Using Oral Arguments to Predict Supreme Court Case Outcomes Lesson

Jaime DyBuncio, June 2020
github.com/jdybuncio
Galvanize g119 cohort

The U.S. Supreme Court in 30 Seconds | Timeline of a Supreme Court Case

Granted

4 Votes needed

≈160 Days
Written Briefing

Argued

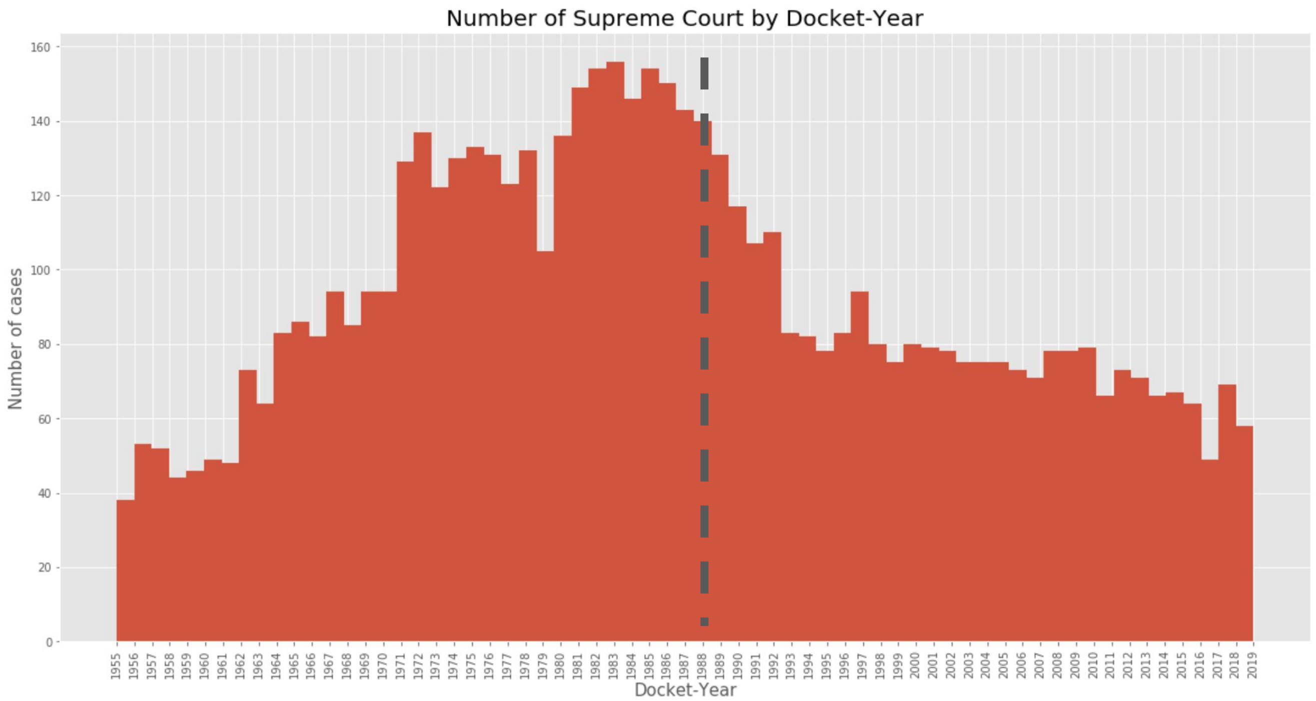
Oral Arguments
(timed session)

