An Efficient Active Learning Framework for New Relations

Lisheng Fu
Ralph Grishman
New York University
Oct. 16, 2013
Relation extraction

Automatic Content Extraction (ACE) 04 Examples:

- the CEO of Microsoft (Employment)
- a military base in Germany (Located)
- U.S. businessman (Citizen)
Relation extraction

- High annotation cost for supervised learning
- Limited performance for semi-supervised methods
- Active learning can be more feasible in practice, especially for new relations
Active Learning of Relation Extraction (Previous work)

- Not much for relation extraction
- Mostly include uncertainty-based strategies (Roth & Small 2008, etc.)
- Effective
- Different views of data help (Distributional similarity between relations) (Sun and Grishman 2012)
Active Learning of Relation Extraction (Previous work)

1. Ask users for a small set of seeds for one relation
2. Train a local feature model and a model based on distributional similarity
3. Run co-testing between these two models
4. Ask users queries from the contention set (Yes/No whether the example is the target relation)
5. Update models
Active Learning of Relation Extraction (Previous work)

...will become the next [US President]. (unlabeled)

<table>
<thead>
<tr>
<th>Local Features (Maximum Entropy)</th>
<th>Distributional Similarity (Nearest Neighbor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>US President</td>
</tr>
<tr>
<td>Entity Type</td>
<td>Texas Governor (e.g. labeled)</td>
</tr>
<tr>
<td>Chunk</td>
<td></td>
</tr>
<tr>
<td>Dependency</td>
<td>US senator</td>
</tr>
<tr>
<td>Parse</td>
<td></td>
</tr>
</tbody>
</table>

Based on the context of the phrase in a large unlabeled corpus

NOT Employment

Employment

0.971

0.462
Problems of Previous Work

• Lack of negative examples in the initial seed set
  - not representative of corpus
  - requires more queries answered as no
• Does NOT use examples on which models agree
  - slows learning
• Does Not favor discovery of novel examples
Improve the Balance of Positive and Negative Examples in the Seed Set

- Select negative examples automatically
  - improve balance without burdening users
  - use entity type constraints to reduce risk of selecting positives

- Queries will be closer to the boundary
  - speeds learning
Experiments (Early Iterations)

Dataset: ACE 04

30 iterations = 150 examples

F1 score is the average of 6 types

B: Baseline

N: Non Relation Selection
Interleaving Self-training

- Examples on which two models agree = agreement set.
- After each iteration, bootstrap positive and negative examples of high confidence from the agreement set.
- Less risk of semantic drift than with pure self-training which is based on a single model.
- Magnifies speed-up of active learning.
Experiments (Early Iterations)

Dataset: ACE 04
30 iterations = 150 examples
F1 score is the average of 6 types
B: Baseline
N: Non Relation Selection
S: Interleaving Self-training
Encouraging Novelty

- Distributional similarity model is better source of novel examples
  - examples that might contain new lexical features
  - underweighted relative to the local feature model in the previous work.
- Favor examples marked as positive by the distributional similarity model in selecting queries
Experiments (Early Iterations)

Dataset: ACE 04

30 iterations = 150 examples

F1 score is the average of 6 types

B: Baseline

N: Non Relation Selection

S: Interleaving Self-training

G: Encouraging Novelty
Experiments (Further Annotation)

Dataset: ACE 04
200 iterations = 1000 examples
Most frequent type in ACE 04
Entity Type Constraints

• Use hard rules to answer queries automatically
• Usually not a problem in supervised learning
• Could save annotation depending on the type
• Poses risk if the NE tagger is not good

<table>
<thead>
<tr>
<th>Type</th>
<th># queries in total</th>
<th># queries that filters apply</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMP-ORG</td>
<td>1000</td>
<td>91</td>
<td>9.1%</td>
</tr>
<tr>
<td>PHYS</td>
<td>1000</td>
<td>106</td>
<td>10.6%</td>
</tr>
<tr>
<td>GPE-AFF</td>
<td>590</td>
<td>84</td>
<td>14.2%</td>
</tr>
<tr>
<td>PER-SOC</td>
<td>234</td>
<td>64</td>
<td>27.3%</td>
</tr>
<tr>
<td>ART</td>
<td>151</td>
<td>54</td>
<td>35.8%</td>
</tr>
<tr>
<td>OTHER-AFF</td>
<td>105</td>
<td>56</td>
<td>53.3%</td>
</tr>
</tbody>
</table>

# queries saved under the perfect tagger
Conclusion & Future Work

• We further boosted active learning by adding examples automatically and favoring novelty examples

• Find a better distributional similarity metric between relations or other different views

• Analyze different behaviors of different types of relations
Q & A

• Thank you.
Seeds and Queries

Question 2

[(firestone)] is [(decatur)]'s third-largest employer.

Is this a EMP-ORG relation?

[Yes]  [No]  [More Context]  [Not Sure]  [Quit]