



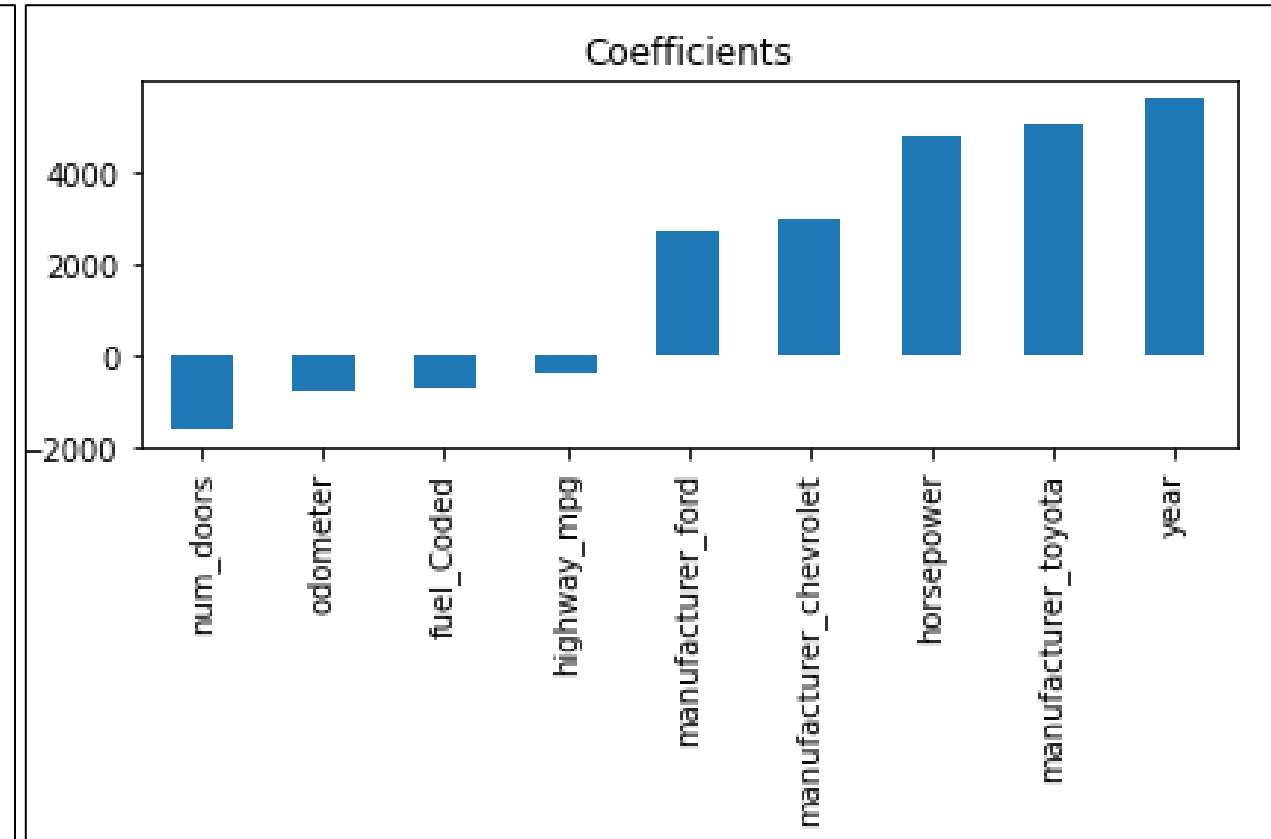
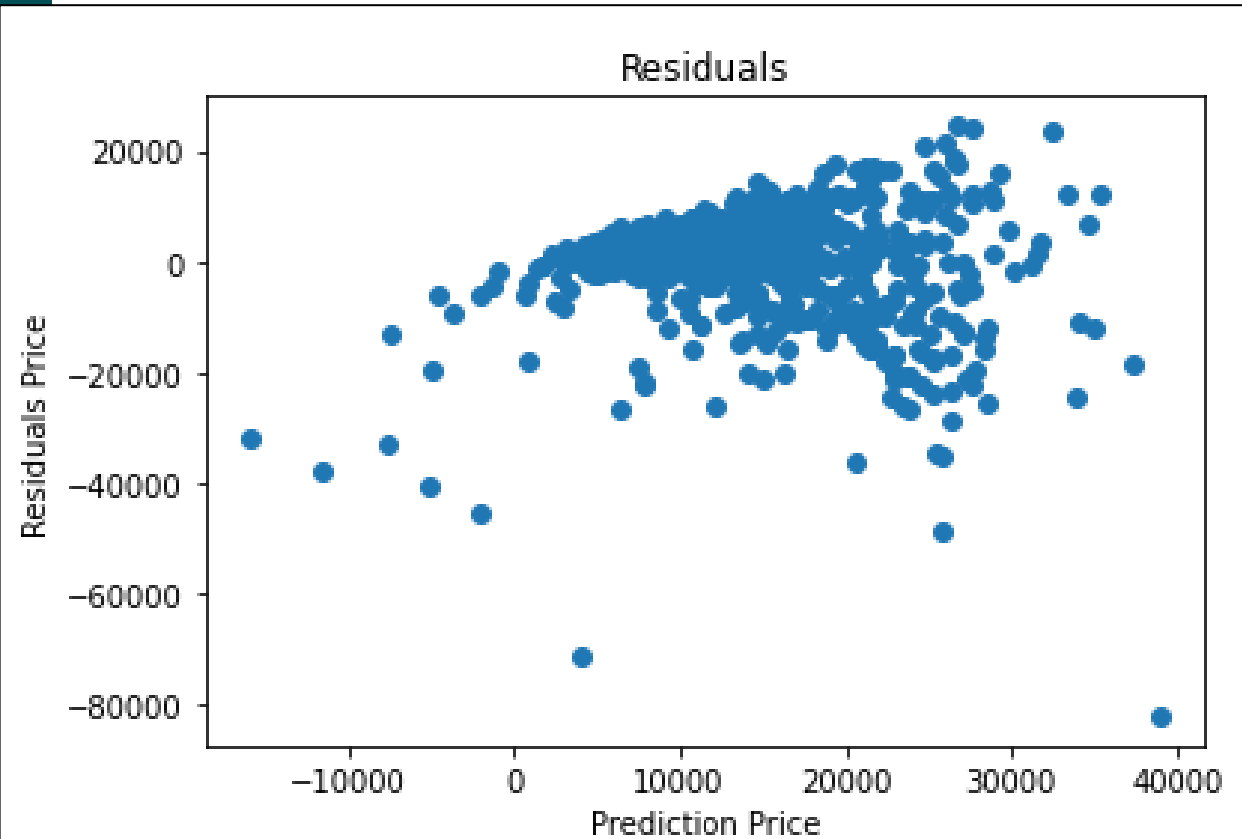
# MO CARS USED CARS PRICE ESTIMATION AND SCAM DETECTION