≈90 Days
Weekly Conference

Decided

Opinions Read

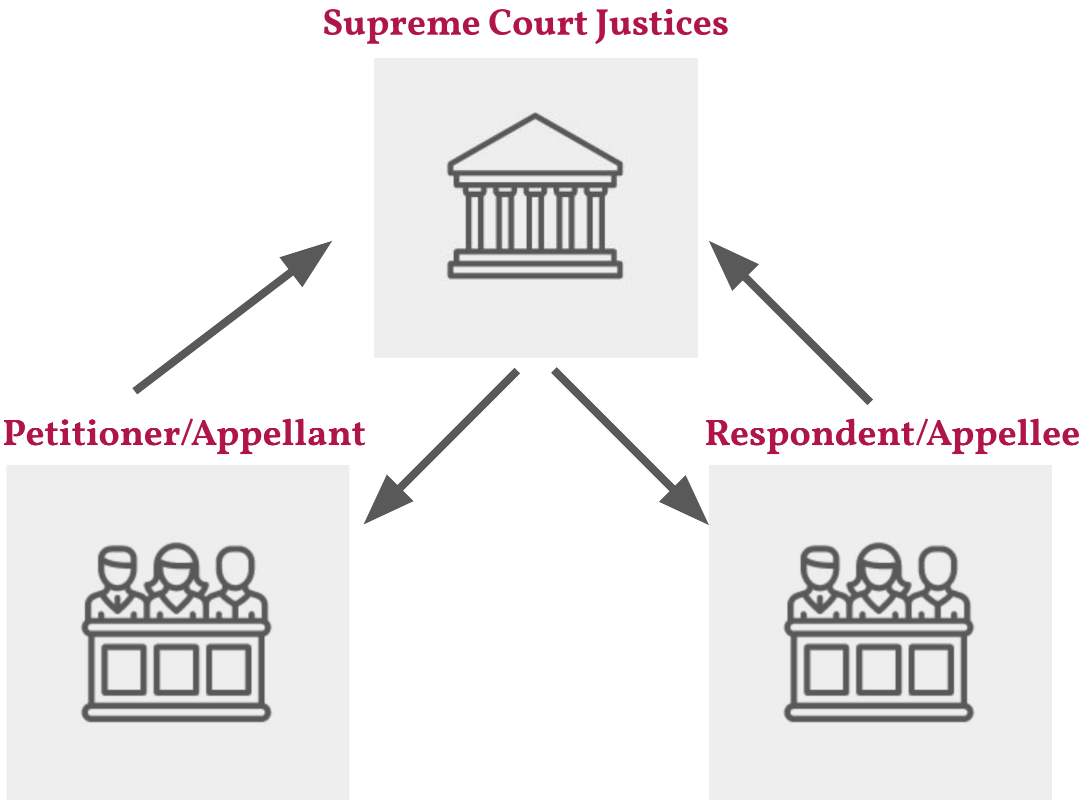
The U.S. Supreme Court in 30 Seconds | Annual Caseload



Sample

- 6,000 Supreme Court Cases from 1955-Current w/ One Oral Argument
    - Most recently, Supreme Court takes **60-70 cases per Year**

