The Genealogy of Ideology: Identifying Persuasive Memes and Predicting Agreement in the U.S. Courts of Appeals

Daniel L. Chen, Adithya Parthasarathy, Shivam Verma

NYU

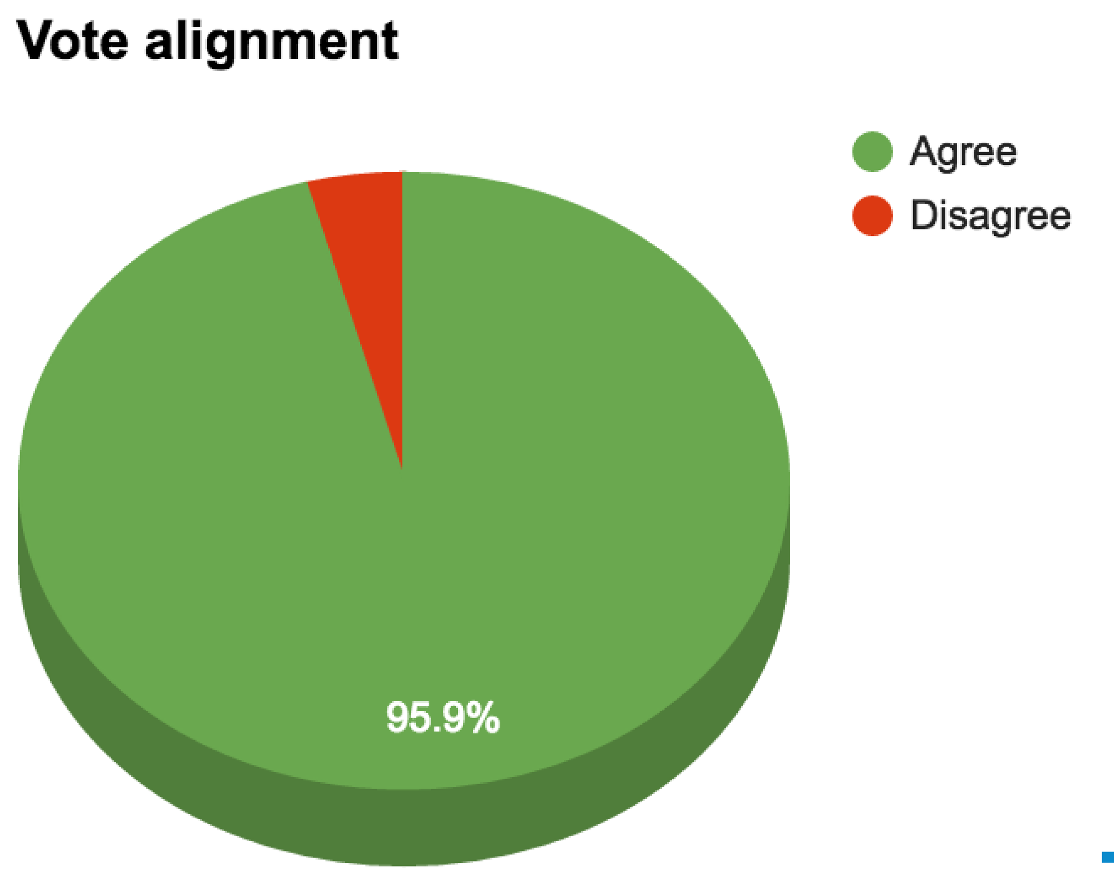
Goal

Analyze the voting patterns of the judges and predict if two judges are likely to agree or disagree with each other.

