**Multi-Grained Named Entity Recognition**

Congying Xia1,5 Chenwei Zhang1 Tao Yang2 Yaliang Li3 Nan Du2 Xian Wu2 Wei Fan2 Fenglong Ma4 Philip S. Yu1,5

1University of Illinois at Chicago 2Tencent Medical AI Lab 3Alibaba Group 4University of Baffulo 5Zhijiang Lab

**BACKGROUND**

- **Named Entity Recognition** is a fundamental task in NLP
  - Input: an utterance
  - Output: identified entities, such as person names, locations, and organizations
  - Applications: syntactic parsing, question answering, relation extraction

- **Non-overlapping Named Entity Recognition**
  - Sequence Tagging models (LSTM/CNN + CRF)
  - Fail to detect nested named entities which are embedded in longer entity mentions

- **Nested Named Entity Recognition**
  - Nested NER models are designed explicitly for recognizing nested named entities
  - Nested NER models usually do not perform well on non-overlapping entities

**RESULTS**

- **Nested NER task**
  - Datasets: ACE-2004, ACE-2005

- **Non-overlapping NER task**
  - Dataset: CoNLL-2003

- **Performance on different types of sentences**
  - Split the test data in ACE-2005 into two portions: sentences with/without overlapping entities

**MGNER**

- **Detector**: detect possible entities in various granularities
  - **Word Processor**: word embedding, postage embedding, character level embedding
  - **Sentence Processor**: sentence LSTM + Elmo embedding
  - **Detection Network**: generate all possible word segments and estimate the probability of each proposal as being an entity or not

- **Classifier**: detect entity positions in various granularities
  - **Entity Processor**: context-aware entity representation with self-attention
  - **Classification Network**: classify candidates into pre-defined categories

**CONCLUSIONS**

- MGNER: state-of-the-art performance on both Nested NER and Non-overlapping NER
- High modularity and each component in MGNER can adopt a wide range of neural networks