The U.S. Supreme Court in 30 Seconds | Parties Involved



Data Science & The Supreme Court

Goal

Predict if the Petitioner of the Case will win

Application

1. Influence Petitioner Strategy during oral arguments

2. Craft strategy during waiting period b/w Argument & Decision (i.e. Op-eds)

Creating a Dataset from Oral Argument PDFs

PDF of Oral Argument

Oyez.org

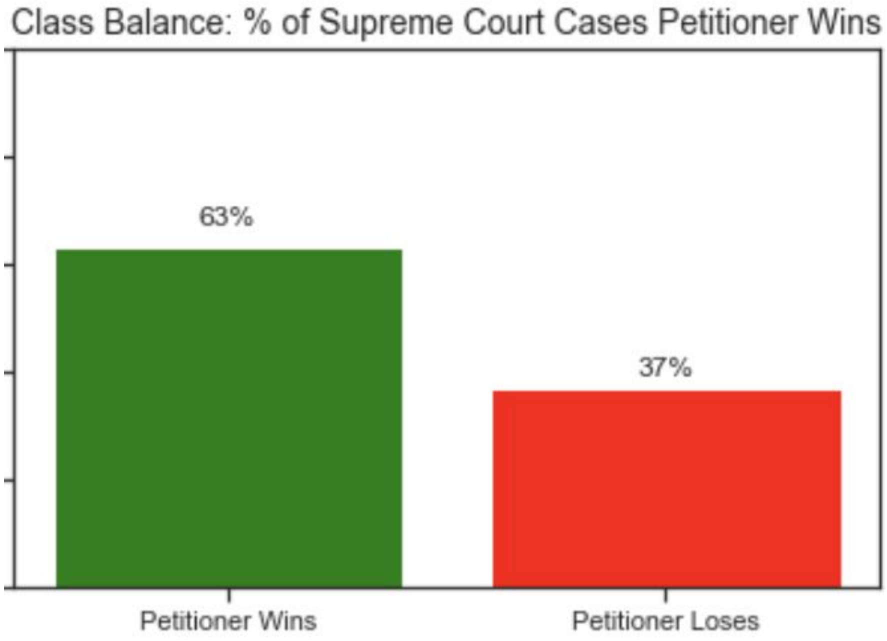
- **Supplement w/ Case Data**
  - Add Label if Petitioner Wins
  - Identifies 10 types of speaker directions
    (Ex: Respondent to Justice. Justice to Respondent. Amicus on behalf of Petitioner to Justice, etc.)

- **50 Features Created** (10 speaking Directions * 5 Features per Direction)

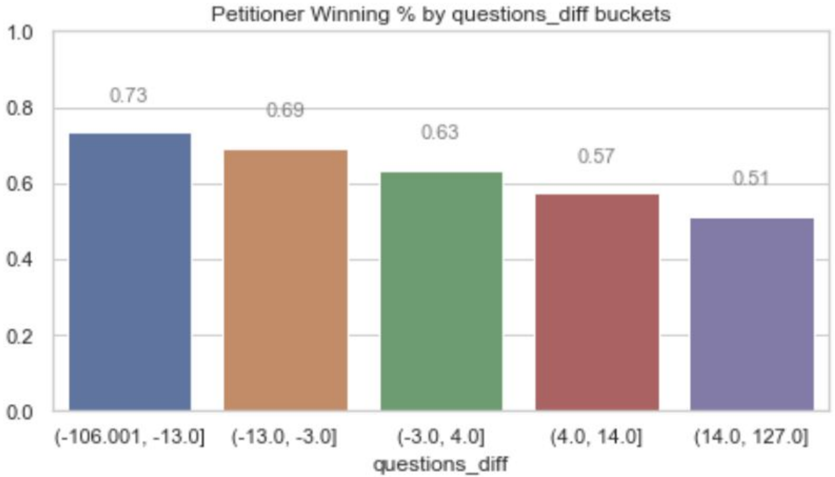
My Parsing Script & Functions

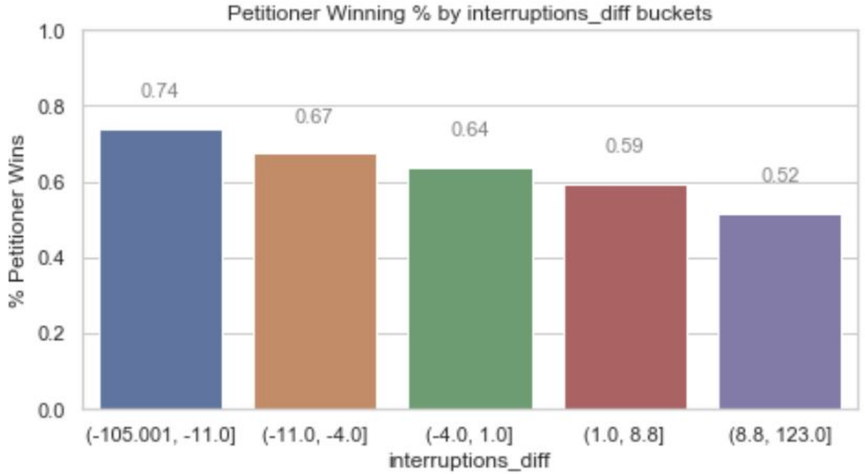
Party	Petitioner-to-Justice	Justice-to-Petitioner
(1) Words	61	3
(2) Time	13.9 secs	1.23 secs
(3) Interruptions	1	1
(4) Questions	0	0
(5) Corpus	"I think it is very hard to imagine.."	"Well, that's another"
(6) Petitioner Wins	1	0

