Jue WANG

Email: zjuwangjue@gmail.com
{NLP / Information Extraction}

Education

Zhejiang University, PhD Student in Computer Science — Sep 2018 - Jun 2023(Expected)
National Scholarship 2019-2020.

Université Paris Saclay (CentraleSupélec), Double Degree — Sep 2016 - Jun 2018
Excellent Eiffel Scholarship 2016.

Zhejiang University, Bachelor in Electrical Engineering — Sep 2014 - Jun 2018
GPA: 3.89/4.00 (top 10%), The First Prize Scholarship, Excellent Student Scholarship.

Research Experience

Effective Slot Filling via Weakly-Supervised Dual-Model Learning — 2021

- Long paper at AAAI 2021. First Author.
- A weakly-supervised method for slot-filling. Supplied with very little labeled data (e.g., 100 sentences), the proposed approach significantly outperforms conventional supervised and semi-supervised methods.

Two are Better Than One: Joint Entity and Relation Extraction with Table-Sequence Encoders — 2020

- Long paper EMNLP 2020. First Author.
- A novel table-filling approach for joint extraction of entities and relations with SoTA results on ACE04, ACE05, CoNLL04, and ADE datasets. Specifically, we propose a *table-sequence encoders* architecture to learn two kinds of representations for NER and RE, and leverage the attention weights maintained by `BERT` to further improve the performance.

Pyramid: A Layered Model for Nested Named Entity Recognition — 2020

- Long paper ACL 2020. First Author.
- A novel layered model for nested NER with SoTA results on ACE04, ACE05, GENIA, and NNE datasets. The model consists of multiple neural network layers stacked in a pyramid shape. And the hidden states at layer l represent l-grams of the input text, which are labeled only if their corresponding text regions represent complete entity mentions.

EXACT: Attributed Entity Extraction by Annotating Texts — 2019

- Accepted to SIGIR 2019. Assistant Developer and Presenter.
- Propose `EXACT` for extracting attributed entities from textual documents by performing explorative annotation tasks.

Internship

SUTD - StatNLP, NLP Intern — Feb 2020 - Jun 2020

- Remote internship due to COVID-19 with Prof. Lu Wei.
- Publish a long paper at EMNLP 2020 (*Two are Better Than One: Joint Entity and Relation Extraction with Table-Sequence Encoders*).

ByteDance - ByteCamp, Camper in Algorithm Track — Aug 2019 - Sep 2019

- Multimodal Classification. Achieve 3rd place in Algorithm Track.
- A hybrid model of tensor fusion (*Tensor fusion network for multimodal sentiment analysis*), multimodal attention (*Hybrid Attention based Multimodal Network for Spoken Language Classification*), and our implemented *author gate mechanism*.

Rokid - NLP Department, NLP Intern — Jun 2018 - Dec 2018

- Use `TextCNN` to replace the original word2vec based classification module. Moreover, by pre-computing the convolution hidden states of common adjacent words, we avoids most online computation and dramatically improve the response speed. Pre-computed hidden states are stored in the hash table and updated weekly by the script.
- Implement `BERT+CRF` to replace the original NER module, with F1 0.91 => 0.96;
- Involved in `IntentTree` Project, which aims to build a domain-independent tree (with intents to be its branches) to address the problem of domain transferability. Domains are constantly changing while intents are relatively stable.
- Maintain the word embedding module (word2vec and ELMo).

Others

- **Patents**: Two innovation patents pending.
- **Languages**: English, French and Chinese.
- **Frameworks**: PyTorch, Tensorflow, Keras.