Ivan Bilan

Neural Models for Relation Argument Extraction

Deep Learning for NLP (WS 17/18, LMU)

Instructor: Dr. Benjamin Roth

Argument Extraction

- Identifying relevant facts for a query entity in the sentence

i.e.: Ramon, who was killed in action, was promoted posthumously from lieutenant to captain, the military spokeswoman said.

- killed in action – per:cause_of_death
- captain – per:title

NER Bottleneck

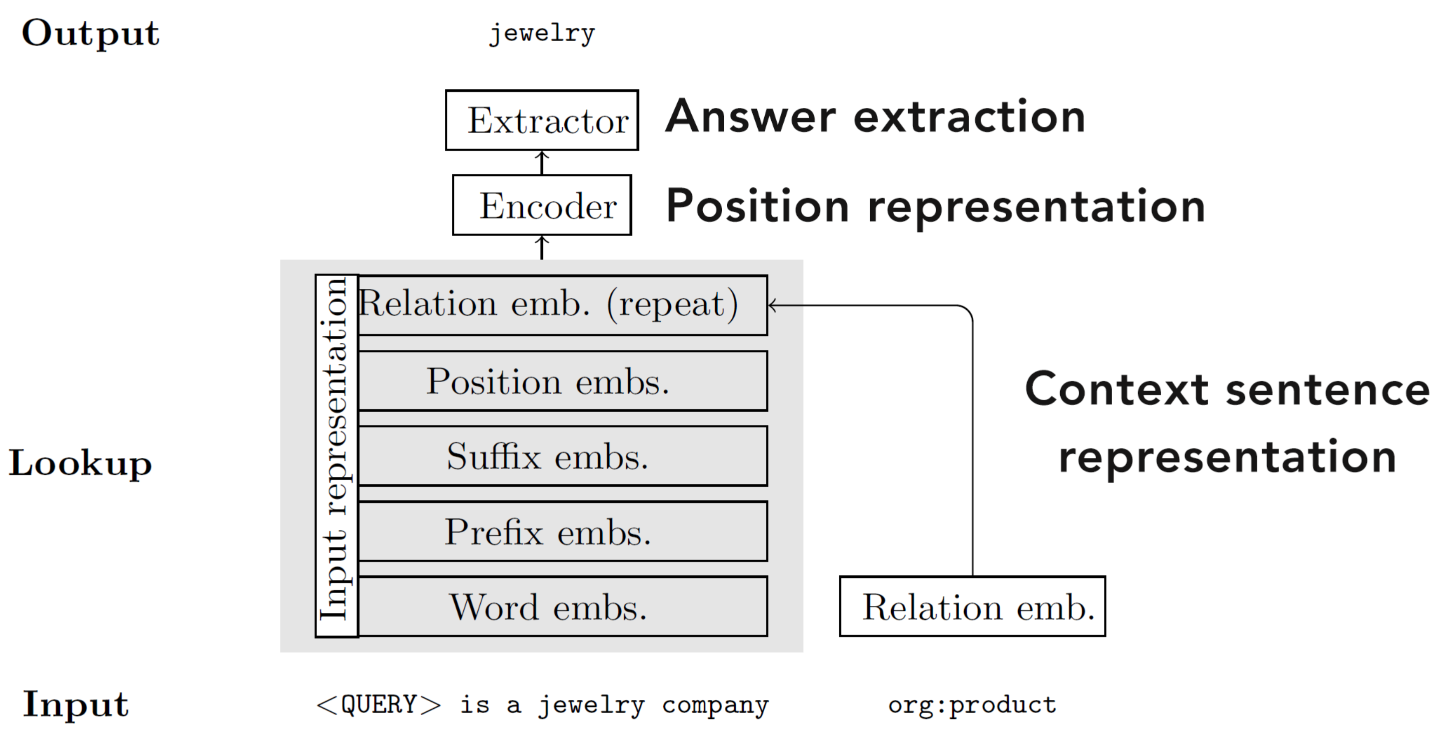
- In easy cases, standard NER approaches work well enough
- NER can bring down Recall by 30% overall:
  - miss the correct answer
  - return incorrect answer span

- **[Haig]PER attended the [US Army]ORG academy at [West Point]LOC**
- **[Haig]PER attended the [US Army academy at West Point]ORG**

Open-type Relation Argument Extraction Model

- Predicts a slot filler without using existing NER libraries given a query and a relation
- It also decides whether relation holds or not
- Achieves ~82% F-Score on a distantly supervised Wikipedia Dataset with 12 relation classes

Roth et al.