# WHAT IS OUR MAIN ISSUE?

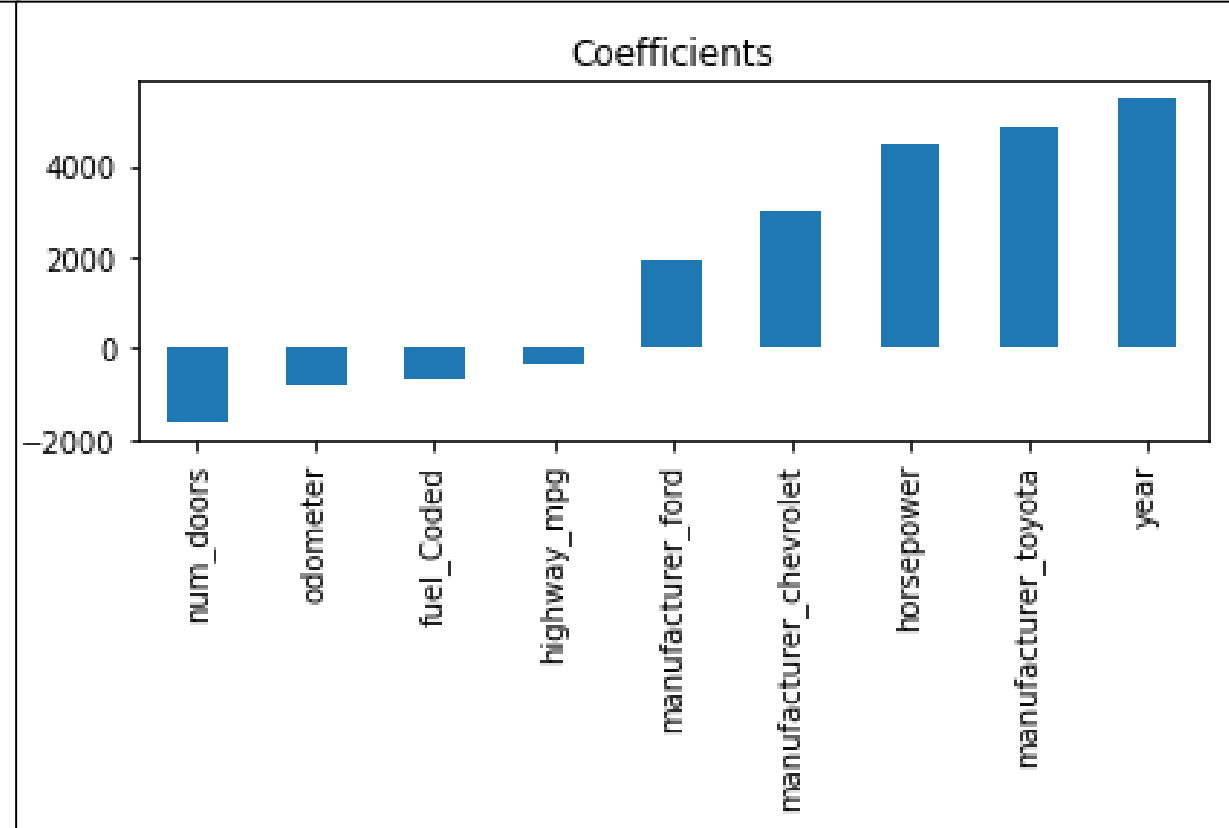
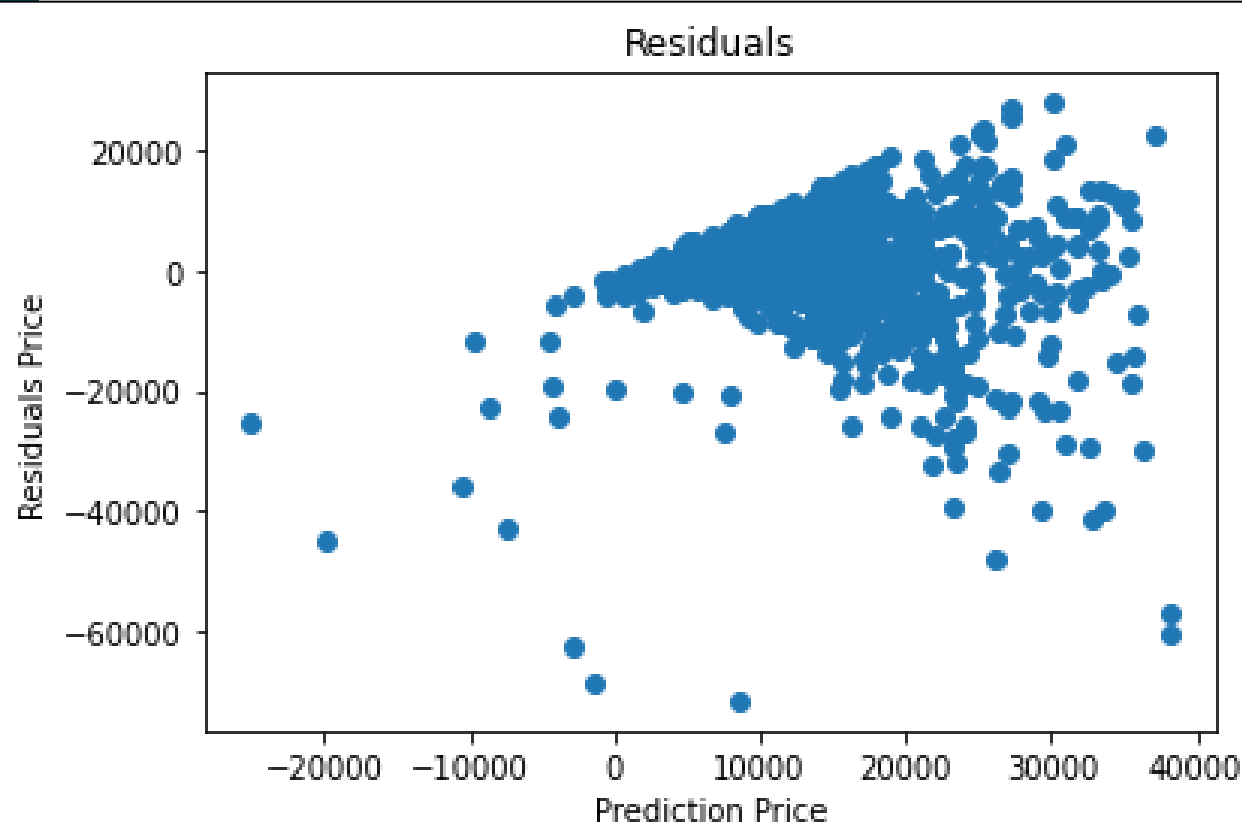
---

