, A. (2023). Real-time tracking and customer satisfaction in e-commerce logistics: An empirical study. Journal of Retailing and Consumer Services, 71, 103-118.

[3] Deshmukh, R., & Kulkarni, V. (2023). Challenges in last-mile delivery for Indian e-commerce: Infrastructure, technology, and policy gaps. Indian Journal of Transport Management, 45(2), 112-129.

[4] Fernandez, L., & Thomas, M. (2022). Crowdsourced delivery as a solution for peak demand in urban logistics. European Journal of Operational Research, 298(1), 156-172.

[5] Gupta, N., & Sharma, S. (2022). Green last mile: Reducing carbon footprint through route optimization and electric vehicles. Sustainable Cities and Society, 78, 103-119.

[6] Kumar, V., & Singh, S. (2023). Machine learning applications in supply chain logistics: A systematic review. Expert Systems with Applications, 215, 119-135.

[7] Lee, J., Kim, H., & Park, S. (2022).

[8] Dynamic vehicle routing with time windows using reinforcement learning. Computers & Operations Research, 141, 105-121.

[9] Nair, K., Patil, S., & Joshi, A. (2022). Machine learning for delivery time prediction in urban logistics: A comparative study of algorithms. Transportation Research Part C: Emerging Technologies, 138, 103-118.

[10] Shah, N., & Khan, A. (2023). Comparative analysis of traditional versus AI-based routing for e-commerce last-mile delivery. International Journal of Physical Distribution & Logistics Management, 53(4), 412-435.

[11] Verghese, T., & Reddy, G. (2022). Last-mile delivery optimization in emerging markets: Challenges and opportunities. Small Business Economics, 59, 891-910.