Zhihong Shao

E-mail: zhshaothu@gmail.com
Web: https://ZhihongShao.github.io

Research Interests

My interests are in natural language processing and deep learning. I am particularly interested in how we can build a robust and scalable AI system that can leverage diverse skills (e.g., retrieval, logical reasoning, and numerical reasoning) to aggregate possibly-heterogeneous information and answer natural language questions precisely regardless of their complexity. My recent work focused on (i) comprehensive exploitation of a large-scale corpus for multi-answer question answering [1], (ii) improving the reliability of (weakly-supervised) neuro-symbolic models for complex discrete reasoning [3], (iii) robust classification and text matching (e.g., paraphrase detection and natural language inference) [2].

Education

Tsinghua University, Beijing, China — September 2019 - Present
Ph.D. Student, Computer Science and Technology
Advisor: Minlie Huang

Beihang University, Beijing, China — September 2015 – July 2019
B.E., Computer Science and Technology
GPA: 3.86/4, *Rank:* 2/213

Publications

[1] Answering Open-Domain Multi-Answer Questions via a Recall-then-Verify Framework
Zhihong Shao, and Minlie Huang
Annual Meeting of the Association for Computational Linguistics (ACL), 2022.
(Best QA system on the AmbigNQ leaderboard)

[2] AdvExpander: Generating Natural Language Adversarial Examples by Expanding Text
Zhihong Shao, Zhongqin Wu, and Minlie Huang
IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP), vol. 30, pp. 1184-1196, 2022.

[3] A Mutual Information Maximization Approach for the Spurious Solution Problem in Weakly Supervised Question Answering
Zhihong Shao, Lifeng Shang, Qun Liu, and Minlie Huang
Annual Meeting of the Association for Computational Linguistics (ACL), 2021.

[4] Long and Diverse Text Generation with Planning-based Hierarchical Variational Model
Zhihong Shao, Minlie Huang, Jiangtao Wen, Wenfei Xu, and Xiaoyan Zhu
Empirical Methods in Natural Language Processing (EMNLP), 2019.

Preprint

[5] CoTK: An Open-Source Toolkit for Fast Development and Fair Evaluation of Text Generation
Fei Huang, Dazhen Wan, **Zhihong Shao**, Pei Ke, Jian Guan, Yilin Niu, Xiaoyan Zhu, and Minlie Huang
Arxiv abs/2002.00583, 2020.

Research Experience

CoAI Lab, Tsinghua University — Sep 2019-Present, Beijing, China
Ph.D. Student (Supervisor: Minlie Huang)
Worked on open-domain multi-answer question answering [1], weakly-supervised training for neuro-reson reasoning [3], robust classification and text matching [2], and data-to-text generation [4].

Huawei Noah's Ark Lab Jun 2020-Oct 2020, Shenzhen, China
Research Intern (Supervisor: Lifeng Shang)
Worked on neuro-symbolic reasoning [3] under a weakly-supervised setting, where we selected high-quality symbolic reasoning processes for training via mutual information maximization.

Awards

2nd Prize, Comprehensive Scholarship, Tsinghua University 2021
3rd Prize, the National Final of "LAN QIAO CUP" C/C++ Group 2018
China National Scholarship 2017
1st Prize, National College Students Mathematics Competition (non-math-major) 2016
China National Scholarship 2016

Services

Reviewer/Program Committee: ACL, EMNLP, NLPCC, ARR

Teaching Assistant

Artificial Neural Network 2019 Fall, 2020 Fall, 2021 Fall
Instructor: Minlie Huang

Object-Oriented Programming 2020 Spring, 2021 Spring, 2022 Spring
Instructor: Minlie Huang
Also gave guest lectures and made assignments