

Predicting Generous Taxi Tips

Analysis of Factors Influencing Rider Gratuity

➤ ISSUE / PROBLEM

- **Challenge:** NYC TLC needs to understand what factors drive generous tipping ($\geq 20\%$) to help improve driver earnings and service quality across the taxi system.
- **Business Need:** Identify actionable insights that can guide policy decisions and driver support programs to increase overall tip rates and driver satisfaction.

➤ RESPONSE

Solution: Built a Random Forest machine learning model using 2017 NYC taxi trip data to predict generous tippers and identify key influencing factors.

Approach:

- Analyzed 22M+ taxi trips with credit card payments.
- Engineered features including trip characteristics, temporal patterns, and vendor information.
- Developed classification model with 73% accuracy across all performance metrics.
- Conducted comprehensive bias and ethical assessment.

➤ KEY INSIGHTS

- **Primary Findings:**
 1. **Vendor Identity Dominates** - Taxi company (*VendorID*) is the strongest predictor (60%+ importance)
 2. **Trip Value Matters** - Higher fares and longer durations strongly correlate with generous tips
 3. **Time Has Minimal Impact** - Rush hours, days of week, and months show negligible influence
 4. **Group Dynamics** - Passenger count slightly affects tipping behavior
- **Model Performance:** Consistent 73% accuracy, precision, recall, and F1-score indicate reliable predictions without overfitting.

➤ IMPACT

Immediate Benefits:

- **System-Wide Insights:** Clear understanding of tip-driving factors for policy development.
- **Driver Support:** Evidence-based guidance for maximizing earnings potential.
- **Service Optimization:** Focus resources on high-impact areas (longer trips, premium services).

Strategic Recommendations:

- **Promote Credit Card Payments** - Strongest predictor of generous tipping.
- **Vendor Quality Programs** - Address significant company-based differences
- **Zone-Based Incentives** - Target high-fare areas during peak demand.
- **Bias Monitoring** - Implement ongoing fairness audits to prevent discrimination.
- **Ethical Deployment:** Model suitable for system-wide insights and policy guidance, but requires safeguards against individual customer targeting to ensure equitable service.

