Food Delivery Service Customer Churn Analysis Report

Executive Summary

This analysis examines customer churn in a food delivery service, where churn is defined as a customer who has not placed an order in the last 3 months. The study identifies key behavioral patterns that predict customer churn and evaluates the effectiveness of two logistic regression models at predicting which customers are likely to stop using the service.

My analysis reveals that delivery issues, payment problems, and order frequency are all significant predictors of customer churn. The combination of these three factors provides substantially better predictive performance than using delivery issues alone. This report outlines my methodology, findings, and recommendations for reducing customer attrition.

Data Analysis Findings

Comparison of Behavioral Patterns

The boxplots and statistical tests reveal significant differences between customers who churned and those who remained active:

1. **Number of Orders**: Active customers place significantly more orders than those who churn. The mean number of orders for active customers is 31.16 compared to 21.77 for churned customers ($p < 0.001$). This indicates a strong relationship between engagement frequency and retention.
2. **Delivery Issues**: Customers who churned experienced more delivery problems. The mean number of delivery issues for churned customers is 2.39 compared to 1.60 for active customers ($p < 0.001$). This suggests service quality directly impacts customer retention.
3. **Payment Issues**: Churned customers encountered more payment problems. The mean number of payment issues for churned customers is 1.12 compared to 0.79 for active customers ($p < 0.001$). This indicates that technical friction in the payment process contributes to customer attrition.

All these differences are statistically significant with p-values well below 0.001, confirming these factors are strongly associated with churn behavior.

Predictive Modeling Results

Model 1: Single Predictor (Delivery Issues)

The first model used only delivery issues as a predictor of churn. The results show:

- **Accuracy**: 59.0%
- **Precision**: 60.3%
- **Recall**: 70.4%
- **F1 Score**: 65.0%

The visualization of this model shows a clear linear relationship between the number of delivery issues and churn probability. As delivery issues increase from 0 to 4, the predicted probability of churn increases from approximately 39% to over 75%. This confirms that delivery issues are a significant predictor of customer churn.

Model 2: Multiple Predictors (Delivery Issues, Payment Issues, Number of Orders)

The second model incorporated three predictors: delivery issues, payment issues, and number of orders. The results show:

- **Accuracy**: 69.0%
- **Precision**: 69.5%
- **Recall**: 75.9%
- **F1 Score**: 72.6%

The multivariable model substantially outperforms the single-predictor model across all metrics, with a 10 percentage point improvement in accuracy and a 7.6 percentage point improvement in F1 score. The ANOVA test comparing the two models yielded a p-value of 5.971e-15, confirming that the addition of payment issues and number of orders significantly improves the model's predictive power.

The visualization of predicted probabilities by number of orders shows a negative correlation - as the number of orders increases, the probability of churn decreases. However, this relationship is less linear than with delivery issues, suggesting a more complex interaction between engagement frequency and churn behavior.

Key Coefficients and Their Interpretation

The coefficients from the multiple predictor model provide valuable insights:

- **Delivery Issues** (coefficient = 0.497): Each additional delivery issue increases the log odds of churn by 0.497, indicating a significant positive association with churn.
- **Payment Issues** (coefficient = 0.637): Each additional payment issue increases the log odds of churn by 0.637, showing an even stronger positive association with churn than delivery issues.
- **Number of Orders** (coefficient = -0.060): Each additional order decreases the log odds of churn by 0.060, confirming that more frequent customers are less likely to churn.

All three predictors are highly significant with p-values below 0.001, confirming their importance in predicting churn behavior.

Conclusions and Recommendations

Based on my analysis, I recommend the following actions to reduce customer churn:

1. **Improve Delivery Reliability**: The strong correlation between delivery issues and churn suggests that addressing delivery problems should be a priority. Investing in delivery infrastructure, better tracking systems, or partnering with more reliable delivery services could significantly reduce churn.
2. **Resolve Payment Processing Issues**: Payment problems have an even stronger association with churn than delivery issues. Streamlining the payment process, offering more payment options, and proactively addressing failed payments could help retain customers.
3. **Encourage Frequent Ordering**: Customers who place more orders are less likely to churn. Implementing loyalty programs, personalized recommendations, or subscription models could increase order frequency and improve retention.
4. **Early Intervention System**: Develop an early warning system based on the predictive model to identify customers at high risk of churning. This would allow for targeted interventions before customers leave the service.
5. **Segmented Retention Strategy**: Different segments of customers may have different churn drivers. Consider developing targeted retention strategies based on customer segments and their specific pain points.

The improved performance of the multivariable model demonstrates that customer churn is a multifaceted problem requiring a comprehensive approach. By addressing delivery issues, payment problems, and encouraging frequent engagement, the company can substantially improve customer retention and long-term profitability.