Expert Finding in Legal Community Question Answering

Arian Askari, Suzan Verberne, Gabriella Pasi
Leiden University and University of Milano-Bicocca
{a.askari, s.verberne}@liacs.leidenuniv.nl gabriella.pasi@unimib.it

Problem

An Applicant + A problem that should be solved legally in court by hiring a lawyer

- There are no studies on expert finding in the legal domain.
- Differences with other domains in terms of expertise and authority, and content:
  - Asker knows almost nothing about the law
  - Answerer is a lawyer (professional user)
  - Content is a combination of formal and informal language and it may contain emotional language

Proposed Method

Goal: Find a lawyer that have specific expertise

1. Comment-based: concatenate of comments on lawyer’s responses
2. Recency-based: concatenate of latest responses by the lawyer
3,4. Sentiment-Based: we create two profiles based on positive sentences and negative sentences in lawyer’s responses

All Lawyers’ Answers → Generate profiles

Model 2 BM25 → Re-rank by VBD

Compute relevance score per profile using VBD

There is a need to find lawyer based on the expertise for a case of an applicant since the lawyer’s impact is the greatest when they work in their expert field.

Result & Insights

<table>
<thead>
<tr>
<th>Model</th>
<th>MAP</th>
<th>MRR</th>
<th>P@1</th>
<th>P@5</th>
<th>P@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 LM</td>
<td>22.8 %</td>
<td>40.9 %</td>
<td>23.9 %</td>
<td>19.7 %</td>
<td>13.2 %</td>
</tr>
<tr>
<td>Model 1 BM25</td>
<td>3.7 %</td>
<td>7.0 %</td>
<td>2.9 %</td>
<td>1.4 %</td>
<td>2.6 %</td>
</tr>
<tr>
<td>Model 2 LM</td>
<td>19.4 %</td>
<td>21.9 %</td>
<td>13.5 %</td>
<td>12.7 %</td>
<td>7.8 %</td>
</tr>
<tr>
<td>Model 2 BM25</td>
<td>21.0 %</td>
<td>36.6 %</td>
<td>22.5 %</td>
<td>18.3 %</td>
<td>11.5 %</td>
</tr>
<tr>
<td>VBD</td>
<td>37.3 %</td>
<td>70.7 %</td>
<td>60.5 %</td>
<td>55.6 %</td>
<td>25.9 %</td>
</tr>
<tr>
<td>VBD + Profiles (Weighted)</td>
<td>39.3 %</td>
<td>73.2 %</td>
<td>64.9 %</td>
<td>57.1 %</td>
<td>27.7 %</td>
</tr>
</tbody>
</table>

- Model 1: Candidate Level
- Model 2: Document Level
- Vanilla BERT Document-based (VBD): We re-rank experts’ answers by a pretrained BERT model with a linear combination layer stacked atop (VBD) the classifier using Model 2 BM25 as initial ranker.
- The weights of profiles:
  - Comments > Recency > Sentiment positive > Sentiment Negative

Dataset

- Our dataset has been scraped from the Avvo QA forum. Avvo is a legal online platform where anyone could post their legal problem for free and receive responses from lawyers.