- WHICH CARS CURRENTLY LISTED ON CRAIGSLIST SHOULD THE COMPANY BUY?
- HOW DO WE AVOID SCAMS?

# DETECTING BEST LISTED CARS FOR PROFIT: MULTIPLE LINEAR REGRESSION MODEL



# DETECTING BEST LISTED CARS FOR PROFIT (CONT): FITTING OUR MODEL TO OUR DATA



# DETECTING BEST LISTED CARS FOR PROFIT (CONT): PREDICTIONS

## Multiple Linear Regression Formula

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

Y : Dependent variable

$\beta_0$  : Intercept

$\beta_i$  : Slope for  $X_i$

X = Independent variable

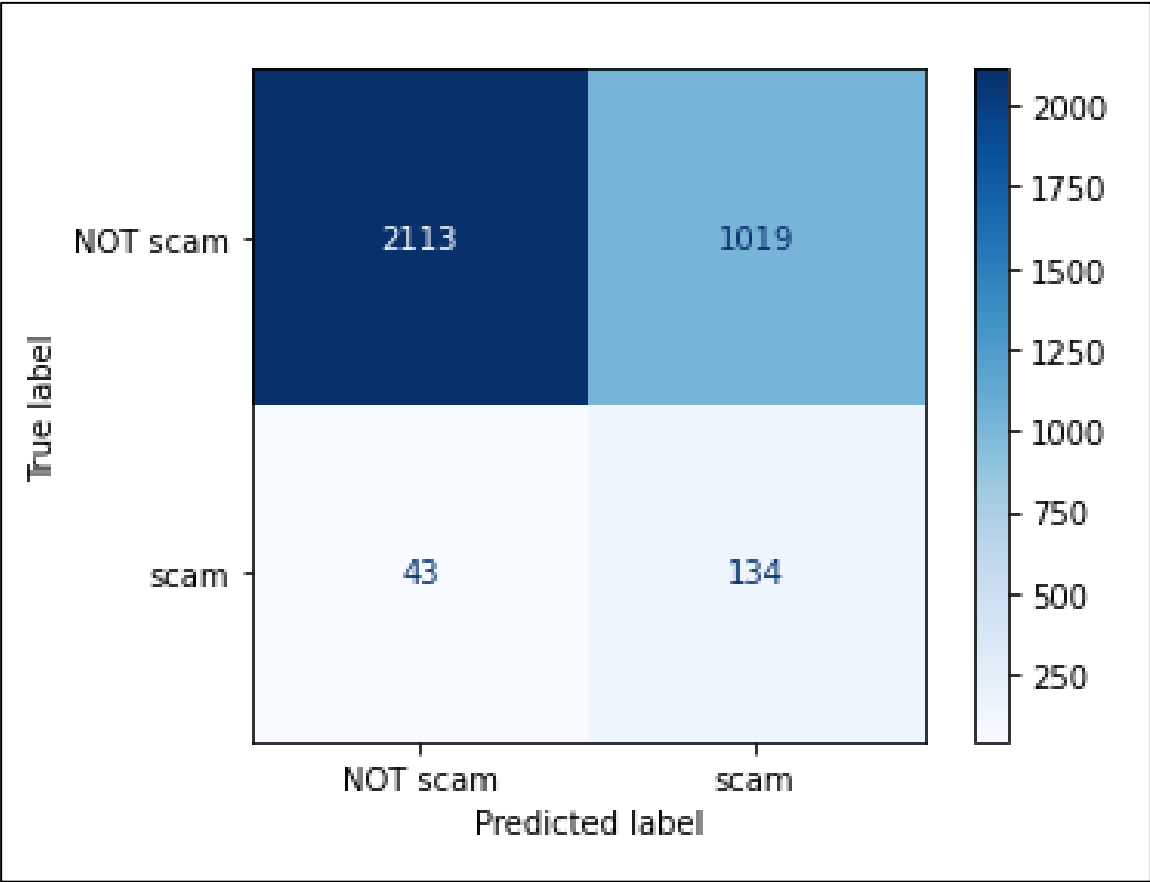
```
Model coefficient and intercept
```

```
[ 4895.45524596 -397.90410615 -2481.85990593 -2670.99393419  
 -1165.30552793  4101.15066856  2897.62694823  3160.92854955  
  3139.20073613]  
24006.74220469204
```

