Tianyu Yu

E-Mail: yiranytianyu@gmail.com; Tel.: +86-15179283506

EDUCATION

Tsinghua University   **Beijing, China**

M.Eng. in **Computer Technology**   09/2021 - present

Beihang University   **Beijing, China**

B.Eng. in **Software Engineering**   09/2017 - 06/2021

GPA: **91.37** (Ranking: **1/165**)

AWARDS & HONORS

Nomination Award of Shenyuan Model, Beihang University (10/4000+).   12/2020

National Scholarship for Undergraduates (Top 1%) for **Twice**   11/2018 & 11/2019

Outstanding Winner of Excellent Learner Scholarship of Beihang University **(Top 1)** for **Twice**   11/2018 & 11/2019

Merit Student Award of Beihang University for **Twice**   11/2018 & 11/2019

1st Prize in Social Work Excellence Scholarship   11/2019

Publications

Tianyu Yu, Tianrui Hui, Zhihao Yu, Yue Liao, Sansi Yu, Faxi Zhang, Si Liu*. Cross-Modal Omni Interaction Modeling for Phrase Grounding. ACM MM 2020

RESEARCH EXPERIENCE

Visually Grounded Commonsense Knowledge Extraction.   2021.04~present

❖   Discover commonsense knowledge from images via VRD.

Multiple Semantic Level Entity Set Expansion   2020.11~present

❖   Introduced "negative semantic" and "negative seed entity" in entity set expansion task.

❖   Constructed two large entity set expansion datasets based on Wikipedia-en and *Harry Potter* series.

❖   Devised an algorithm to take use of negative semantic information from input, which improved the expansion accurary on multiple semantic level.

Sentence-Level Pretraining for Document-Level RE   2020.07~2021.03

❖   Designed Sentence-Level ELECTRA pretraining task to learn better sentence representation from large corpus.

❖   Designed Entity Context Prediction pretraining task to learn better entity representation from large corpus.

* Improved F1 score on DocRED comparing to BERT-based baseline by fine-tuning using devised tasks.

Cross-Modal Omni Interaction Modeling for Phrase Grounding   2019.10~2020.07

* Adressed the phrase grounding accuracy problem as the primary researcher.
* Devised a novel model to capture complex spatial and semantic relationship among image regions and phrases through multi-level multi-modal interaction.
* The new method improved the grounding accuracy by 6.15% on Flickr30K Entities and 21.25% on ReferItGame.

Joint Extraction of Relation and Evidence for Document-Level RE   2019.10~2020.05

* Jointly extract evidence and relation in RE task to help both evidence extraction and relation extraction.
* Devised a bilinear based entity feature extraction module to improve the relation extraction performance.
* Usinig Deep Sets to ensure the permutation invariance of evidences and using beam search to improve efficiency.
* Improved relation extraction F1 score on DocRED using a novel evidence-guided-attention mechanism.

Identification of Real-Time Requirements in Natural Language Specification   2018.04~2019.07

* Built the web crawler to collect sufficient industrial documents of software engineering requirements, translated them into structured data, and labelled the data manually according to the nature of the requirements.
* Achieved feature extraction from data using the models of TF-IDF and word2vec+BoW, and customized the stop-word list with respect to the predefined keyword set of real-time requirements.
* Tested and compared the performance of different classifiers on the collected dataset, and further assessed the results comparing to other research work in terms of the dataset, performance, classification approach and so on.

Research of the Sentiment Analysis of Chinese Sentences   2018.09

* Proposed to apply LSTM with CNN to the sentiment analysis of Chinese sentences, established the model, and completed the training with datasets like COAE.
* Developed the front- and back-end of the website based on Django and Bootstrap for users to input texts and files with prediction results visualized in different colors.
* Implemented the function of feedback to collect users' reports about the accuracy of prediction for future incremental learning.

OTHER INFORMATION

Computer Skills: C, C++, Python, Java, SQL, Pytorch, Tensorflow, Scipy

Languages: Mandarin (Native), English (Proficient)