EDA | Class Balance



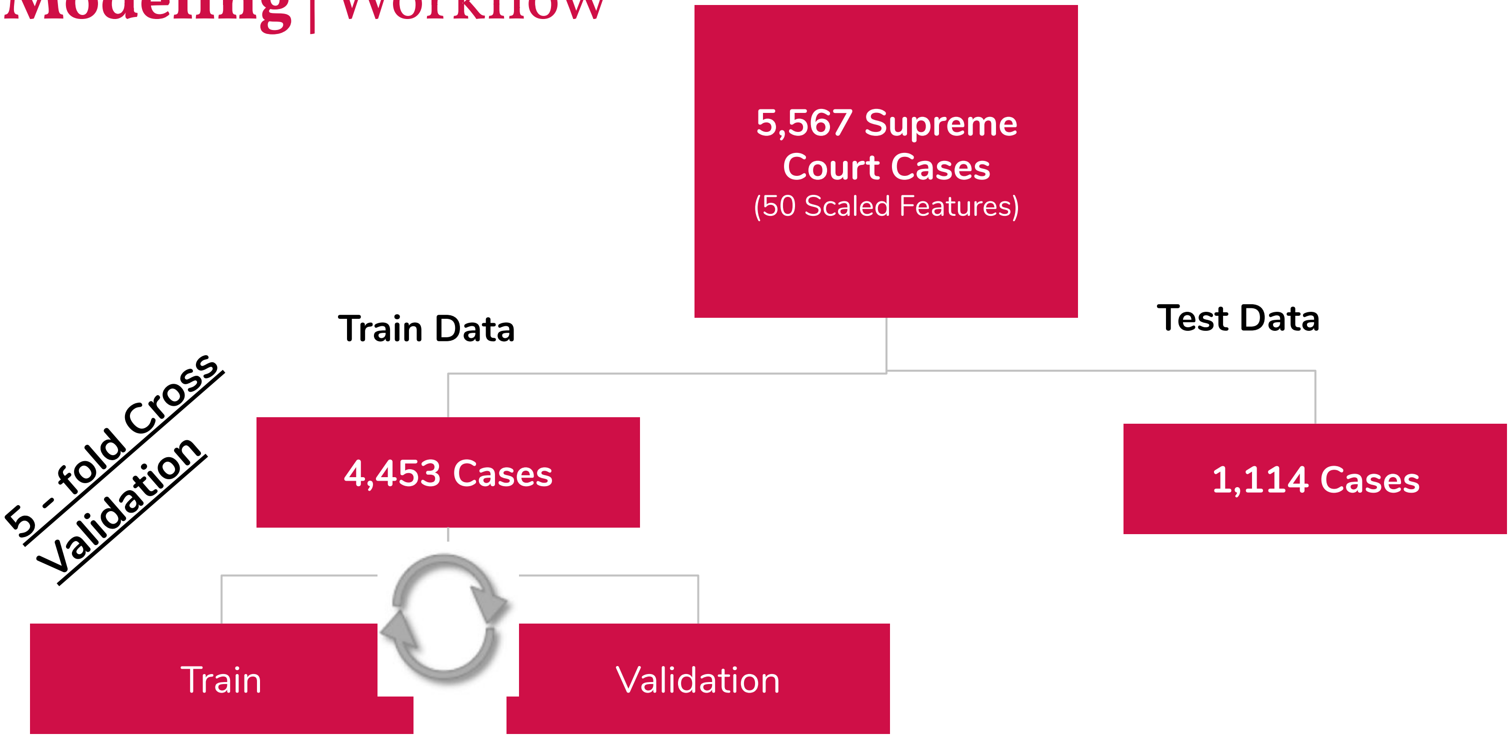
EDA | Interesting Features





Note: Petitioner Win Rate is always above 50%

Modeling | Workflow



- Supervised Learning Algorithms Trained & Evaluated:
  - Logistic Regression
  - Random Forest
  - Gradient Boost
  - LSTM
- Tested varying Features used
- Gridsearch parameters to test models

