INTRODUCTION

In our past work, we demonstrated Stacking with auxiliary features (SWAF) for improving performance of slot filler systems.

Stacking uses supervised learning and thus relies on training data.

Sometimes, we might not have access to individual systems and only to the system outputs.

We first use unsupervised ensembling to combine such systems without training data and then use stacking to combine the supervised ensemble and the unsupervised ensemble.

We obtain state-of-the-art results on two KBP tasks for 2015:

1. Cold Start Slot Filling (CSSF)
2. Tri-lingual Entity Discovery and Linking (TEDL)

OVERVIEW OF THE KBP TASKS

1. Cold Start Slot Filling (CSSF)

- Extract information (titles) about specific attributes (slots) for a set of entities (queries) from a given corpus.
- Query entities can be PER/ORG/GPE.
- The font size of your title should be bigger than your name(s) and only to the system outputs.
- The title box is automatically set to fit the title box. You can manually override this feature and customize the title box as needed.

2. Tri-lingual Entity Discovery and Linking (TEDL)

- Find all entity mentions in a corpus of English, Chinese, and Spanish documents and link to a BaseKB KB.
- If there is no KB entry for the entity, systems are expected to cluster all the mentions for that entity using a NIL ID.
- Entities can either be PER/ORG/GPE/LCO/FAC.

ENSEMBLING ALGORITHM

- For systems common between 2014 and 2015, we use stacking to aggregate the outputs.
- For systems without training data, we use constrained optimization approach described in Weng et al. to aggregate confidence scores.
- Use the aggregated unsupervised ensemble as one more system to the stacker.
- Unsupervised ensemble from 2014 is used as training data for this new aggregated system.
- Instances classified as correct are kept and post-processed.
- The final output is made to look like its produced by a single system by resolving any conflicts.

RESULTS

<table>
<thead>
<tr>
<th>Approach</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle Voting Baseline (≥ 3)</td>
<td>0.4384</td>
<td>0.2720</td>
<td>0.3357</td>
</tr>
<tr>
<td>Top ranked CSSF system in 2015 (Angeli et al.)</td>
<td>0.3989</td>
<td>0.0588</td>
<td>0.1462</td>
</tr>
<tr>
<td>Stacking approach by Rajani et al.</td>
<td>0.6566</td>
<td>0.4312</td>
<td>0.5371</td>
</tr>
<tr>
<td>Combining supervised and unsupervised ensembles</td>
<td>0.6797</td>
<td>0.4314</td>
<td>0.5480</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approach</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle Voting Baseline (≥ 4)</td>
<td>0.5114</td>
<td>0.4681</td>
<td>0.4954</td>
</tr>
<tr>
<td>Top ranked TEDL system in 2015 (Sil et al.)</td>
<td>0.695</td>
<td>0.5472</td>
<td>0.6110</td>
</tr>
<tr>
<td>Stacking approach by Rajani et al.</td>
<td>0.813</td>
<td>0.515</td>
<td>0.630</td>
</tr>
<tr>
<td>Combining supervised and unsupervised ensembles</td>
<td>0.686</td>
<td>0.624</td>
<td>0.653</td>
</tr>
</tbody>
</table>

Table 2: Results on the 2015 CSSF task using the official NIST scorer and CEAF metric.

![Figure 1](image1.png)

This PowerPoint 2007 template produces a 36”x48” page.

RESEARCH POSTER PRESENTATION DESIGN © 2015

ACKNOWLEDGEMENTS

This research was supported by DARPA DEFT program under AFRL grant FA8750-13-2-0026.

CONTACT

Nazneen Fatema Rajani: mmrajani@cs.utexas.edu
Ray Mooney: mooney@cs.utexas.edu

REFERENCES


CONCLUSIONS

- Stacking-based approach to ensembling both supervised and unsupervised systems is very promising on CSSF and TEDL tasks.
- Our model outperforms top ranked systems from the 2015 competition as well as several other ensembling methods on both tasks.
- We obtained state-of-the-art results for CSSF and TEDL.
- Adding the unsupervised ensemble along with the shared systems leads to a substantial increase in recall.

TABLE 1: Number of supervised and unsupervised systems for each of the tasks and languages

<table>
<thead>
<tr>
<th>System</th>
<th>Number of supervised systems (N)</th>
<th>Number of unsupervised systems (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSSF</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>TEDL</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1: Number of supervised and unsupervised systems using SWAF

ACKNOWLEDGEMENTS

This research was supported by the DARPA DEFT program under AFRL grant FA8750-13-2-0026.

CONTACT

Nazneen Fatema Rajani: mmrajani@cs.utexas.edu
Ray Mooney: mooney@cs.utexas.edu

REFERENCES


ACKNOWLEDGEMENTS

This research was supported by the DARPA DEFT program under AFRL grant FA8750-13-2-0026.

CONTACT

Nazneen Fatema Rajani: mmrajani@cs.utexas.edu
Ray Mooney: mooney@cs.utexas.edu

REFERENCES