QA-Driven Relation Extraction

Zeqiu Wu, Anirud Yadav, Wenxuan Zhou
Outline

• Introduction
• Challenges and Solution
• System Design
• Experiments
• Future Work
Outline

- Introduction
- Challenges and Solution
- System Design
- Experiments
- Future Work
Project Overview

We leverage the feedback signals from an existing question answering (QA) system, Jacana (Yao et al., 2013) to improve the performance of one of the state-of-the-art relation extraction (RE) model, CoType (Xiang et al., 2017).
**Task Definition:**

For each sentence, we would like to extract every entity pair if a target relation type between the 2 entities is expressed in that sentence, in the form of a relation tuple: (entity mention 1, relation type, entity mention 2).

The Women’s March was a worldwide protest on January 21, 2017. The protest was aimed at Donald Trump, the recently inaugurated president of the United States. The first protest was planned in Washington, D.C., and was known as the Women’s March on Washington.
CoType

CoType is a system that *jointly extracts typed entities and relations* with labeled training data heuristically obtained from knowledge bases (*Distant Supervision*).
Outline

• Introduction
• Challenges and Solution
• System Design
• Experiments
• Future Work
Two observations in CoType

Noisy training data from distant supervision

- Randomly selecting negative examples introduces *false negatives* due to the incompleteness of knowledge base.

- *False positives* come from the assumption that if two entities participate in a relation, any sentence that contains those two entities might express that relation.

  - Jim Trump, Relation: *None*, Chicago
    - Jim Trump was born in Chicago.
    - Donald Trump, Relation: *president_of*, United States
    - Donald Trump, Relation: *travelled_to*, United States
    - Donald Trump flew back to United States.
Two observations in CoType

Ignorance of quality difference of training examples:

- Each training relation mention is treated equally (weight = 1) in constructing the heterogeneous network. However, different training relation mentions can have different levels of quality in indicating the relation type.

Donald Trump, Relation: president_of, United States

S1: Donald Trump is the president of United States.

S2: Donald Trump sits in the most respected seat in United States.

Stronger indication
Proposed Solution

Use learning-to-rank mechanism from existing QA system(s) to detect false positive/negatives as well as the quality score of each training relation mention.
QA System in Our Project

We leverage a model called Jacana, which was designed for a special QA task (Answer Sentence Selection), to provide rank scores

- Input: A question and a set of sentences
- Output: Score of each sentence indicating whether the sentence contains the answer

<table>
<thead>
<tr>
<th>Question</th>
<th>Sentences</th>
<th>Score (0 - 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who is the president of United States?</td>
<td>Donald Trump is the president of United States.</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Donald Trump sits in the most respected seat in United States.</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Donald Trump flew back to United States.</td>
<td>0.1</td>
</tr>
</tbody>
</table>
Outline

• Introduction
• Challenges and Solution
• System Design
• Experiments
• Future Work
Pipeline

1. Get training data for Jacana by constructing question & positive/negative sentences
2. Train Jacana
3. Get training data (entity/relation mentions) for CoType by linking sentences to KB
4. Construct question & sentence as input for Jacana
5. Run Jacana on the constructed Q&S to get rank scores
6. Update label & weight for each training example based on the rank scores
7. Run CoType with updated training data
Question Generation

• Each target relation type maps to a question template
  • E.g. ‘place_of_birth’ -> ‘Where was $PERSON born?’

• For each positive relation mention, we construct a question with the template and entity mention 1.
  • E.g. (Donald Trump, place_of_birth, New York) -> ‘Where was Donald Trump born?’

• For each negative relation mention, we construct a question for each possible relation type.
  • E.g. (Jim Trump, None, Chicago) -> ‘Where was Jim Trump born?’, ‘Where did Jim Trump travel?’, etc
Training Data Update

• Range of rank score (s) from Jacana (linear regression): [0, 1]

• For each positive relation mention:
  • If $s > \theta_1$ : remain as positive, weight (w) = normalized s
  • Else: new label = ‘None’, w = normalized s

• For each negative relation mention:
  • If $\text{max}(s) > \theta_2$ : new label = $\text{argmax } s(y)$, w = normalized s
  • Else: remain as negative, w = normalized s
Outline

• Introduction
• Challenges and Solution
• System Design
• Experiments
• Future Work
Experiments

Datasets

• QA:
  • 100 questions and 4710 sentences from TREC QA news articles (APW, XIE, NYT) (Wang et al., 2007)

• RE:
  • Training: 100K+ sentences with 184K positive and 650K negative examples from TREC QA news articles
  • Test: 300+ sentences from NYT with 300+ human labeled relation mentions
  • 3 entity types (Person, Location, Organization), 18 relation types
Experiments

Results

- We simply use the same parameter values as claimed in CoType paper

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CoType</td>
<td>0.264113</td>
<td>0.335038</td>
<td>0.295378</td>
</tr>
<tr>
<td>CoType_QA_Label</td>
<td>0.219485</td>
<td>0.501279</td>
<td>0.305296</td>
</tr>
<tr>
<td>CoType_QA_Full</td>
<td>0.317352</td>
<td>0.355499</td>
<td>0.335344</td>
</tr>
</tbody>
</table>

CoType_QA_Label: only update label
CoType_QA_Full: update both label & weight
Case Study

• In 1995, Jordan regained sovereignty over its territories occupied by Israel and exchanged ambassadors with Tel Aviv.
  • Linked by freebase: (Israel, contains, Tel Aviv) with weight 1.0
  • Predicted by QA: (Israel, None, Tel Aviv) with weight 0.87

• Cameron spent the night at Mercy Hospital near downtown San Diego, and Beltran spent the night at Scripps Clinic in La Jolla.
  • Linked by freebase: (Mercy Hospital, located in, San Diego) with weight 1.0
  • Predicted by QA: (Mercy Hospital, located in, San Diego) with weight 0.72
Outline

• Introduction
• Challenges and Solution
• System Design
• Experiments
• Future Work
Future Work

• Extracted typed entities and relations can serve as additional features for QA system. Thus, potentially, QA and IE can mutually enhance each other.

• How to leverage the rank scores can be formalized as an optimization problem with systematic mathematical equations.

• Current question construction is one direction only. Two directional questions may help achieve better results with more constraints.
Reference