- Using model w/ highest avg F1-Score on Validation Set:
  - Predicted outcome of Test Data and compared to actuals to evaluate final model performance

Performance vs. Validation Set(s)

Modeling | Numerical

Random Forest Parameters used in Grid Search

Parameter	Grid Search Values
n_estimators	[100,200,500,1000,1200]
max_features	['sqrt', 0.25,.50,None]
min_samples_split	[3, 5, 10, 50]
min_samples_leaf	[5, 10, 50]
max_depth	[3, 5, 10, None]

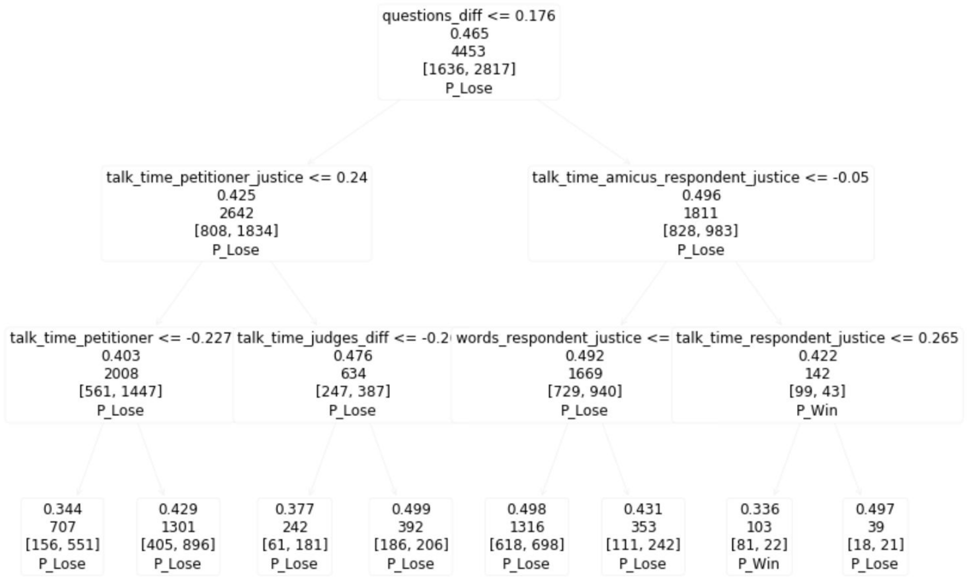
(best model's params are underlined)

	F1	Recall	Precision
Baseline_Train	0.77	1.00	0.63
Model_Train	0.78	0.98	0.65

- **GridSearched** over 1,400 models
- **Tested:** Logistic Regression, Random Forest, Gradient Boost Classifiers
- **Top 3 Features:** Petitioner - Respondent Difference in: *Judge Talk Time*, *Questions*, & *Interruptions*

Modeling | Numerical w/ NLP Features

Example of one Decision Tree in Random Forest Model



	F1	Recall	Precision
Baseline_Train	0.77	1.00	0.63
Model_Train	0.79	0.98	0.65

- **GridSearch** 15 models w/ NLP Features
- **NLP Features**: Tf-idf Vectors of the: Petitioner Words to the Judges, Judges words to the Petitioner, and both combined
  - Tested: 500, 5,000, and 10,000 max features in Tf-idf vector
  - Removed Stop Words & Punctuation
- **Top 3 Features:** Petitioner - Respondent Difference in *Judge Talk Time*, and *Interruptions*. and *Talk Time of the Justices to the Respondent*

