



Trust in Law Enforcement:

Understanding
the Drivers of
Police
Sentiment
Scores

by Norman Lapid

Case File #ML1

Objective:

Build a model to predict the degree to which police are trusted by members of their community.

Location: Chicago, IL, USA



Case Learning #1:

Lay the groundwork. Investigation requires thorough research.



Case Learning #2:

Make the most of limited resources.



- LeaveOneOut:*
 $R^2 \rightarrow \underline{\text{Error}}$ **X X**
- Kfold($n_splits=10$):*
 $R^2 = -42\%$ **X**
- Kfold($n_splits=6$):*
 $R^2 = 16\%$
- RepeatedKFold*
($n_splits=10$, $n_repeats=10$):
 $R^2 = 62\%$ ✓
- RepeatedKFold*
($n_splits=6$, $n_repeats=10$):
 $R^2 = 71\%$ ✓ ✓

Case Learning #3:

Follow proper procedure.

1. Construct a **Pipeline**

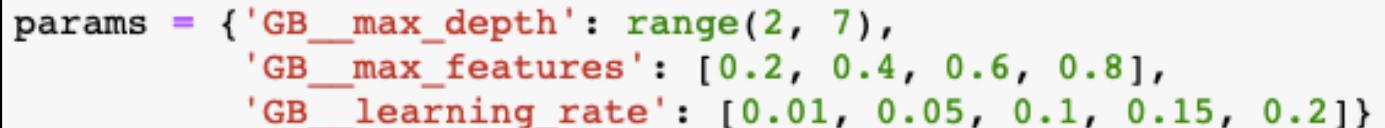
- a. Apply transforms (scaling)
- b. Apply an estimator (regression)



POLICE PROCEDURE
& INVESTIGATION
GUIDE FOR WRITERS

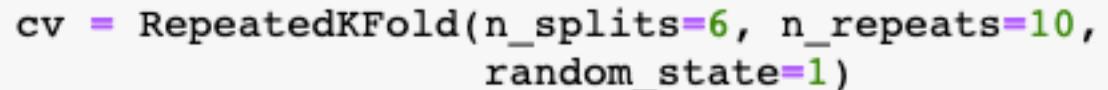
```
steps = [('scaler', RobustScaler()),  
         ('Ridge', Ridge(random_state=1))]  
pipeline = Pipeline(steps)
```

2. Define grid search parameters



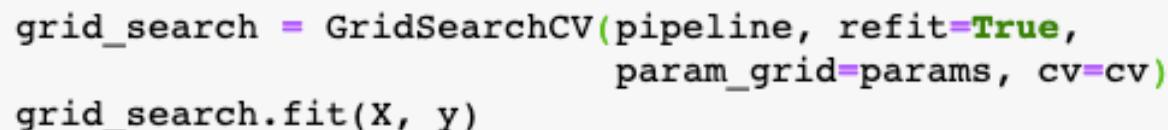
```
params = {'GB__max_depth': range(2, 7),  
          'GB__max_features': [0.2, 0.4, 0.6, 0.8],  
          'GB__learning_rate': [0.01, 0.05, 0.1, 0.15, 0.2]}
```

3. Instantiate cross-validator



```
cv = RepeatedKFold(n_splits=6, n_repeats=10,  
                    random_state=1)
```

