

Non-Technical Presentation: Analyzing Terry Stops and Arrest Decisions

Background

Terry Stops are brief interactions initiated by law enforcement based on reasonable suspicion of criminal activity. These encounters carry significant implications for community trust and effective policing. Our analysis aims to identify the factors influencing arrest decisions during these stops and develop predictive models to improve reliability and fairness.

Key Questions

1. How do race and other demographic factors impact arrest decisions?
2. What variables most strongly predict the likelihood of an arrest?
3. How can insights from this analysis guide improvements in policing practices?

Objectives

- Analyze police stop data to uncover patterns, including potential racial profiling and arrest trends.
 - Develop predictive models to support informed decision-making.
 - Enhance community trust through transparency and data-driven practices.
-

Metrics of Success

1. **Accuracy Score:** Measures the proportion of correct predictions. Higher accuracy reflects overall model reliability.
 2. **Precision:** Represents the proportion of correct positive predictions. High precision ensures fewer false arrests.
-

Data Understanding

The dataset, sourced from police records of Terry Stops, includes key features such as demographics, stop reasons, and outcomes.

- **Dataset Overview:** Rows represent individual stops, while columns capture their characteristics.
 - **Descriptive Statistics:** Distribution, averages, and counts of critical variables were analyzed.
-

Data Preparation and Analysis

Data Quality Checks: Addressed missing values, duplicates, and outliers to ensure accurate analysis.

- **Missing Values:** Filled where feasible or excluded incomplete records.
- **Duplicates:** Removed redundant entries to prevent skewed results.
- **Outliers:** Investigated to determine relevance or exclusion.

Exploratory Data Analysis (EDA): Explored relationships between variables like race, gender, and stop reasons to uncover arrest patterns.

Modeling

We developed two models:

1. **Baseline Model:** Logistic Regression
 - Straightforward and interpretable, ideal for binary outcomes (arrest or no arrest).
2. **Advanced Model:** Decision Trees
 - Captures non-linear relationships and feature interactions for deeper insights.

Why These Models?

Logistic Regression provides a simple starting point, while Decision Trees offer nuanced analysis and improved flexibility.

Evaluation

Model Comparison: Both models were assessed on accuracy, precision, and recall.

- **Best Performing Model:** Decision Trees outperformed Logistic Regression in accuracy and recall, effectively identifying true arrests.
 - **Precision:** Decision Trees also had fewer false positives, demonstrating better reliability.
-

Recommendations

1. **Continuous Monitoring:** Regularly evaluate model performance to maintain fairness and effectiveness.
 2. **Data Enrichment:** Include additional contextual factors, such as community demographics, to refine predictions.
 3. **Stakeholder Training:** Equip law enforcement with skills to interpret model outputs responsibly and understand their implications.
-

Conclusion

This analysis of Terry Stops using logistic regression and decision trees highlights critical factors influencing arrest decisions. Data-driven insights can support fair and informed policing practices, fostering trust within communities. Continuous improvement of these models is essential to mitigate biases and adapt to evolving societal dynamics.

Thank you so much.