Implementation of Stop and Frisk in New York (2012)

David Koh

Introduction

One of the foundations of policing under former NYPD police commissioner Ray Kelly and former mayor Michael Bloomberg is a technique called “stop and frisk” which gave police officers in New York City the ability to observe individuals, stop them based on a reasonable suspicion, and to frisk them as well if the police officers still believed something was going on. This is a technique that has been gaining popularity over the years and has recently been highlighted for its controversial implementation.

The purpose of this project is to use data science techniques to analyze how “stop and frisk” was implemented in New York City in 2012.

Datasets

The New York City Police Department (NYPD) requires paperwork to be filed for each stop. In 2012 there were 532,911 stops. I was able to obtain a dataset from the New York Civil Liberties Union (NYCLU) website. For each stop the officer documented details about the suspect’s race, sex, location, time of stop, etc. using Form UF-250. There are 2 ways the NYPD reports this stop-and-frisk data: a paper report released quarterly and an electronic database released annually.

1. I explored this data to see what types of people were being stopped, frisked, and if this actually led to arrests.
2. I also used the data to try and predict the kinds of people who were most likely to be frisked.

Approach

For my analysis I primarily used the *pandas* package in Python to perform statistical analysis. After cleaning the data I was able to perform some basic analysis based by focusing on sex and race. From the analysis it is clear to see that disregarding all other variables, there are a lot of Black men who have been stopped, at nearly 50% of the total population. For comparison there are 1.8M Black men in New York City and in 2012 there were stops of 250,000 Black men. The second most stopped group of people was White Hispanic men. These two groups combine for nearly 70% of the total number of people stopped in 2012.

Furthermore, the percentages of men and women by race who were subsequently frisked and then arrested were similar to the percentages of the people who were stopped. Choosing to focus on a comparison of Black men versus White men the percentage of stops in Black men led to 5% of those stopped being arrested while the percentage of stops in White men that led to arrests was higher at 6%.

Classification of Those Who Were Frisked

To determine whether or not a person would be frisked I decided to use 2 models: Random Forests and Logistic Regression. I prepared the data by going through and removing features that would be highly correlated with a frisk such as those features that stated “reason for frisk” as these features would always match up with those who were frisked. There were also several categorical features, such as “hair color”, “race”, “sex”, and “borough.” For each of these I created dummy features so that eventually I was left with 60 different features.

After the data was cleaned and prepped I fit the data to a Random Forests model. Using the default score function I was able to get an accuracy of over 98% but this was too good to be true. After performing a 5-fold cross validation the accuracy dropped to 73% even though this would still prove to be the most accurate model to predict whether or not someone was frisked. I also tried to implement PCA with 15 components before using Random Forests. Doing so decreased the accuracy to 67%.

For comparison I also tried performing a Logistic Regression. Using Logistic Regression I was able to get an accuracy of 56%. To see how it would be affected using PCA I used that in conjunction with Logistic Regression which again decreased the accuracy, this time to 52%.

Therefore in my analysis Random Forests was the best model in predicting whether or not a person was frisked. Since PCA creates new features (aka components) to make predictions I decided to compare the features used without PCA against just using “race”, “sex”, and “age” with Random Forests, which yielded an accuracy of XX%. Although using “race”, “sex”, and “age” with Random Forests still did a better job of predicting than the other combinations using more features helped better predict whether or not someone was frisked.

Conclusion

When it comes to crime and society as a whole, as the technology and methods used to catch criminals becomes better and more widely available we must be ever more vigilant that they are not used to discriminate. Ironically with this push towards colder, more “scientific” ways of policing it may be important now more than ever to allocate resources properly so that we are building stronger communities and working together to prevent crime rather than stepping backwards in time to discriminate against certain peoples based on prejudices and biases that will not go away.

References:

1. <http://www.theatlantic.com/features/archive/2014/03/is-stop-and-frisk-worth-it/358644/>
2. <http://www.nyclu.org/content/stop-and-frisk-data>
3. <http://www.nyc.gov/html/nypd/html/analysis_and_planning/historical_nyc_crime_data.shtml>
4. <http://www.nyc.gov/html/dcp/html/census/demo_tables_2010.shtml>