

1) Refugee Courts

[Time-series prediction] Refugees seeking asylum are assigned to a randomly chosen judge. 400K decisions of 350 judges in over 40 courts. The refugee is either granted asylum or not. For each refugee, we have a record of the judge they were assigned to and the timestamp of the trial, as well as information about the refugee (e.g. nationality, defensive/affirmative, lawyer). The obvious prediction problem here is whether or not the refugee is granted asylum. Part of the challenge of this problem is to set it up in a proper sequential way, to simulate building a model that can then be used for prediction of future trials. The rest of the challenge is to find all the things (via feature generation) that could be predictive (e.g. judge ID, presumed skintone of refugee based on nationality, time-of-day, most recent decisions by the same judge, recent decisions in the courthouse, recent successes of the lawyer, characteristics of the individuals the judges recently saw, etc.), and whether these factors affect disparities.

2) Circuit Courts

[Judge, Text, Citations, Classification] For a prediction problem on cases, we have the voting patterns, N-grams, citation patterns of Circuit Court judges on over 300,000 cases from 1891 to 2013. Judges are randomly assigned in panels of 3. A 5% sample was manually coded for the liberal or conservative voting valence along with 100+ other variables. Using the hand-coded label, predict valence in the larger dataset using the text of the 300,000 cases over a pre-defined set of legal topics. Using only the history of who the judges previously sat with on panels and how the judges' votes aligned with the panelists, predict how the judge votes on the next panel. Using only the citation network, predict the memetic phrases that are likely to be passed along the network in forward citation but not appear in a distant case in the citation graph. Using only the citation network, predict the memetic citations that are likely to be passed along the network in forward citation but not appear in a distant case in the citation graph. Using only the seating network, predict the memetic phrases that are likely to be passed along the network after sitting together but not appear in a distant seat in the seating graph. Predict which judges are more likely to influence how others cite and write after sitting with them. Predict the boundaries of citation clusters using only data on backwards citations. Predict higher court decisions from the way the lower court opinion's fact pattern is summarized. Predict where the fact pattern is located using only the text of the opinions.

3) Circuit Courts

[Text, Citation, Classification] A sample of 6,500 cases have been hand-coded for meaning, like pro-plaintiff or pro-defendant, pro-business or pro-environment, pro-criminal defendant rights or pro-prosecutor, etc., in 25 politically salient legal areas (Capital Punishment, Criminal Appeals, Campaign Finance, Affirmative Action, Gay Rights, Abortion, Gender Discrimination, Sexual Harassment, Title VII, Segregation, Obscenity, Establishment Clause, Americans with Disabilities Act, Piercing the Corporate Veil, Takings, National Labor Review Board, Environmental Protection Act, National Environmental Policy Act, Federalism, First Amendment, Punitive Damages, Standing, Eleventh Amendment, Federal Communications Commission, and Contracts). Using only the (i) text and (ii) precedent it cites, predict the hand-coded label.

4) District Court Criminal Cases

[Time-series prediction] Criminal defendants are assigned to a randomly chosen District Court judge. 1 million sentencing decisions in 94 U.S. District Courts from 1992-2009. The obvious prediction problem here is the sentence length. Part of the challenge of this problem is to set it up in a proper sequential way, to simulate building a model that can then be used for prediction of future trials. The rest of the challenge is to find all the things (via feature generation) that could be predictive (e.g. judge ID, most recent decisions by the same judge, recent decisions in the courthouse, characteristics of the individuals the judges recently saw, etc.). What is interesting from a social and legal perspective is whether extraneous factors like (i) characteristics of the individuals the judges recently saw, (ii) sporting events, and (iii) weather ends up predicting decisions and if so, what factors are relevant and by how much, and whether these factors affect disparities.

5) Supreme Court Oral Arguments

[Audio, Classification] Lawyers always use the exact same sentence when they introduce themselves to the Supreme Court: "Mr. Chief Justice, (and) may it please the Court." We have clipped this data for 1955-2013 comprising over 8000 audio recordings, spoken by many different lawyers over time. Mechanical turk workers have rated 1999-2013 sample (2000+ recordings) based on whether they sound "confident", "trustworthy", "attractive", "masculine", etc. We also have data on the Mturk workers, the Supreme Court cases, and the Supreme Court oral advocates (including their faces). Using the hand-coded label, predict the perceptions on the 1955-2013 set of audio files. We also have the previous best prediction model of Supreme Court outcomes, which incorporates features such as the judges, case characteristics, characteristics of the lower court of origin, and historical trends at the lower courts and of the judges; do addition of features generated from the raw sound file or lawyers' faces predict outcomes?

[Judge, Audio, Text, Classification] We have the entire 1-hour audio files and transcripts for oral arguments in the Supreme Court dating back to 1955 along with the opinions. Features have also already been generated for some raw sound files. Comparing against the previous best prediction model of Supreme Court outcomes, do addition of transcript features predict voting? Can convergence in how judges speak predict voting together over time?