# DETECTING SCAMS: LOGISTIC REGRESSION MODEL

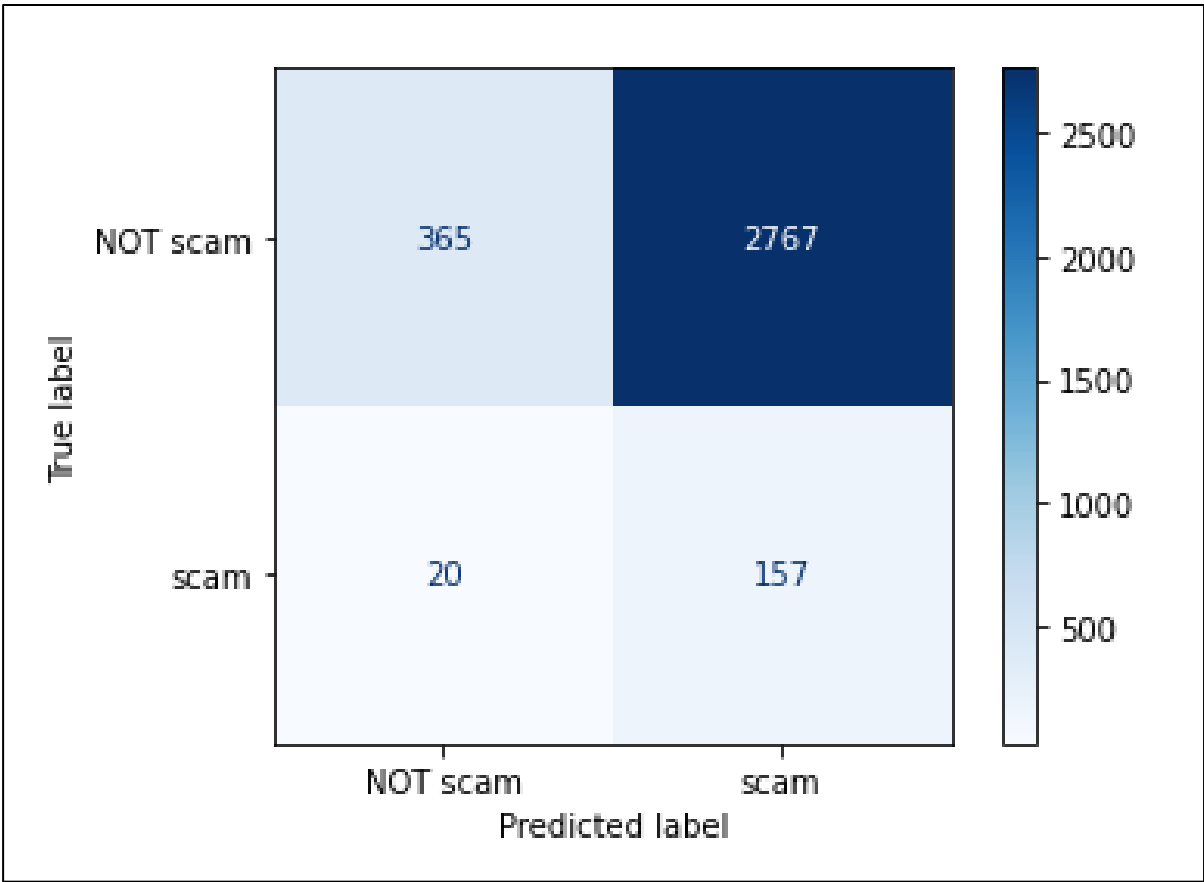
## Model 1

Number of typos, highway mpg, account age in years and condition



## Model 2

Price, year, city mpg, number of emojis and ask down payment



# DETECTING SCAMS: PERFORMANCE SUMMARIES

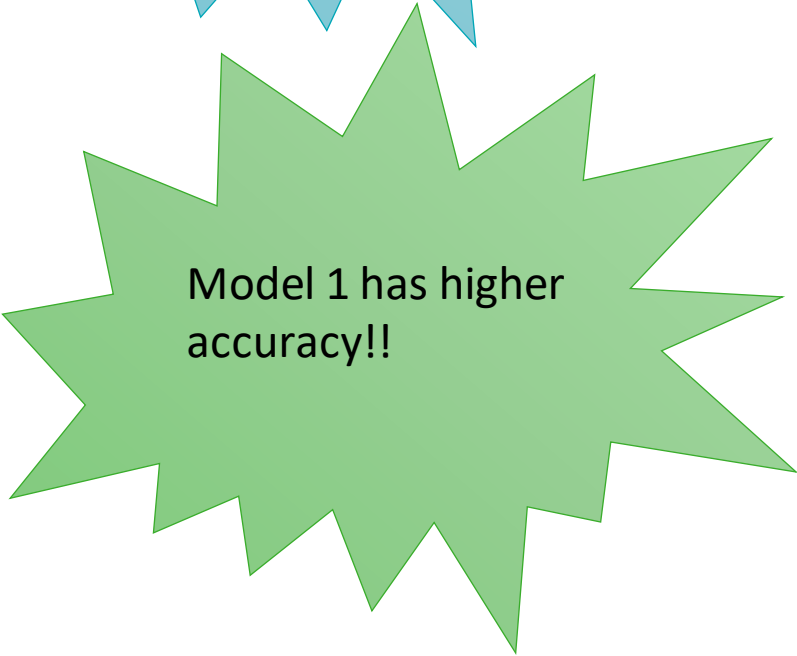
## Model 1

	precision	recall	f1-score	support
