01/04/2021

Prepared by: Marvin Mills

Analyzing Terry Stop Data To Predict “Frisk” Event

AGENDA

01 BACKGROUND & CONTEXT

Business Case
Dataset Context

02 EXPLORATORY DATA ANALYSIS

Pre-model Insights
& Visualizations

03 MODELING

Best Performing Model
Post-modeling Analysis
Can “Frisk Events” Be Predicted?

04 RECOMMENDATIONS + CONCLUSION

BUSINESS CASE

- **Filed & Won** case in NYC (2013) where judge ruled that New York City Police department was liable for a pattern and practice of racial profiling when it came to “stop & frisk”

- **Goal** -- analyze for “stop & frisk” bias in Terry Stop data, and identify whether a “Frisk Event” can be predicted, to support their ongoing efforts to combat injustice

CENTER FOR CONSTITUTIONAL RIGHTS

ABOUT THE DATA

TERRY STOPS – DATA.GOV

Publicly Available
No use of any API
> 45,000 entries
Timeframe: 2015-2020
1 entry = 1 “stop”

MISSING DETAILS

> 13,000 entries for “Call Type” where no info on the “Call Type” was provided

CLASS IMBALANCE

With respect to:

Arrest Flag
Frisk Flag (~1:3.4)

EDA - VARIABLES

The Predictor: Frisk / No Frisk

Additional goal: to do soft analysis for "police quotas" and "time of day" effect

Time_Of_Day	Reported_Quarter	Reported_Year
MO	Q4	2015
MO	Q1	2015
EV	Q1	2015
OV	Q2	2015
OV	Q2	2015
...	...	...
OV	Q4	2020
EV	Q4	2020
AF	Q4	2020
MO	Q4	2020
EV	Q4	2020

Added Features

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45196 entries, 0 to 45195
Data columns (total 23 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Subject Age Group         45196 non-null  object
 1   Subject ID                45196 non-null  int64
 2   GO / SC Num               45196 non-null  int64
 3   Terry Stop ID             45196 non-null  int64
 4   Stop Resolution           45196 non-null  object
 5   Weapon Type               45196 non-null  object
 6   Officer ID                45196 non-null  object
 7   Officer YOB               45196 non-null  int64
 8   Officer Gender            45196 non-null  object
 9   Officer Race              45196 non-null  object
 10  Subject Perceived Race    45196 non-null  object
 11  Subject Perceived Gender  45196 non-null  object
 12  Reported Date             45196 non-null  object
 13  Reported Time             45196 non-null  object
 14  Initial Call Type         45196 non-null  object
 15  Final Call Type           45196 non-null  object
 16  Call Type                 45196 non-null  object
 17  Officer Squad             44612 non-null  object
 18  Arrest Flag               45196 non-null  object
 19  Frisk Flag                45196 non-null  object
 20  Precinct                  45196 non-null  object
 21  Sector                    45196 non-null  object
 22  Beat                      45196 non-null  object
dtypes: int64(4), object(19)
memory usage: 7.9+ MB
```

All Variables

Class Imbalance of the Variable "Frisk Flag" in dataset

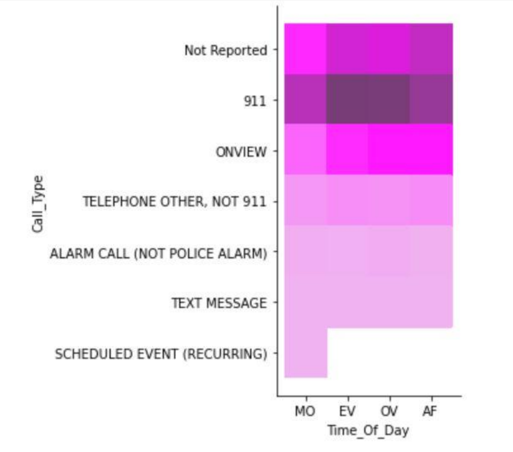
N = Not Frisked
Y = Frisked

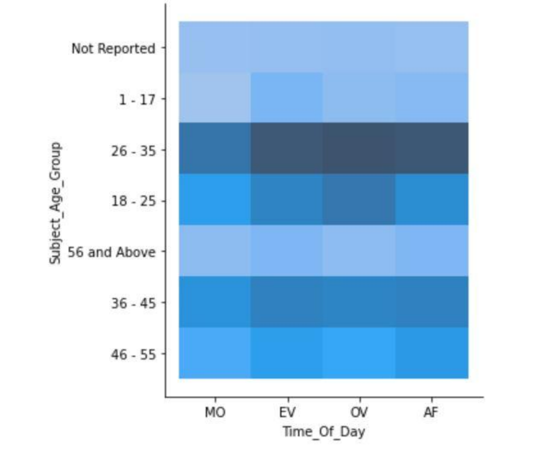
```
N     34660
Y     10058
-       478
Name: Frisk Flag, dtype: int64
```

FRISK VS. ARREST

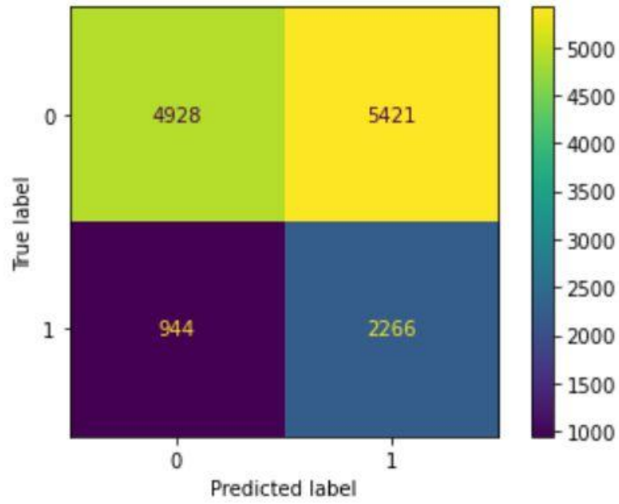
EDA – INSIGHTS

#1: 911 Calls in the evening / overnight
#2: Ages 26-35 in the evening / overnight





MODELING

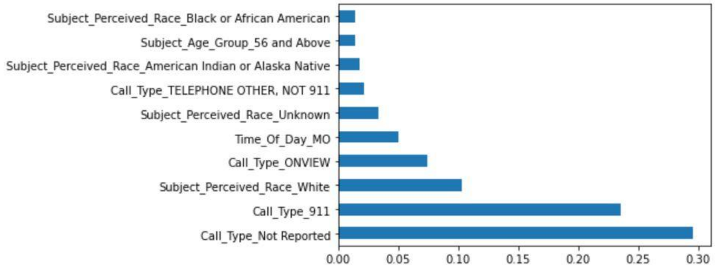



