Cross-lingual complex word identification with multi-task learning
Joachim Bingel and Johannes Bjerva

Languages as tasks

- Lets us share data between languages
- Generalizes better to new languages (without training data)
- Training by alternating between languages
- For languages without training data, predict most similar language and use output layer for that language

Model and Features

- Ensemble of 10 multitask networks and 10 random forests
- Features include length, frequency, character perplexity, semantic specificity (measured by WordNet synsets, hypernyms and hyponyms), inflectional complexity, POS, target-sentence similarity

Results

<table>
<thead>
<tr>
<th>Language</th>
<th>MAE</th>
<th>Rank</th>
<th>$\Delta$ (system)</th>
<th>$F_1$</th>
<th>Rank</th>
<th>$\Delta$ (system)</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>0.066</td>
<td>1</td>
<td>0.012 (TMU)</td>
<td>0.7595</td>
<td>1</td>
<td>0.013 (TMU)</td>
</tr>
<tr>
<td>German</td>
<td>0.075</td>
<td>2</td>
<td>-0.013 (TMU)</td>
<td>0.6621</td>
<td>5</td>
<td>-0.083 (TMU)</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.079</td>
<td>3</td>
<td>-0.007 (TMU)</td>
<td>0.7458</td>
<td>5</td>
<td>-0.024 (TMU)</td>
</tr>
</tbody>
</table>

Analysis

- Word length good predictor, but false positives tend to be long and false negatives short
- False negatives are mostly words rated as complex by very few annotators

Contact
bingel@di.ku.dk