0	0.98	0.67	0.80	3132
1	0.12	0.76	0.20	177
accuracy			0.68	3309
macro avg	0.55	0.72	0.50	3309
weighted avg	0.93	0.68	0.77	3309



## Model 2

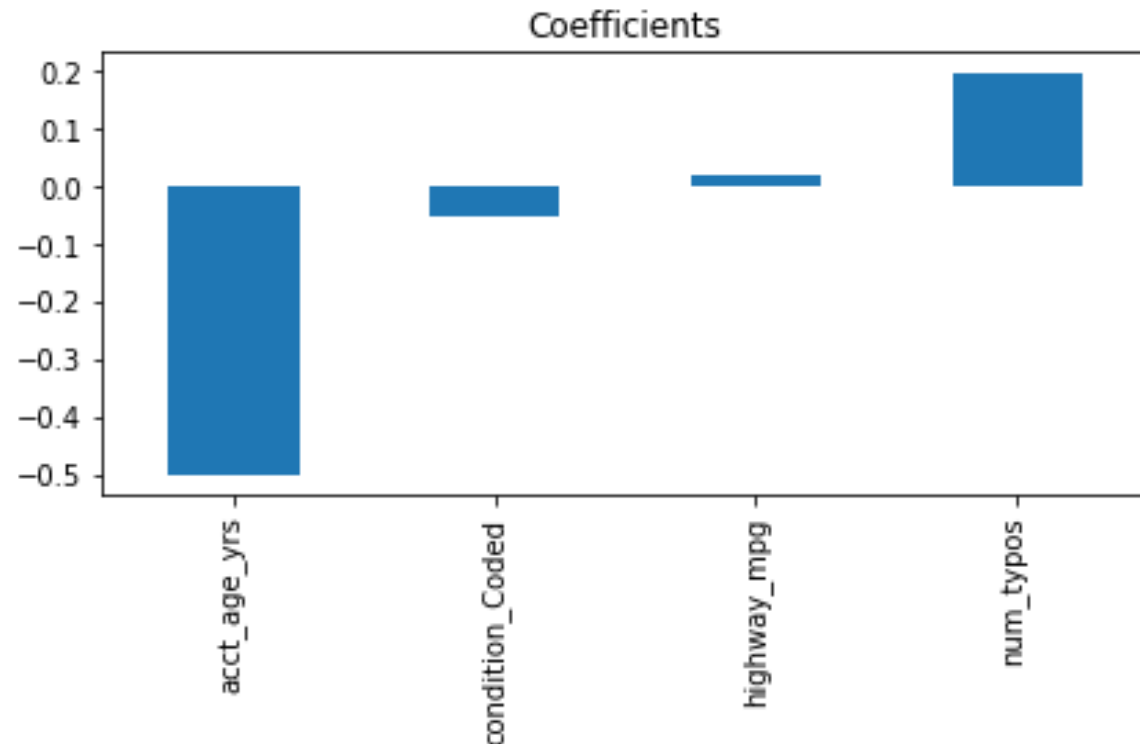
	precision	recall	f1-score	support
0	0.95	0.12	0.21	3132
1	0.05	0.89	0.10	177
accuracy			0.16	3309
macro avg	0.50	0.50	0.15	3309
weighted avg	0.90	0.16	0.20	3309



