Risk Analysis for Intellectual Property Litigation

Mihai Surdeanu*, Ramesh Nallapati*, George Gregory#, Joshua Walker#, Christopher D. Manning*

* Stanford University
# LexMachina (www.lexmachina.com)
Litigation Risk Analysis

Estimate the likelihood of winning at the time of filing

Why:

Should I settle?
Does attorney choice matter?
Where do I file a defensive case if XYZ Co. plans to sue me?
Can I win?
Case Outcome
This is what we try to predict

Case Events
Can be used to estimate merits

Case Entities
Past behavior may influence the current outcome
How?

• 2-step process:
  – Step 1: mine all the *previous* and *concurrent* interactions between the involved entities
    • Parties, attorneys, law firms, judges, districts
    • Not trivial: requires entity resolution
  – Step 2: detect patterns in these interactions and extrapolate to the new case
    • Logistic regression or conditional random fields
Using Historical Information

- Plaintiff
- Judge
- Defendant 1
- Defendant 2
- Attorney

Balance of power among Plaintiff, Judge, Defendents, and Attorneys.
Exploiting the Correlation between Concurrent Cases

81% of such cases have identical outcomes
36% of cases have at least one link predict their outcomes jointly using conditional random fields!
Features

• Past performance:
  – Win rates in a given role (plaintiff or defendant)
  – Win rates in any role
  – Frequency of litigation
  – Judge and district bias

• Concurrent information
  – Outcomes of immediate neighboring cases
The Corpus

20,980 total cases

<table>
<thead>
<tr>
<th>Cases with polarized outcome</th>
<th>Parties</th>
<th>Attorneys</th>
<th>Law firms</th>
<th>Judges</th>
<th>Districts</th>
</tr>
</thead>
<tbody>
<tr>
<td>4,263</td>
<td>12,270</td>
<td>15,706</td>
<td>5,261</td>
<td>1276</td>
<td>88</td>
</tr>
</tbody>
</table>

3,243 cases with some historical information
## Overall Results

<table>
<thead>
<tr>
<th></th>
<th>Baseline (Plaintiff Wins)</th>
<th>Only Past Information</th>
<th>Past and Concurrent Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>52.4</td>
<td>63.4*</td>
<td>64.0*</td>
</tr>
</tbody>
</table>
Ablation Experiment

Drop in performance upon removal of feature

- Plaintiff
- Defendant
- Plaintiff's attorney
- Defendant's attorney
- Plaintiff's law firm
- Defendant's law firm
- Judge
- District

Counsel choice matters
There is bias in the system
Second Ablation Experiment

Drop in performance upon removal of feature

- performance in a given role *
- performance in any role/bias *
- frequency *

Legend:
- * indicates statistical significance.
Conclusions

• Assessed the risk for parties involved in IP litigation using only historical and concurrent features of the participating entities

• Applications
  – Help parties involved in an IP lawsuit make well-informed strategic decisions
  – Reduce the number of cases that reach trial ➔ reduce costs for parties and judicial system

• Future work
  – Including estimators of case merits into the predictive model