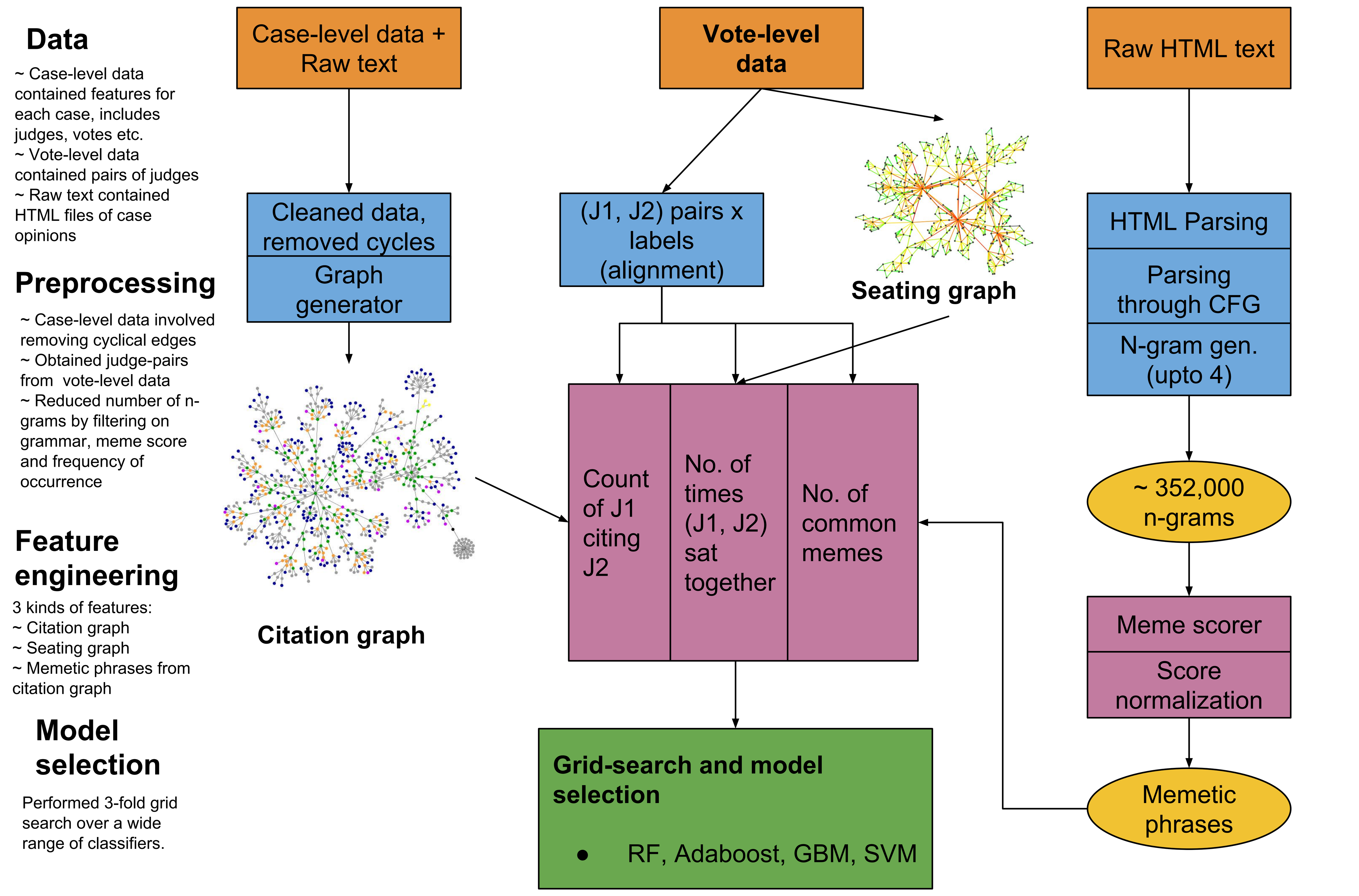
Data

- 387,000 U.S Circuit Court Cases Raw Records
  - 2-grams through 4-grams generated from the case records
  - About 352,000 n-grams in total
- 387,000 U.S Circuit Court Cases Judge Voting Records
  - 426 features for each case, including judge bio and case description
- 18,801 hand-coded U.S Circuit Court Cases Records
  - 868 features for each case, mainly hand-coded case features.

We used weighted sampling for taking care of the unbalanced nature of the dataset.



Pipeline



Feature Engineering

There were three main types of feature engineering involved:

1. **Using Citation graph**
   - Traverse along graph to find out number of times J1 cites J2
2. **Using seating graph**
   - Traverse along graph to find out number of times J1 and J2 sat together previously
3. **Using memetic phrases from raw text**
   - Filter out n-grams based on meme scores, and normalize based on frequency of occurence to remove non-memes
   - Count number of times J1's memes occur in J2's opinions

Results

Classifier	F1S (+1)	F1S (-1)
Majority Classifier	0.98	0.00
SVM	0.48	0.07
Random Forests	0.98	0.36
AdaBoost (DT)	0.94	0.22
AdaBoost (RF)	0.98	0.42

The top features, in decreasing order of importance:

Endogenous	Exogenous	Seating
Wlengthopin	decade2	n grams
totalcites	day	sat together
opinstat	j2score	
votingvalence	distance	
negativecites	state	
liberalvote	treat	
	month	

Memetic Scores

We did POS tagging on the opinion texts, and filtered out the n-grams using a context-free grammar purposed for legal language [1]:

$$
\begin{aligned}
S &\rightarrow \text{TWO} \mid \text{THREE} \mid \text{FOUR}\\
\text{TWO} &\rightarrow \text{A N} \mid \text{N N} \mid \ldots\\
\text{THREE} &\rightarrow \text{N N N} \mid \text{A A N} \mid \ldots\\
\text{FOUR} &\rightarrow \text{N C V N} \mid \text{A N N N} \mid \ldots\\
\text{A} &\rightarrow \text{JJ} \mid \text{JJR} \mid \text{JJS}\\
\text{N} &\rightarrow \text{NN} \mid \text{NNS} \mid \text{NNP} \mid \ldots\\
\text{V} &\rightarrow \text{VB} \mid \text{VBD} \mid \text{VBG} \mid \ldots
\end{aligned}
$$

We use the following expression to score the memeticity of n-grams [2]:

$$P_m = \frac{d_{m\to m}}{d_{\to m}+\delta} \Big/ \frac{d_{m\to \cancel{m}}+\delta}{d_{\to \cancel{m}}+\delta}$$

Phrase	Meme Score
red heat	0.138
salvage services	0.0039
said cars	0.0029
Atlantic coast	0.00216
citizens of different states	0.00212
insurance effected	0.0020
separable controversy	0.0018
taken in tow	0.0017
schooner was	0.00126

Conclusion & Future Work

Obtained best performance with a **AdaBoost classifier using Random Forests** with a weighted F1-score of 0.96. However, the model still has a lot of scope for improvement in the negative samples due to unbalanced nature of the data. We found that the seating history between judges, along with ideological similarity and length of court opinion were important.

Future work would involve using **memetic features such as co-occurence of memes**, influence of J1 on J2 based on memetic citation and other features using the seating and citation graphs. Using representation learning (ala **Word2Vec**) would be another possible avenue.