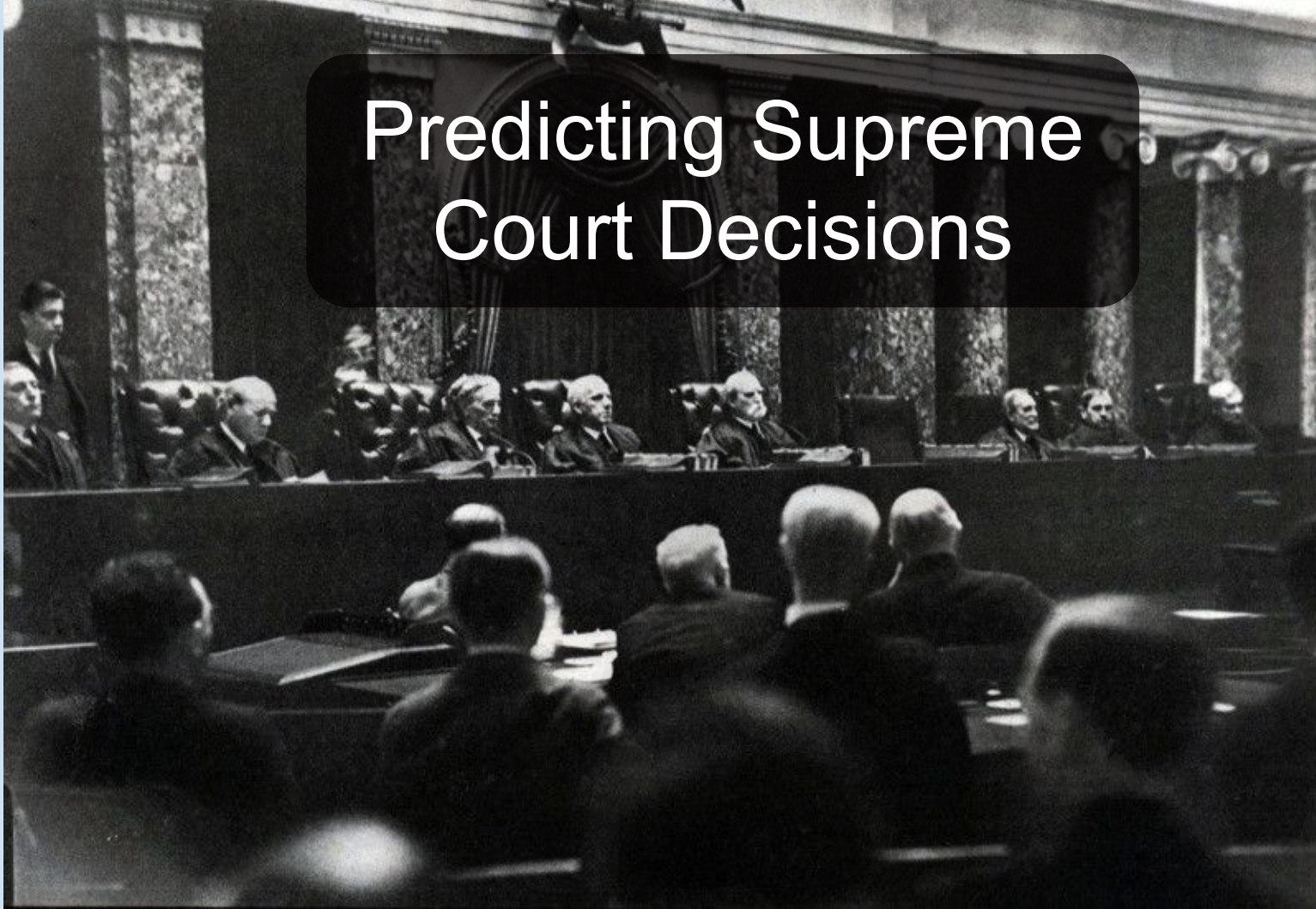


Predicting Supreme Court Decisions



The United States Supreme Court

- ❖ Authority from Article III of the Constitution and Marbury vs. Madison
- ❖ Primarily Court of Appeals
 - Questions of law
- ❖ The Power of Five
 - Five votes interprets (writes?) the law
 - Five votes may determine that a law is unconstitutional and can only be enacted by constitutional amendment.