Modeling | LSTM

Top 20 Words appearing across all Oral Arguments

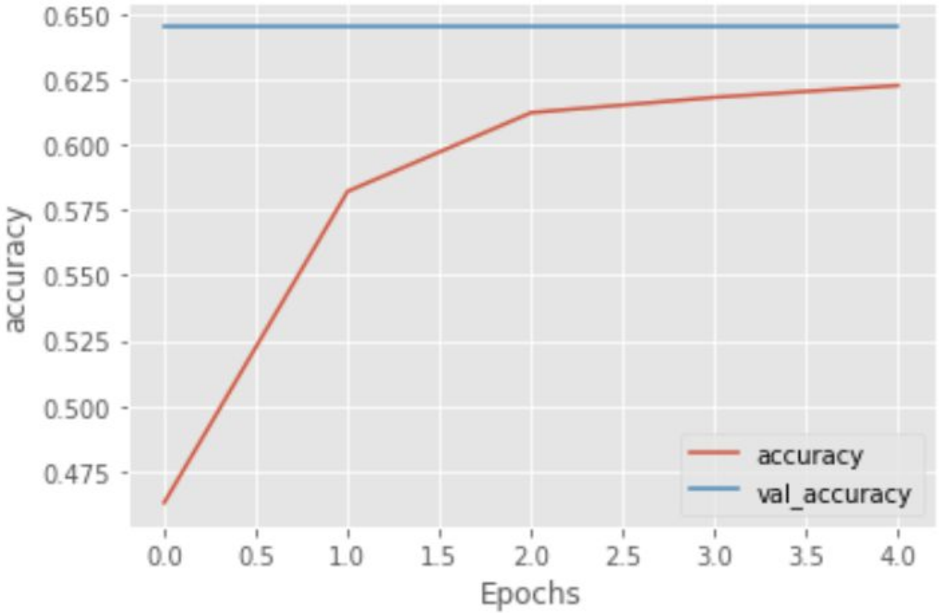
```
{'<OOV>': 1,
 'court': 2,
 'would': 3,
 'case': 4,
 'think': 5,
 'well': 6,
 'state': 7,
 'one': 8,
 'mr': 9,
 'say': 10,
 'thats': 11,
 'yes': 12,
 'question': 13,
 'right': 14,
 'could': 15,
 'honor': 16,
 'statute': 17,
 'dont': 18,
 'justice': 19,
 'may': 20}
```

```
'number 21 docket united states america versus twin city power company et al
mr spritzer mr chief justice honors case involves issue compensation arising
eminent domain proceedings instituted united states 1947 western district so
uth carolina taking made necessary building dam savannah river forms border
south carolina georgia location known clarks  clark hill clark hill dam vita
l part comprehensive plan pictured map reproduced engineers engineers studie
s clark hill dam central part savannah river basin project located point rig
ht tip brown area represent reservoir basin project declared purposes flood
control navigation improvement hydroelectric development condemned lands who
se value question proceeding located miles upstream governments dam clark hi
ll flooded result backing river dam otherwise stated part reservoir basin ta
king fast lands adjacent river respondent course entitled compensation respo
ndent however awarded believe special increment value entitled whether right
view  issu'
```