4. Apply **GridSearchCV**

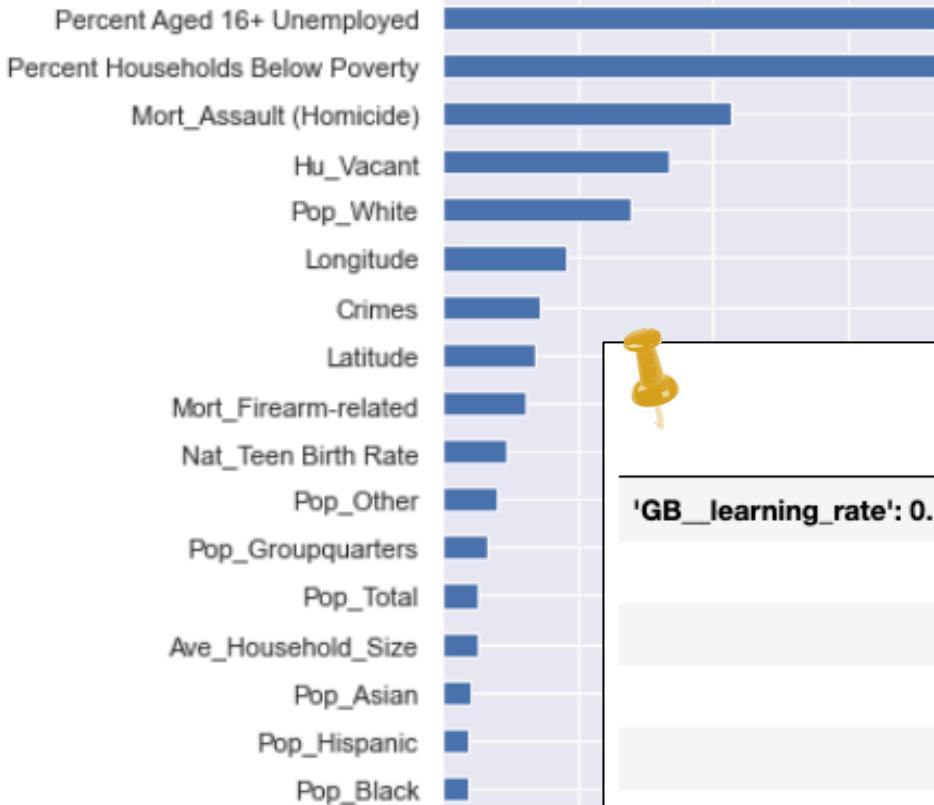


```
grid_search = GridSearchCV(pipeline, refit=True,  
                           param_grid=params, cv=cv)  
grid_search.fit(X, y)
```

Case Learning #4:

To build trust with the public,
interpretability is key.

Model: Gradient Boosting Method

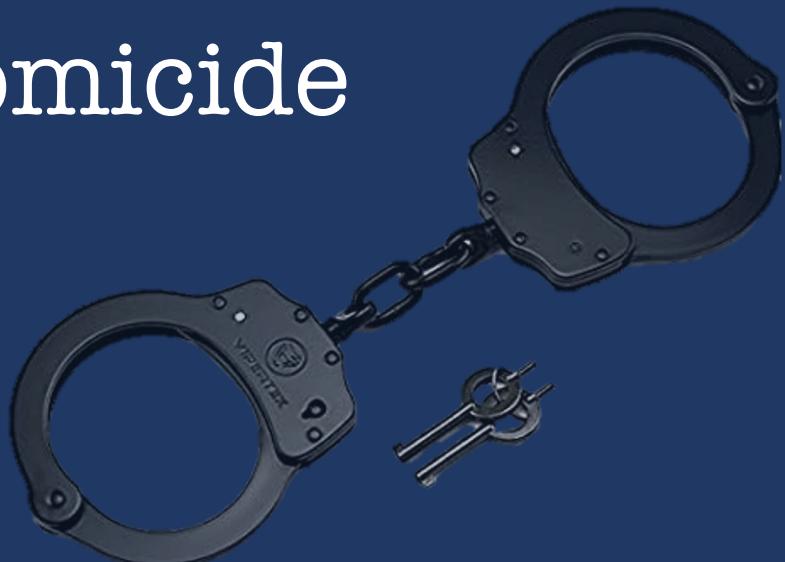


model_top_parameter	mean_test_score	top_predictor
'GB__learning_rate': 0.1, 'GB__max_depth': 4, 'GB__max_features': 0.4	0.706773	Percent Aged 16+ Unemployed
'RF__max_depth': 7, 'RF__max_features': 'sqrt'	0.703165	Pop_White
'Ridge__alpha': 10	0.698531	Longitude
'Lasso__alpha': 0.1	0.688623	Longitude
'kNN__n_neighbors': 7	0.629501	X N/A
'DT__max_depth': 3	0.523752	Percent Aged 16+ Unemployed



Top Suspects:

1. Unemployment
2. Race (% of white population)
3. Urban decay/Vacant housing
4. Per Capita Deaths from Homicide





Thank you.