Significant 5-4 Decisions

- ❖ Obergefell vs. Hodges, 2015
- ❖ National Federation of Independent Business v. Sebelius, 2012
- ❖ Citizens United vs Federal Election Commission, 2010
- ❖ Bush v Gore, 2000
- ❖ Regents of the U. of California v. Bakke, 1978
- ❖ Miranda v Arizona, 1966

The Supreme Court Database (Wash U. Law)

- ❖ Legacy Data (1791-1945) by Case and Vote: 19601 Cases
- ❖ Modern Data (1946-2016) by Case and Vote: 8809 Cases
- ❖ Case Level Sets include 53 Features
 - Many are redundant
 - Many are correlated
 - Many Null Values

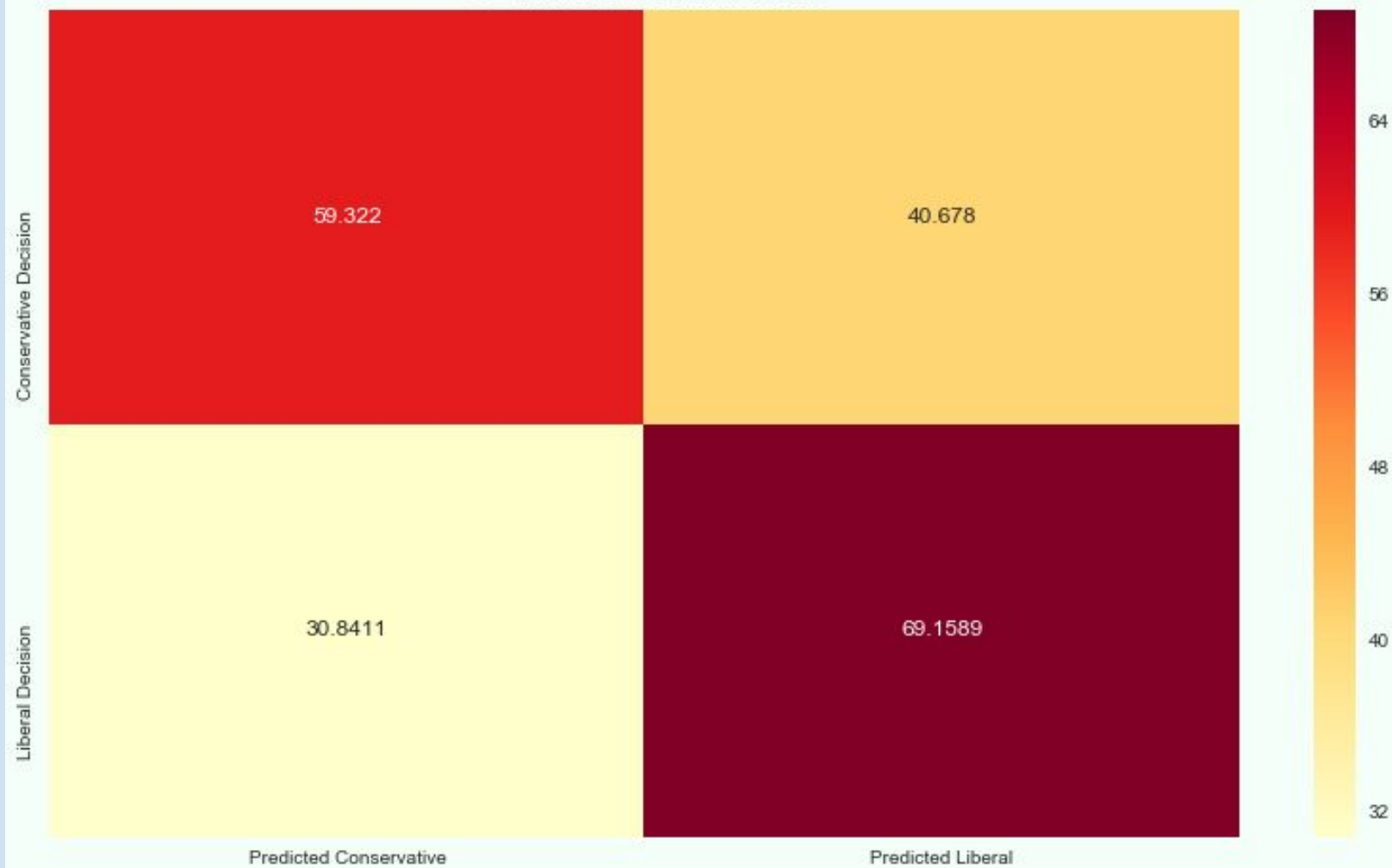
Initial EDA

- Eliminate duplicative columns
- Eliminate columns with large numbers of null values
- Eliminate rows for other null values
- Choose Target as decisionDirection
- Eliminate cases with unspecified decisionDirection

