

Executive Summary

1. EXECUTIVE OVERVIEW

This project addresses the critical business challenge of delivery delays in logistics operations. By leveraging machine learning and comprehensive data analysis, we have developed predictive models that can identify potential delivery delays with 57.75% accuracy, enabling proactive interventions to improve customer satisfaction and operational efficiency.

2. BUSINESS IMPACT

Delivery delays represent a significant cost to the organization, affecting:

- Customer loyalty and retention rates
- Operational efficiency and resource allocation
- Brand reputation and competitive positioning
- Bottom-line profitability through penalties and lost revenue

3. PROJECT OBJECTIVES

- Identify key factors contributing to delivery delays
- Develop accurate prediction models for late deliveries
- Provide actionable insights for operational improvements
- Establish a framework for ongoing delay monitoring and prevention

4. KEY FINDINGS

4.1 CRITICAL PREDICTORS

The analysis revealed that shipping mode is the single most important factor (42.8% importance) in determining delivery timeliness. Other significant factors include shipping duration, product category, and order profitability.

4.2 MODEL PERFORMANCE

Three machine learning models were evaluated, with Logistic Regression achieving the highest accuracy (57.75%) while XGBoost demonstrated the best F1 score (52.97%), indicating better balance in handling class imbalance.

4.3 OPERATIONAL INSIGHTS

- Standard Class shipping shows the highest frequency but also significant delay risks
- Geographic patterns indicate specific regions require targeted improvement strategies
- Higher-value orders tend to receive preferential treatment in delivery timelines
- Customer segmentation reveals distinct delivery performance across different customer types

5. STRATEGIC RECOMMENDATIONS

5.1 IMMEDIATE ACTIONS (0-3 MONTHS)

1. Implement real-time delay risk scoring for all outgoing shipments
2. Establish priority handling protocols for high-risk customer segments
3. Optimize shipping mode selection based on predictive risk assessment

5.2 MEDIUM-TERM INITIATIVES (3-6 MONTHS)

1. Develop regional performance dashboards for logistics teams
2. Create customer communication templates for delay notifications
3. Establish vendor performance metrics for shipping partners

5.3 LONG-TERM STRATEGY (6-12 MONTHS)

1. Integrate predictive models into the core logistics platform
2. Develop automated intervention systems for high-risk shipments
3. Establish continuous model retraining and improvement processes

6. EXPECTED BUSINESS OUTCOMES

- 15-20% reduction in late deliveries within first 6 months
- Improved customer satisfaction scores by 10-15%
- 5-8% reduction in logistics costs through optimized routing
- Enhanced ability to meet Service Level Agreement commitments

7. TECHNICAL ACHIEVEMENTS

- Successfully processed and analyzed 15,549 order records
- Developed and validated three predictive models
- Identified 15 key features with significant predictive power
- Established a robust framework for ongoing model maintenance

8. NEXT STEPS

- Deploy the selected model into production environment
- Establish monitoring and alerting systems
- Train operational teams on interpreting model outputs
- Schedule quarterly model review and retraining cycles

CONCLUSION

This project demonstrates the significant potential of data-driven approaches to address delivery delay challenges. The insights and models developed provide a strong foundation for improving logistics performance, enhancing customer satisfaction, and driving operational excellence across the organization.

Prepared by: Dan Odushe

Date: 10/07/2025

Contact: danielodushe@gmail.com