•••

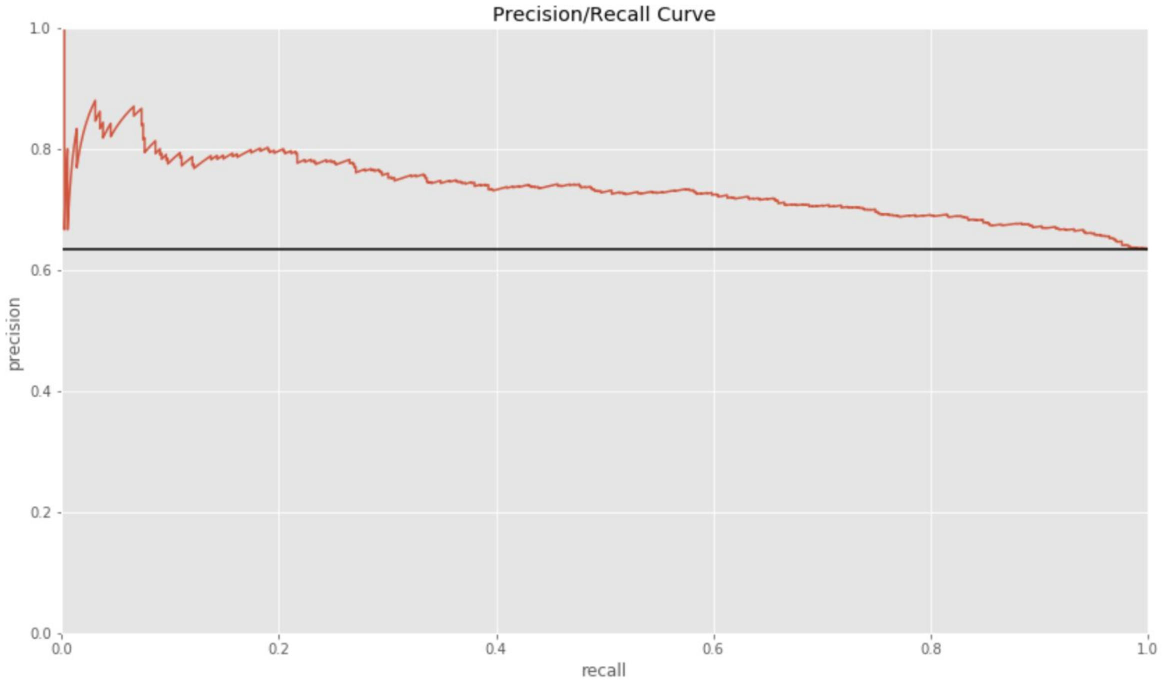
```
[153, 2133, 4292, 86, 30, 3221, 234, 7696, 274, 135, 200, 2285, 3874, 9, 807
4, 9, 124, 19, 719, 4, 1065, 38, 611, 2219, 5035, 3461, 443, 3649, 86, 30, 3
047, 2151, 37, 1089, 928, 390, 36, 299, 1235, 3277, 9570, 896, 2079, 2660, 1
089, 928, 875, 2405, 825, 1, 3009, 3425, 3009, 3425, 3277, 3668, 95, 3049, 2
91, 1, 2672, 5526, 3166, 3166, 3151, 3009, 3425, 3277, 1269, 95, 9570, 896,
6377, 1718, 2010, 31, 14, 5112, 1593, 240, 1196, 6292, 6377, 1718, 2088, 27
6, 3263, 415, 3262, 4974, 8927, 1612, 4806, 872, 1244, 484, 13, 272, 2010, 1
505]
```

Modeling | LSTM



Performance vs. Unseen Test Set

Modeling | Final Results vs. Test Data



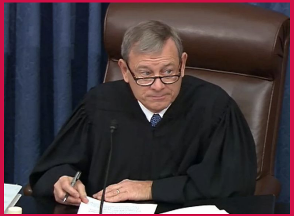
	F1	Recall	Precision
Baseline_Train	0.77	1.00	0.63
Model_Train	0.79	0.98	0.65
Baseline_Test	0.78	1.00	0.63
Model_Test	0.77	0.97	0.64

Predicting 1,114 Supreme Court Cases

Threshold	TP	FP	TN	FN	Recall	Precision	F1
Best Model (Threshold = 0.51)	680	357	52	25	0.96	0.66	0.78
Baseline (Petitioner Wins 100%)	705	409	0	0	1.00	0.63	0.78

“The secret to successful advocacy, is simply to get the court to ask your opponent more questions.”