# DETECTING SCAMS: LASSO VS RIDGE REGRESSION MODEL

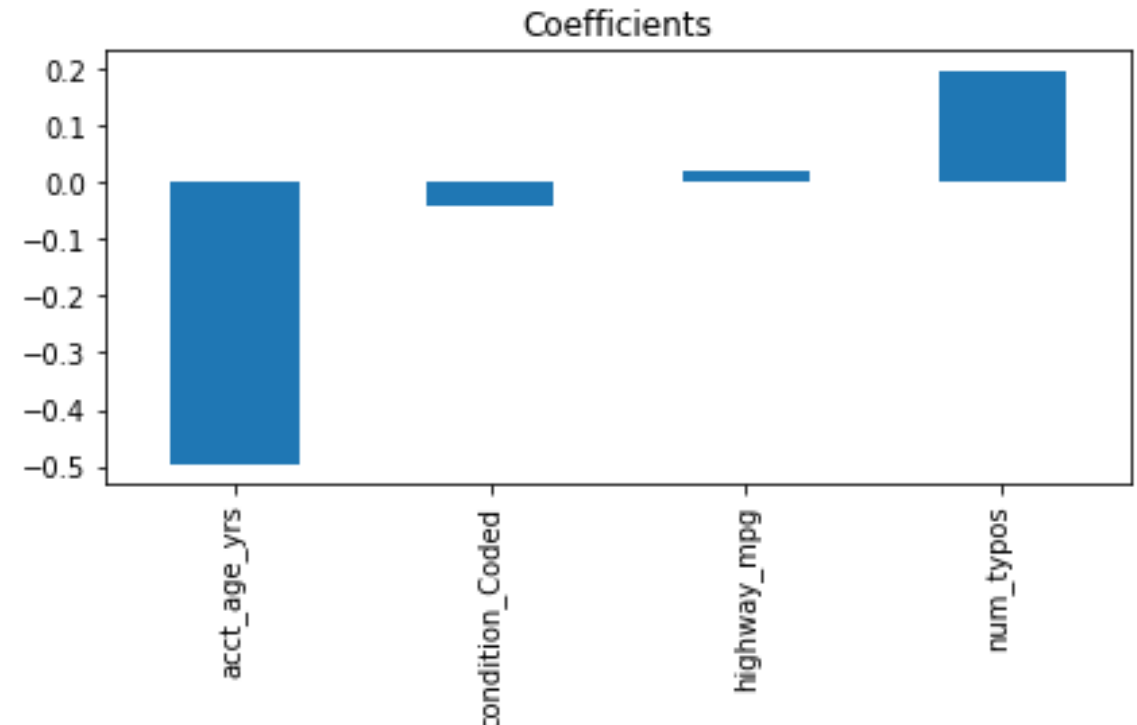
## Lasso Regression

Training set accuracy: 0.6745258066953828  
Test set accuracy: 0.6790571169537625



## Ridge Regression

Training set accuracy: 0.6762638857401949  
Test set accuracy: 0.6790571169537625





# CONCLUSION

- 
- WHICH CARS CURRENTLY LISTED ON CRAIGSLIST SHOULD THE COMPANY BUY?
  - HOW DO WE AVOID SCAMS?

# QUESTIONS?

---