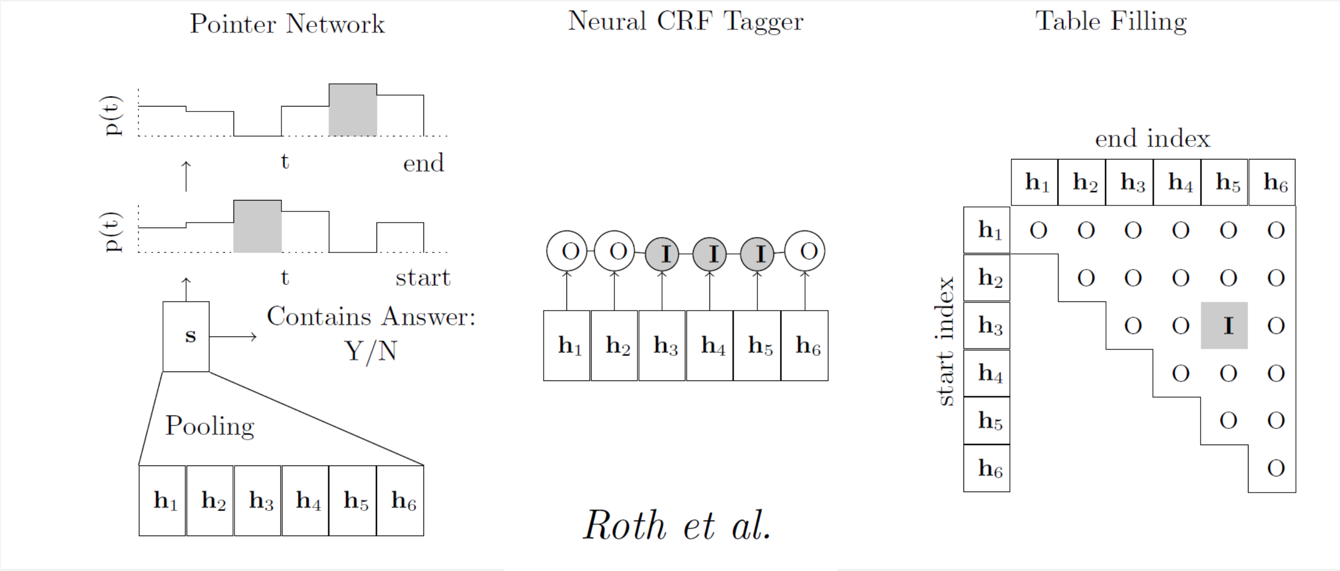
Lookup layer and architecture overview

Open-type Relation Argument Extraction Model

- Encoders (translate the output of lookup layer):
  - Bi-directional **GRU** (200 hidden units)
  - **CNN** encoder (3 layers: 32, 64, 128 filters)
  - Multi-headed **Self-Attention** (8 heads):
    - input representation of each position is used as query to computer attention score
    - using attention scores compute the weighted average of the input representations

Open-type Relation Argument Extraction Model

- Extractors (return answer span in sentence):
  - **Pointer Network** (point to start and end of argument)
  - **CRF Tagger** (predict “I” for answer, “O” for rest)
  - **Table Filling** (pairwise lookup for start and end positions)



TAC Dataset

- 120k hand annotated sentence examples
- very long sentences, containing multiple relations
- 42 relation classes (16 per:, 26 org:)
- includes NER and POS tags
- ~80% of dataset is no_relation examples
- used for relation classification
- best F-Score is 67% using Position-aware Attention

Applying Existing model on TAC

- Do both argument extraction and then relation classification
- Add 2 negative examples for each positive sample:
  - Positive:
    - She garnered **her** greatest success as a **songwriter**. (per:title)
  - Negative:
    - She garnered **her** greatest success as a songwriter. (per:spouse)
    - She garnered **her** greatest success as a songwriter. (org:headquarters)
- For each sentence in the test set, generate a sample for each known relation
  - (She garnered **her** greatest success as a songwriter.) x 42
  - If subject relation of type PER or ORG/LOC generate only related relation classes

Applying Existing model on TAC

- AIG intends to sell shares in American Life Insurance Co. **ALICO**, and AIG Edison Life Insurance Co., said Fumiyasu Sato, a Tokyo-based spokesman for AIG.

Relation	Extracted Object	Confidence Score
org:alternate_names	American Life Insurance Co	0.5846082
org:city_of_headquarters		0.0
org:founded_by		0.0
org:member_of	American Life Insurance Co	0.5126422
org:number_of_employees		0.0
org:parents	American Life Insurance Co	0.44732976
org:shareholders		0.48436856
org:subsidiaries	American Life Insurance Co	0.5975813
org:top_members/employees	**Fumiyasu Sato**	**0.8687201**

Evaluation on TAC

- 18k samples in Test Set
- Best results with Table Filling Extractor + RNN Encoder
- Argument Extraction Results (tuple-level):

Precision:	45.76 %	correctly extracted arguments / all extracted argument samples
Recall:	16.31 %	correctly extracted arguments / gold samples with arguments
F1:	24.04 %	

- Relation Classification:

Precision:	14.69 %	correctly predicted relations / false-positive predictions + gold with relations
Recall:	34.70 %	correctly predicted relations / gold samples with relations
F1:	20.64 %	

References

- “Neural architectures for open-type relation argument extraction”, Benjamin Roth, Costanza Conforti, Nina Pörner, Sanjeev Karn, Hinrich Schütze, Natural Language Engineering 1 (1): 1{37}, 2018
- “Position-aware Attention and Supervised Data Improve Slot Filling”, Yuhao Zhang, Victor Zhong, Danqi Chen, Gabor Angeli, Christopher D. Manning, Stanford University 2017