Comparing Features for Ranking Relationships Between Financial Entities based on Text

Tim Repke, Michael Loster, Ralf Krestel

FEIII Challenge 2017

Financial Entity Identification and Information Integration Challenge 2017

Given:
- set of tuples (financial entity A, role, financial entity B)
- text snippet the tuple was extracted from (10-K, 10-Q filings)
- tuples labelled by experts as (highly) relevant, neutral, irrelevant

Challenge:
- per role, rank relationship tuples by relevance

Our Approach

Bag-of-Words (BOW)
Utilise wording that correlates with relevance

Paragraph Embeddings (EMB)
Dense vector representations of sentences with embeddings trained on full-length filings

Syntax Features (SYN)
Manually derived features independent of wording on character, token, and POS-tag level

Features

The Dataset

Role Distribution

<table>
<thead>
<tr>
<th>Role</th>
<th>Training</th>
<th>Texting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affiliate</td>
<td>186</td>
<td>120</td>
</tr>
<tr>
<td>Agent</td>
<td>61</td>
<td>40</td>
</tr>
<tr>
<td>Counterpart</td>
<td>64</td>
<td>108</td>
</tr>
<tr>
<td>Guarantor</td>
<td>34</td>
<td>28</td>
</tr>
<tr>
<td>Insurer</td>
<td>19</td>
<td>47</td>
</tr>
<tr>
<td>Issuer</td>
<td>129</td>
<td>98</td>
</tr>
<tr>
<td>Seller</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>Servicer</td>
<td>21</td>
<td>57</td>
</tr>
<tr>
<td>Trustee</td>
<td>420</td>
<td>304</td>
</tr>
<tr>
<td>Underwriter</td>
<td>21</td>
<td>40</td>
</tr>
</tbody>
</table>

All         | 975      | 900     |

Vote

BOW: 0.88 (0.08) EMB: 0.89 (0.09) SYN: 0.94 (0.06) Vote: 0.95 (0.04)

35 1 25 0.09 0.09 0.19 0.19
11 36 1 60 0.09 0.09 0.14 0.14
7 36 1 60 0.09 0.09 0.14 0.14
4 12 (0.9) 4 (0.88) 7 (0.74) 12 (0.82) 8 (0.74) 4 0.06
60

Features

Results

Approach: Random, Worst, BOW, EMB, SYN

<table>
<thead>
<tr>
<th>Approach</th>
<th>NDCG (avr)</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>0.88 (0.08)</td>
<td>--</td>
</tr>
<tr>
<td>Worst</td>
<td>0.72 (0.04)</td>
<td>--</td>
</tr>
<tr>
<td>BOW</td>
<td>0.88 (0.08)</td>
<td>0.34 (0.13)</td>
</tr>
<tr>
<td>EMB</td>
<td>0.89 (0.09)</td>
<td>0.24 (0.14)</td>
</tr>
<tr>
<td>SYN</td>
<td>0.94 (0.06)</td>
<td>0.44 (0.11)</td>
</tr>
<tr>
<td>Vote</td>
<td>0.95 (0.04)</td>
<td>0.43 (0.12)</td>
</tr>
</tbody>
</table>

Confusion Matrix

- Good performance for highly relevant samples
- Predictions biased towards relevant
- Weighting during training doesn’t resolve bias
- Classification performance only secondary, using predictions to calculate ranking score partly resolves bias

Vote (BOW+EMB+SYN): soft vote of BOW, EMB, and SYN

5-fold cross-validation
- Trained on ~900 training samples, leaving out samples from 5 documents
- Tested on all testing samples, evaluation per role and aggregated

Implications on NDCG

- Best-case (random): average NDCG of 100 random sorts
- Worst-case (worst): NDCG of inverse perfect order

Challenge Results

Confusion Matrix

- Good performance for highly relevant samples
- Predictions biased towards relevant
- Weighting during training doesn’t resolve bias
- Classification performance only secondary, using predictions to calculate ranking score partly resolves bias

Approach: Random, Worst, BOW, EMB, SYN

<table>
<thead>
<tr>
<th>Approach</th>
<th>NDCG (avr)</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>0.88 (0.08)</td>
<td>--</td>
</tr>
<tr>
<td>Worst</td>
<td>0.72 (0.04)</td>
<td>--</td>
</tr>
<tr>
<td>BOW</td>
<td>0.88 (0.08)</td>
<td>0.34 (0.13)</td>
</tr>
<tr>
<td>EMB</td>
<td>0.89 (0.09)</td>
<td>0.24 (0.14)</td>
</tr>
<tr>
<td>SYN</td>
<td>0.94 (0.06)</td>
<td>0.44 (0.11)</td>
</tr>
<tr>
<td>Vote</td>
<td>0.95 (0.04)</td>
<td>0.43 (0.12)</td>
</tr>
</tbody>
</table>

Relevance Distribution

- Highly relevant
- Relevant
- Neutral
- Irrelevant

Training
- Ranking

Features

Syntax Features (SYN)
Manually derived features independent of wording on character, token, and POS-tag level

Example: Histogram of snippet length (characters)

Contact
Tim Repke (tim.repke@hpi.de)
Rudolf-Breitscheid-Str. 187
14482 Potsdam, Germany