6) New Orleans District Attorney office

[Time-series prediction] We have data on arrested individuals constituting the universe for the New Orleans District Attorney office for over a decade with many stages of random assignment (screener, federal prosecutor, and judge). The obvious prediction problem here is the sentence length, whether the individual was guilty, and whether the individual was sent home without trial. Part of the challenge of this problem is to set it up in a proper sequential way, to simulate building a model that can then be used for prediction of future cases. The rest of the challenge is to find all the things (via feature generation) that could be predictive (e.g. judge ID, most recent decisions by the same judge, recent decisions in the courthouse, characteristics of the individuals the judges recently saw, characteristics of the screener, what the screener recently saw, characteristics of the prosecutor, etc.). What is interesting from a social and legal perspective is whether extraneous factors like (i) the screener identity, (ii) prosecutor identity, and (iii) judge ends up predicting decisions and predicts inequality of treatment of individuals by race and gender, and if so, what factors are relevant and by how much, and whether these factors affect disparities.

7) Societal Attitudes

We have data on social attitudes of Americans since 1972 and the social attitudes of a global sample of individuals. Predict the attitudes using demographic characteristics or any other social, political, economic feature. What is interesting from a social and legal perspective is whether legal factors like (i) U.S. court decisions in a sample of 6,500 cases have been coded for meaning, like pro-plaintiff or pro-defendant, pro-business or pro-environment, pro-criminal defendant rights or pro-prosecutor, etc., in 25 politically salient legal areas (Capital Punishment, Criminal Appeals, Campaign Finance, Affirmative Action, Gay Rights, Abortion, Gender Discrimination, Sexual Harassment, Title VII, Segregation, Obscenity, Establishment Clause, Americans with Disabilities Act, Piercing the Corporate Veil, Takings, National Labor Review Board, Environmental Protection Act, National Environmental Policy Act, Federalism, First Amendment, Punitive Damages, Standing, Eleventh Amendment, Federal Communications Commission, and Contracts) affect attitudes in the direction the law intended, in the opposite direction the law intended, or polarizes individuals, (ii) the extraneous factors like judicial biographies randomly assigned to cases affect attitudes. Compare the predictive model with one for world sample.

8) World War I Desertion

We have digitized several World War I British archival datasets, including universe of deserters reported in military diaries, police gazettes, and handwritten military trials, commuted and executed capital sentences, geocoded casualties, maps, officer lists, ethnicity dictionaries, and order of battle. The obvious prediction problem here is desertion using any design features, such as previous desertion, casualties, battles, Irish or British identity, Easter Rising, etc.

9) Pharmaceutical company payments to physicians

We have several data sources on pharmaceutical company payments to physicians, including data from legal settlements disclosing \$316 million in pharmaceutical company payments to 316,622 physicians across the U.S. from 2009-2011 and data after the Affordable Care Act published the payments, along with the biographies of physicians. Construct a predictive model of physician payments from pharmaceutical companies. The effects of disclosure laws on what is being disclosed are typically unknown since data on disclosed activity rarely exist in the absence of disclosure laws. Does the predictive model change before and after mandatory disclosure?

10) Federal judicial clerks

We have data on judges' decisions and data on some clerks and their subsequent careers. Construct a predictive model of judicial quality, that is, the opinions' subsequent citations, reversal, length, and previous citations using characteristics of the case, the judges' experience, and the clerks' biographies. Does the periodic unravelling of the market for judicial clerks predict judicial quality beyond the best prediction models? Can clerks' biographies or face features be predicted judges' ideology and/or market unraveling?

11) Auxiliary data: weather and sporting events

We have data on daily weather in all the U.S. courthouses and data on sporting events (NFL, NBA, etc.) for the cities where the courthouses reside. These comprise a novel feature to merge into the respective datasets.

12) Population-representative U.S. data

We have population-representative U.S. data on labor markets, marriage markets, voting outcomes, crime outcomes, economic growth, property prices, etc. Predict the outcomes using demographic characteristics or any other social, political, economic feature. What is interesting from a social and legal perspective is whether legal factors like (i) U.S. court decisions in a sample of 6,500 cases have been coded for meaning, like pro-plaintiff or pro-defendant, pro-business or pro-environment, pro-criminal defendant rights or pro-prosecutor, etc., in 25 politically salient legal areas (Capital Punishment, Criminal Appeals, Campaign Finance, Affirmative Action, Gay Rights, Abortion, Gender Discrimination, Sexual Harassment, Title VII, Segregation, Obscenity, Establishment Clause, Americans with Disabilities Act, Piercing the Corporate Veil, Takings, National Labor Review Board, Environmental Protection Act, National Environmental Policy Act, Federalism, First Amendment, Punitive Damages, Standing, Eleventh Amendment, Federal Communications Commission, and Contracts) affect outcomes and (ii) the extraneous factors like judicial biographies randomly assigned to cases affect attitudes, and if so, what factors are relevant and by how much.

13) Machine-learning meets algorithmic mechanism design

Consider a task like data entry from web searches for an inexperienced human rater found on online labor markets. After assigning a pilot task, predict who is the best at completing the task and predict the optimal incentive scheme using a fully randomized set of time-varying incentives.

14) Social violence

We have data on social violence after a financial crisis. Construct a predictive model of ethnic-religious conflict using any features of interest. Do measures of religious fundamentalism matter beyond the best prediction models?

15) Circuit Court Oral Arguments

A large number of Circuit Court oral arguments are available and nothing is known about whether linguistic features are relevant. We have all the Circuit Court meta-data from #2 and #5 training set for clipping the introductory sentences.