Executive summary templates

Use the Layout dropdown menu to select a template or build your own using these layouts as inspiration.

Title

Subtitle

Project Overview

Details

Key Insights

Image Alt-Text Here

Next Steps

Title

Subtitle

ISSUE / PROBLEM

Image Alt-Text Here

RESPONSE

IMPACT

Image Alt-Text Here

KEY INSIGHTS

Title

Subtitle

ISSUE / PROBLEM

RESPONSE

KEY INSIGHTS

IMPACT

Image Alt-Text Here

Title

Subtitle

OVERVIEW

PROJECT STATUS

KEY INSIGHTS

NEXT STEPS

Optimizing Taxi Driver Earnings Through Predictive Tipping Analysis

Executive Summary 6: Data-Driven Insights for Enhanced Revenue and Driver Satisfaction

Overview

This project aimed to develop a machine learning model to predict generous tipping behavior (tips ≥ 20%) among taxi customers, enabling drivers to optimize their earnings. By analyzing trip characteristics and other relevant features, we sought to provide drivers with actionable insights to inform their decision-making. The Random Forest model emerged as the most effective, demonstrating a robust balance between precision and recall, crucial for minimizing both missed opportunities and driver disappointment.

Problem

Taxi drivers rely heavily on tips to supplement their income. However, predicting which customers will tip generously is challenging. Incorrect predictions can lead to wasted time and resources, or missed opportunities for higher earnings. The client, the New York City Taxi & Limousine Commission (TLC), sought a data-driven solution to this problem, aiming to provide a tool for drivers to better anticipate tipping behavior.

Solution

We developed and evaluated two machine learning models: Random Forest and XGBoost. After thorough testing and comparison, the Random Forest model was selected due to its superior F1-score and recall. This model leverages trip data, time of day, day of week, and vendor information to predict the likelihood of a customer being a generous tipper.

Details

Model Selection: The Random Forest model achieved an F1-score of 0.7235 and a recall of 0.7791 on the test data, indicating a strong ability to identify generous tippers.

Key Features: Vendor ID, predicted fare, mean trip duration, and mean trip distance were identified as the most influential factors in predicting generous tips.

Error Analysis: The model exhibited a higher rate of false positives than false negatives, meaning it is more likely to predict a generous tip when one is not given. This is considered acceptable, as it is better for the driver to be pleasantly surprised than disappointed.

Ethical Considerations: We emphasized the importance of transparency and fairness in model deployment, ensuring that it is used to assist drivers, not discriminate against customers.

Model Performance Table:
The table below presents the evaluation metrics for both the Random Forest (RF) and XGBoost (XGB) models, assessed through cross-validation (CV) and on the held-out test dataset. Key metrics include:

- **Precision:** The proportion of predicted generous tippers who were actually generous.
- **Recall:** The proportion of actual generous tippers who were correctly identified.
- **F1-Score:** The harmonic mean of precision and recall, providing a balanced measure.
- **Accuracy:** The overall proportion of correct predictions.

The Random Forest model demonstrated superior performance, particularly in recall and F1-score, making it the recommended choice.

	model	precision	recall	F1	accuracy
0	RF CV	0.674919	0.757312	0.713601	0.680233
0	RF test	0.675297	0.779091	0.723490	0.686538
0	XGB CV	0.669726	0.723553	0.695512	0.666557
0	XGB test	0.677219	0.745488	0.709716	0.679004

Next Steps

To ensure successful implementation and continued improvement, we recommend the following actions:
Pilot: Deploy the Random Forest model in a pilot program for driver feedback.
Data: Gather cash tip data and customer history to improve accuracy.
Features: Explore real-time and driver-specific features.
Monitor: Continuously monitor and retrain the model.
Train: Educate drivers on model use and limitations.
Review: Conduct regular ethical reviews.

Title

Subtitle

Overview

Objective

Results

Next Steps