Modelling/Return to EDA

- Choose Random Forest based on preliminary comparisons with Legacy Data
- Choose Case Level Data
- Feature Engineering
 - Coding features with too many categories
 - Create Dummies for Justices
 - Eliminate Additional Features based on importance
- Separate Training and Testing Data

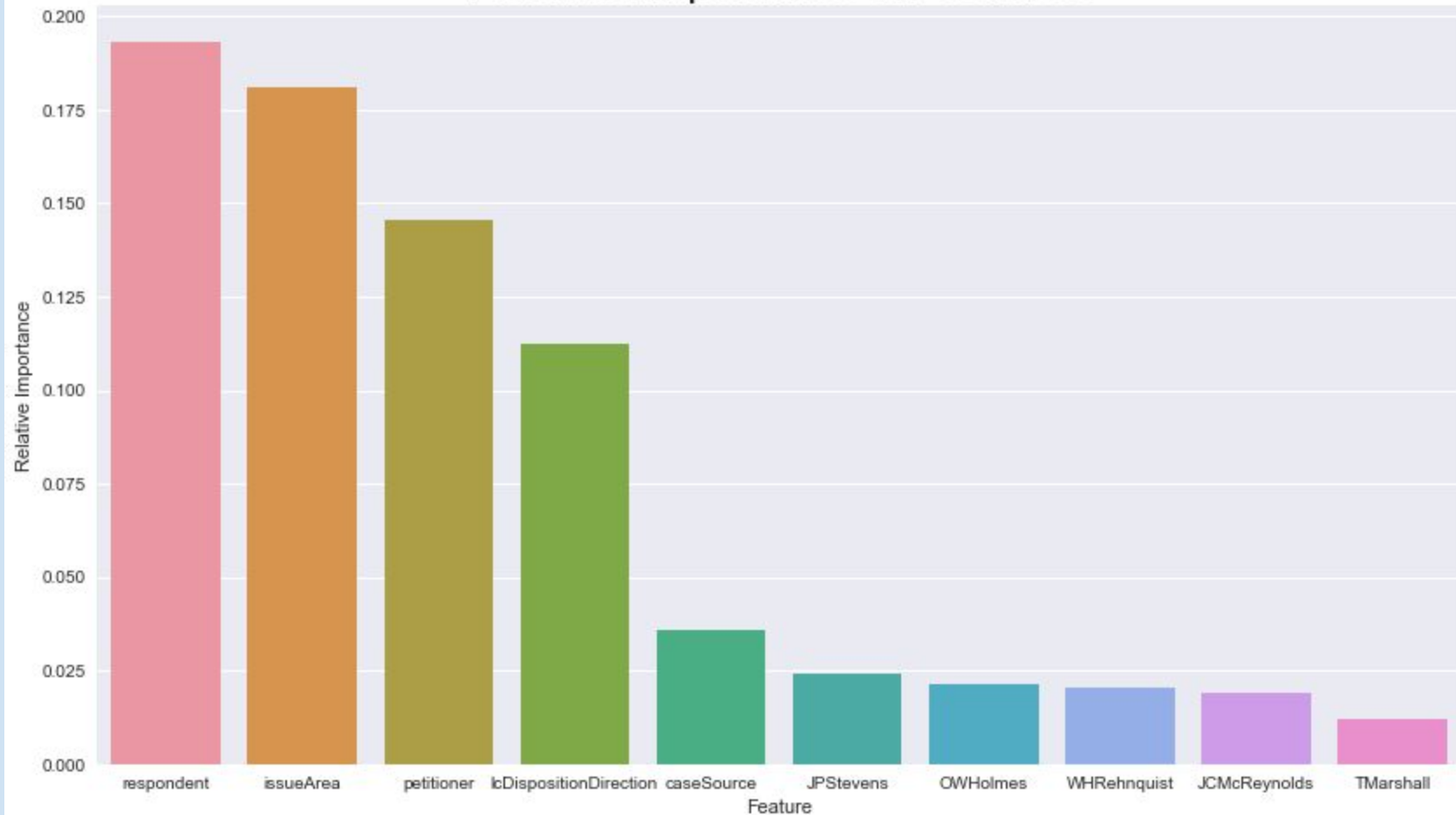
Confusion Matrix



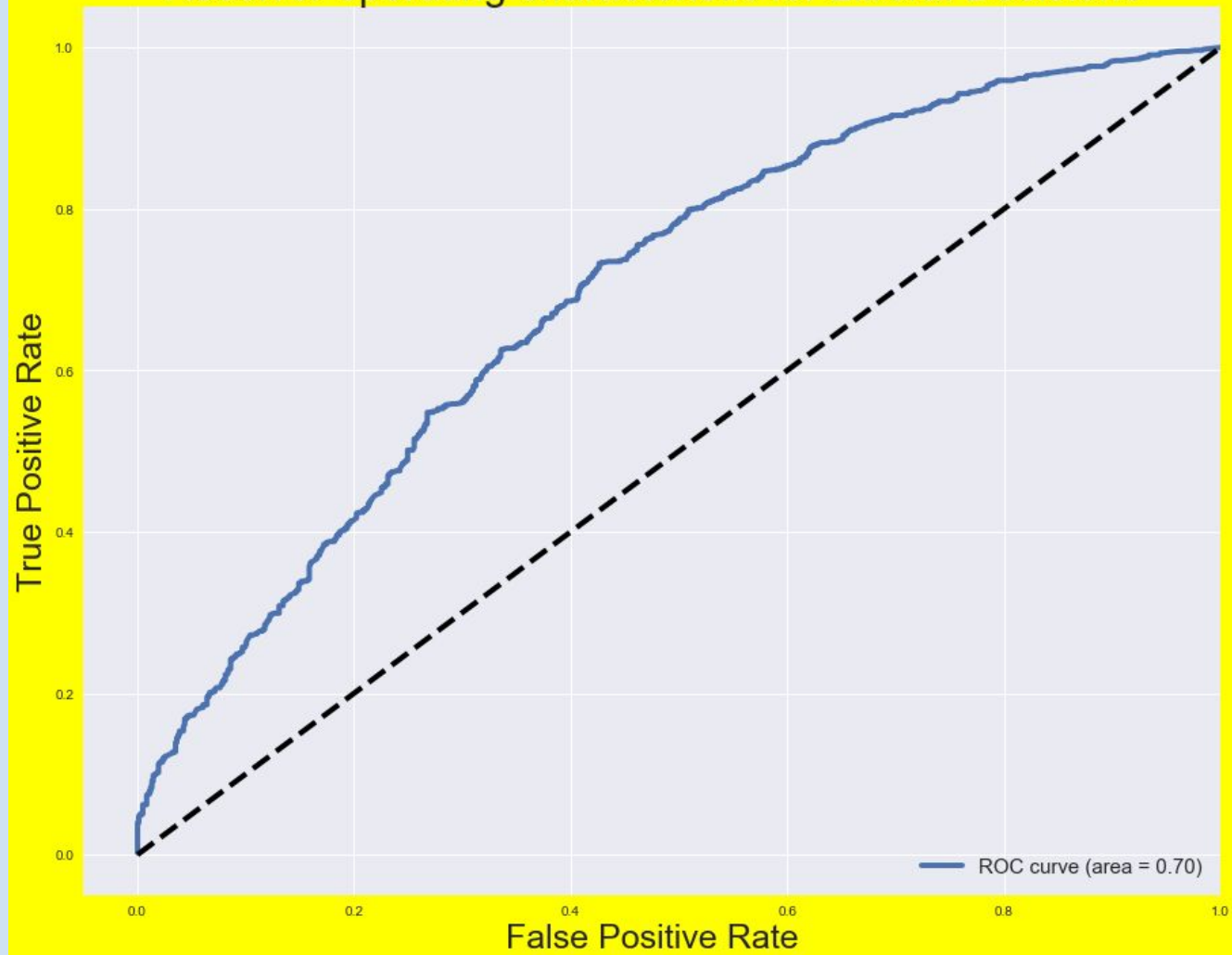
Classification Report

	Precision	Recall	F-1	Support
Conservative	.65	.59	.62	826
Liberal	.64	.69	.66	856
Avg/Total	.64	.64	.64	1568