Chief Justice Roberts



Questions?

Thanks!

Jaime DyBuncio, June 2020
github.com/jdybuncio
Galvanize g119 cohort

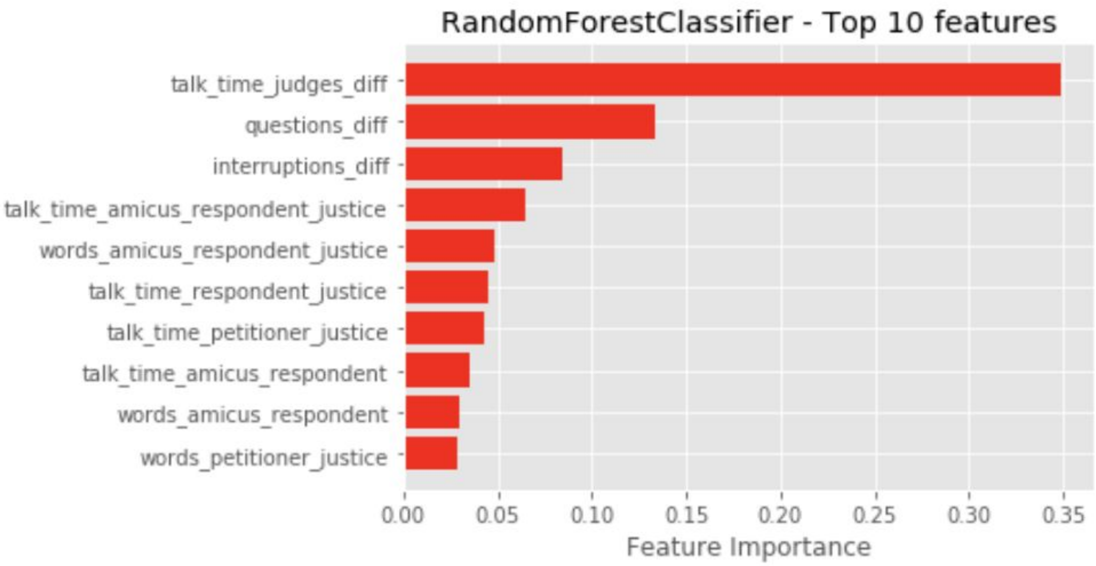
CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Please keep this slide for attribution.

APPENDIX

```
{'corpus_petitioner': (84516, 1912.0, 676.0),
 'corpus_petitioner_justice': (51537, 640.0, 301.0),
 'corpus_respondent': (80443, 1705.0, 622.0),
 'corpus_respondent_justice': (46442, 648.0, 300.0),
 'corpus_amicus_neutral': (14209, 47.0, 22.0),
 'corpus_amicus_neutral_justice': (8154, 17.0, 10.0),
 'corpus_amicus_petitioner': (13852, 48.0, 23.0),
 'corpus_amicus_petitioner_justice': (9457, 22.0, 13.0),
 'corpus_amicus_respondent': (11560, 32.0, 16.0),
 'corpus_amicus_respondent_justice': (7431, 14.0, 9.0)}
```

APPENDIX



1. **Tested: Logistic, RF, GBoost models - GridSearch over 1,400 models**
2. **Criterion** = Gini
3. **Trees:** 1,000
4. **Max Features** = 0.25
5. **Max Depth** = 3
6. **Min Samples Split** = 3
7. **Min Samples Leaf** = 10
8. **Features:** Difference in Judge Talk Time, Difference in Questions, Difference in Interruptions

APPENDIX

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

True Positive + False Positive = Total Predicted Positive

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

True Positive + False Negative = Actual Positive

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$= \frac{True\ Positive}{Total\ Predicted\ Positive}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$= \frac{True\ Positive}{Total\ Actual\ Positive}$$

Precision - % of Predicted Petitioner Wins actually are cases they Win

Recall - % of Actual Petitioner Wins model is able to Project