```
]: scoring(y_test, forest_clf_pred)

Accuracy score is: 0.5305701010398997
---
F1 score is: 0.4158942828301367
Precision score is: 0.294783400546377
Recall score is: 0.7059190031152648
```

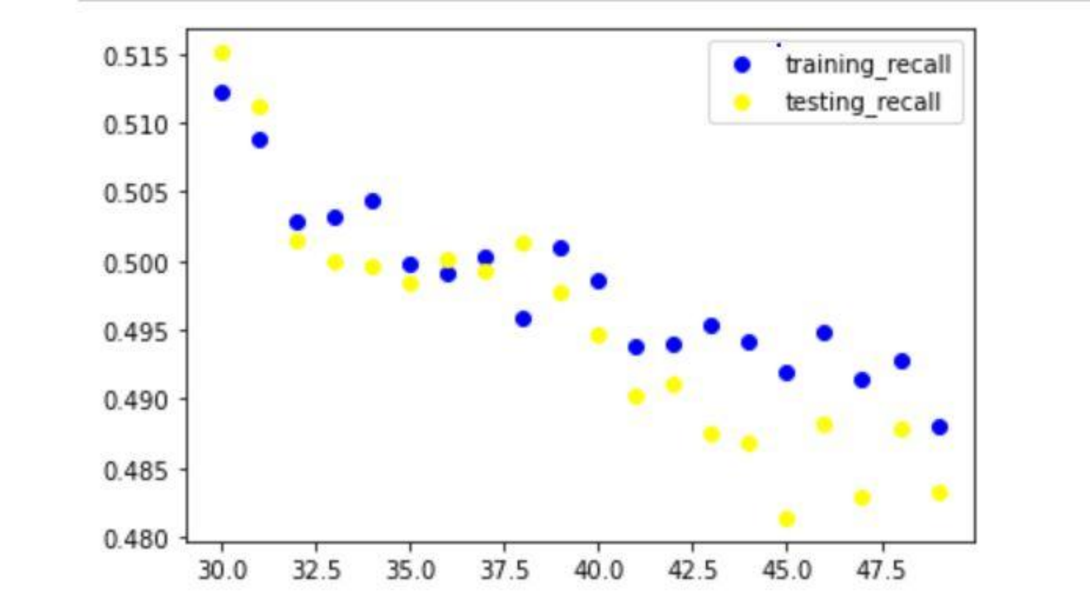
Baseline: higher accuracy and lower for all other scores

GOAL: **improve F1 & Recall scores**



Feature: **Call Type Not Reported**

MODELING



LOGISTIC REGRESSION + SMOTE DATA

As the test size increased, the SMOTE Data decreased, and Recall score showed the strongest sensitivity to this effort.

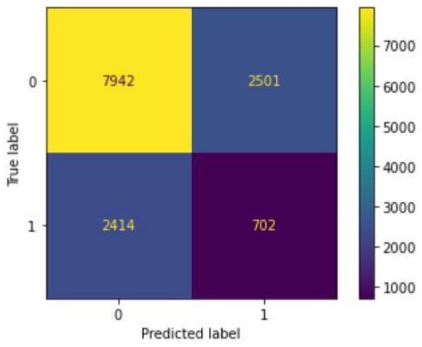
As True Positives go up, False Positives go up

MODELING

F1, Recall & Precision

```
Accuracy score is: 0.6472453720775868
---
F1 score is: 0.23655227454110136
Precision score is: 0.23780487804878048
Recall score is: 0.23531279771355987
```

```
]: plot_confusion_matrix(dummy_clf, y_test, dum_predictor)
plt.show()
```



Post-Modeling Analysis

Key Takeaways:

- Slight hyperparameter tuning didn't improve model significantly
- Models seem optimized for accuracy

CONCLUSION

To Close:

Recommendations & Next Steps

1. Analysis shows “frisk events” are not predictable given the data construction
2. Use dataset more skewed to non-emergency stops
3. Over & Undersampling using SMOTE

Any Questions?
Thank you.
github.com/elem333

CREDITS

Appreciation:

- Presentation template by Slidesgo
- Icons by Flaticon
- Data.gov - Terry Stops Data
- Yish Lim, Flatiron Instructor