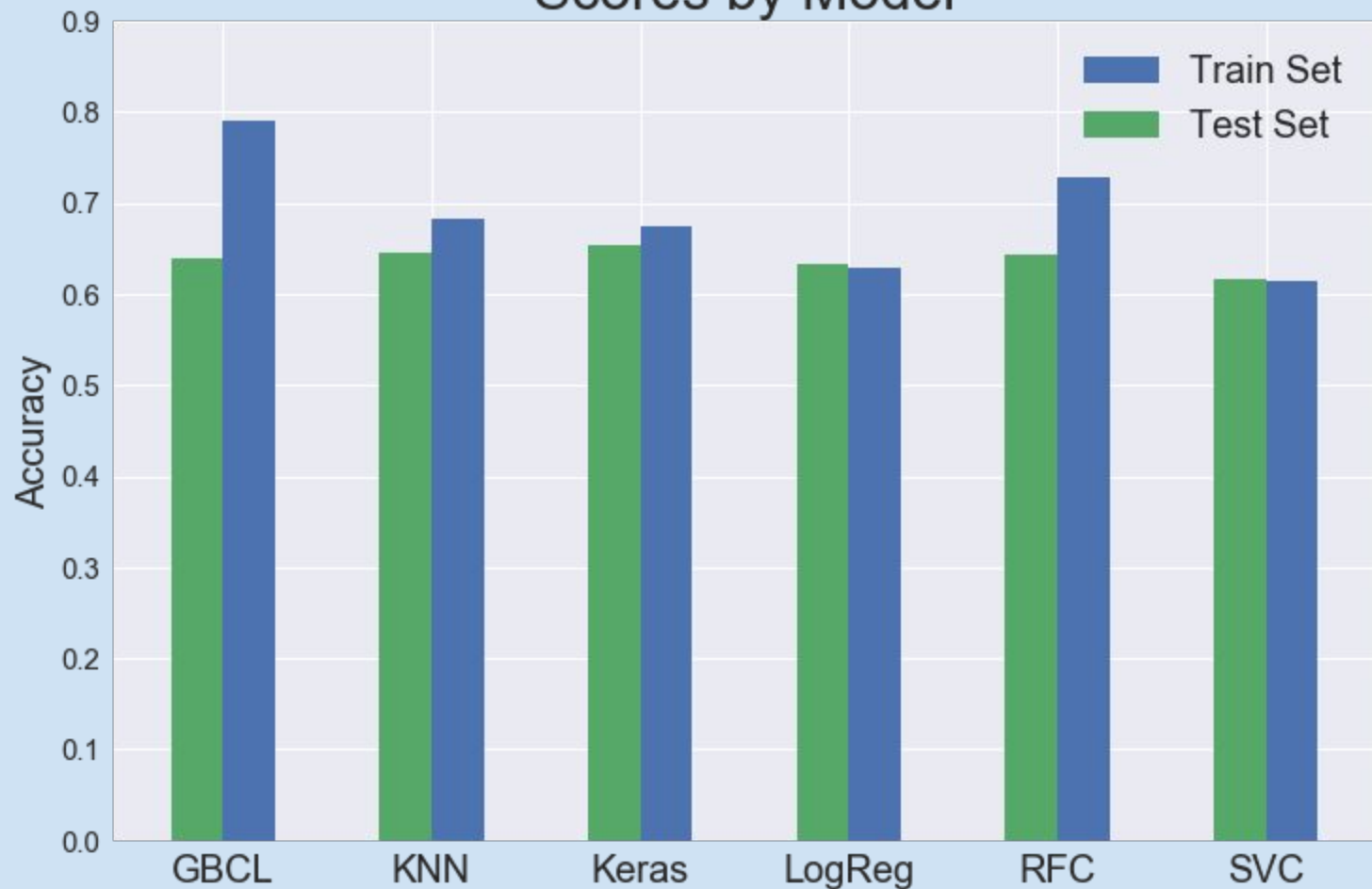
Relative Importance of Features



Receiver operating characteristic for Liberal Decisions



Scores by Model



Next Steps

- Refine Model
 - Time-Series
 - Justice Level Data
 - Review feature selection
- Another Look at Boosted Models
- Assess Coding of Features
- Apply to 2017 and Beyond

