Presidential Elections by County

Angie Rincon

Business Problem

- Building a model to predict presidential elections at a county level
- Initial model trained on 2016 election results
- Help campaigns predict how counties will vote in future elections
- Help campaigns determine areas for outreach

Data and Methods

- Election results, previous election results, population, and density information came from Gary Hoover and can be found at data.world.
- Alaska election results were found at thecinyc.com.
- Missing Alaska demographic information was found on Wikipedia.
- Additional information was found at the United States Census Bureau.

Target

```
df_all.Target.value_counts()
```

```
Trump       2653
Clinton      488
Name: Target, dtype: int64
```

Best Model

```
pipeline = imbpipe(steps = [
    ('sm', SMOTE()),
    ('ss', StandardScaler(with_mean = False, with_std = False)),
    ('linsvc', LinearSVC(class_weight = 'balanced'))
])
```

```
results['test_score'].mean()
```

```
0.8756135291732597
```

```
results['train_score'].mean()
```

```
0.8785950018179409
```

```
confusion_matrix(y_train, y_hat_train)
```

```
array([[1909,   80],
       [  19,  347]], dtype=int64)
```

Conclusions

- Unknown at this time

Next Steps

- Add more census information
- Additional iterations
- PCA
- Explore Logistic Regression more

Thank you!

Angie Rincon

- GitHub
- LinkedIn