

Data Report

Stop Resolution: The Art of Prediction.

Business Understanding.

The Terry Stops presents a business opportunity to improve the efficiency and fairness of law enforcement actions. By developing a predictive model that can assist officers in determining the likelihood of an arrest being made during a Terry Stop, the law enforcement agencies can make informed decisions and potentially reduce the number of false arrests and incidents of police misconduct. However, it is important to approach this problem with caution and transparency, considering the ethical concerns raised by the use of gender and race data. The goal is to provide a tool that can help improve policing, while avoiding biases and discrimination.

Objectives.

Main Objective.

To create a predictive model for Terry Stops that accurately predicts the outcome of the stop (arrest made or not).

Specific Objectives.

To take into consideration key factors such as the presence of weapons and the time of the call in the model.

To ensure that the model is ethically sound and avoids any biases or discrimination related to gender and race.

Data Understanding.

Data Source.

This dataset was provided by the City of Seattle and is managed by the Seattle Police Department. It was created on April 13, 2017 and last updated on February 6, 2023.

Data Description.

The dataset contains **54873** rows and **23** columns, each row representing a unique Terry Stop record as reported by the officer conducting the stop. The columns in the dataset include information about the subject of the stop, such as the perceived age group, perceived race, and perceived gender.

The dataset also includes information about the officer, such as the officer's gender, race, and year of birth. Additionally, the dataset includes information about the resolution of the stop, any weapons found, the date and time the stop was reported, and information about the underlying Computer Aided Dispatch (CAD) event. The data is updated daily and is licensed under the public domain.

Features

- Subject Age Group
- Terry Stop ID
- Stop Resolution
- Weapon Type
- Officer Gender
- Call Type
- Arrest Flag
- Frisk Flag

- Year
- Month
- Officer Age
- Time Category
- Location

Data Preparation.

Loading the data.

The data set was imported into a Jupyter notebook for processing. The data frame was displayed to visualize the data and to examine the relationships between the variables.

Cleaning data.

The data was analyzed and cleaning was done by checking for duplicates and missing values. Which were then removed. The categorical data types were converted to numerical data types and feature engineering to as part of data preprocessing.

Modelling.

The data was analyzed using machine learning techniques, specifically logistic regression and random forest. The logistic regression model was used to identify the most important features that predict the outcome of a Terry stop, while the random forest model was used to build a more complex model that can capture non-linear relationships between features.

Results.

The results of the analysis showed that the most important features for predicting the outcome of a Terry stop are the arrest flag, frisk flag, type of call, and year of the Terry stop. Other important features include the month of the stop and the individual's race and gender. The machine learning model was able to accurately predict the outcome of Terry stops with an overall accuracy of around 75%.

Conclusion

The results of this analysis provide valuable insights into the factors that are most important in determining the outcome of a Terry stop. These insights can be used to inform the development of policies and procedures to help ensure that Terry stops are conducted fairly and equitably. It is important to note, however, that the analysis should be regularly reviewed and updated as new data becomes available, to ensure that the results remain relevant and accurate over time.