Classification of Legal Expertise and Work from Profile Data

Chase Basich, Austin Chambers

Problem

One of the biggest challenges law firms face is:

1. Locating attorneys having the right **expertise**.
2. Locating **past work** to help with current work or to pitch for new business.

Without solutions to the above problems, firms lose a tremendous amount of time and opportunity.

Solution

- To classify legal work and attorney expertise according to the firm's unique set of legal practices
- These classifications solve the above problems by enabling expertise/work search by the practice classification.

Data

- Datasets were acquired by scraping the public profiles of attorneys and matter narratives, short summaries of cases, from the websites of two firms.

	Profiles	Narratives	Practices
Firm A	143	521	85
Firm B	3659	0	101

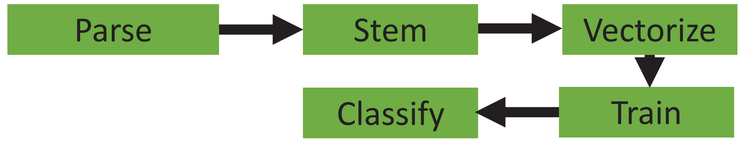
- Narratives and profiles were inconsistently tagged to a subset of relevant multiple practices.

Sample Narrative: "Obtained judgment on behalf of a Michigan retail store in an age discrimination case"

Sample Profile: "John Smith represents employers in responding to and handling charges of employment discrimination before the EEOC and state

Expected Classification: **Employment Law Practice**

Classification Pipeline



Features

- **Narrative**

Each narrative consists of a plain text description of the case, which is stemmed and vectorized, as well as the name of each attorney who worked on the case. Each narrative was labeled with at least one practice area.

- **Attorney**

Each attorney has a plaintext biography which is stemmed and vectorized. Each attorney was labeled with at least one practice area

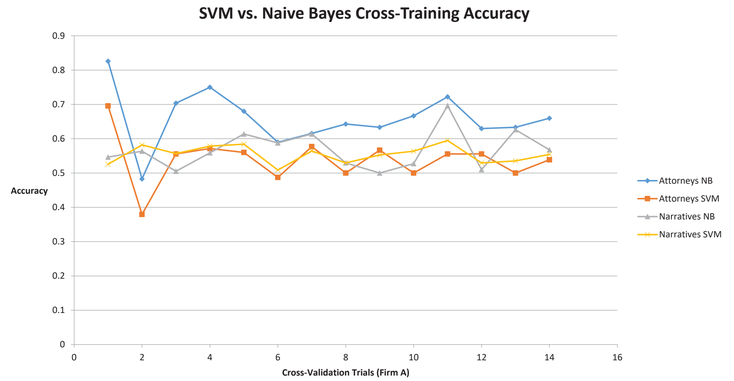
Classification Schemes

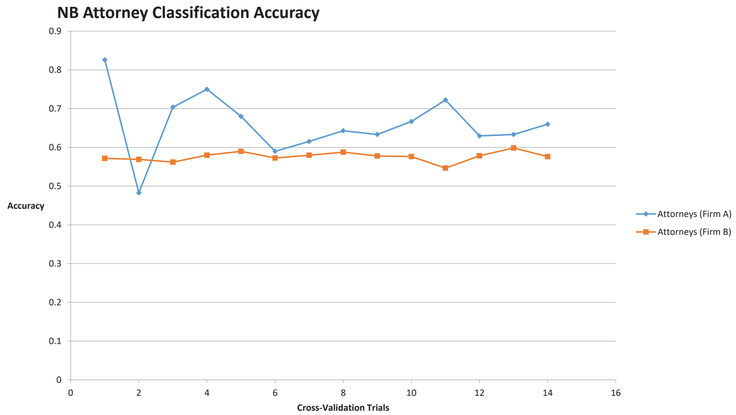
To evaluate our models we used four classification schemes. We classified each lawyer and narrative using only the features from the lawyer and narrative, respectively. We also classified each lawyer and narrative while adding the features from the corresponding narratives and corresponding lawyers.

Models

We compared two different models: a multi-class one-vs-rest SVM and a TF-IDF Naïve Bayes model. The same features were used to evaluate each model, with the Naïve Bayes classifier outperforming the SVM.

Results





Analysis and Future Work

With its high rate for attorney classification, our classifier could be used as a recommending tool for a law firm: given a case it could classify the case and select an attorney with matching areas of expertise. We also found that the success of the classification increased by 10% if a description of each practice area was provided.

Our next step would be to extend this to provide multiple classifications per narrative and per attorney, and to use the similarities to better match an attorney to a case.