RANRAN HAORAN ZHANG

windchimeran.github.io @ hzz5361@psu.edu
scholar.google.com/citations?user=aDqdjcUAAAAJ&hl

RESEARCH INTEREST

My research interests lie in natural language processing. My research aims to build a universal framework that can **extract and deduce logical information from heterogeneous textual data**.

- Information Extraction
- Executable Semantic Parsing
- Indirect Supervision
- Education x NLP

EDUCATION

Penn State University (PSU) — Jan 2022 – Present
Ph.D. Computer Science — **Advisor: Dr. Rui Zhang**

University of Illinois at Urbana-Champaign (UIUC) — Aug 2019 – May 2021
M.S. Information Management — **Advisor: Dr. Heng Ji**

Changsha University of Science & Technology (CSUST) — Jul 2014 – Jun 2018
B.S. Computer Science — **Advisor: Dr. Daojian Zeng**

PUBLICATIONS AND MANUSCRIPTS

(* refers to equal contributions)

Under Review'22 — ***ConEntail : An Entailment-based Framework for Universal Zero and Few Shot Classification with Supervised Contrastive Pretraining*** Ranran Haoran Zhang, Aysa Xuemo Fan, Rui Zhang.

NAACL'21 Demo — ***COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation***. Qingyun Wang, Manling Li, Xuan Wang, Nikolaus Parulian, Guangxing Han, Jiawei Ma, Jingxuan Tu, Ying Lin, Ranran Haoran Zhang, Weili Liu, Aabhas Chauhan, Yingjun Guan, Bangzheng Li, Ruisong Li, Xiangchen Song, Heng Ji, Jiawei Han, Shih-Fu Chang, James Pustejovsky, David Liem, Ahmed Elsayed, Martha Palmer, Jasmine Rah, Cynthia Schneider, Boyan Onyshkevych. Retrieved from here.

EMNLP'20-Findings — ***Minimize Exposure Bias of Seq2Seq Models in Joint Entity and Relation Extraction.*** Ranran Haoran Zhang*, Qianying Liu*, Aysa Xuemo Fan, Heng Ji, Daojian Zeng, Fei Cheng, Daisuke Kawahara and Sadao Kurohashi. Retrieved from here.

AAAI'20 — ***CopyMTL: Copy Mechanism for Joint Extraction of Entities and Relations with Multi-Task Learning.*** Daojian Zeng*, Ranran Haoran Zhang*, Qianying Liu. Retrieved from here.

IEEE Access'19 — ***User-oriented paraphrase generation with keywords controlled network.*** Daojian Zeng, Ranran Haoran Zhang, Lingyun Xiang, Jin Wang, Guoliang Ji. Retrieved from here.

INTERNSHIP

Chatbot for Alzheimer's Caregiver.

Research Intern @ Benten Tech

Aug 2021 - Nov 2021

- Information Retrieval based FAQ module.
- Neural based chit chat module with long-term memory.
- Deployed on Alexa Skill.

PROJECT

Universal NLP Classifier

Research Intern @ PSU

May 2021 - Present

- Proposed an indirect supervision meta task to solve all NLP classification tasks.
- Integrated heterogeneous labeled textual data for a large-scale intermediate pretraining.
- Provided new insights about generation-vs-classification framework of pretrained language models.

Attribution Verification of News Articles

Team Member, Researcher @ UIUC BLENDER Lab

Jan 2021 - Mar 2021

- Verified whether or not the news article is from the purported source.
- Extracted features from both news text and news images.
- Integrated my part to the team repository via docker.

COVID-19 Knowledge Graph

Team Member, Researcher @ UIUC BLENDER Lab

Jun 2020 - Sep 2020

- Built COVID-KG by information extraction.
- Generated drug repurposing report based on COVID-KG automatically.
- Released the KG and visualization tools to help medical research.

Sequence-to-Unordered-Multi-Tree for Joint Extraction of Relations and Entities

Team Leader, Researcher @ UIUC BLENDER Lab

Nov 2019 - Mar 2020

- Formulated the output sequence to unordered-multi-tree structure to minimize the model bias of the Seq2Seq model in relation extraction.
- Designed AB-Test to disentangle model bias from data bias.
- Implemented a toolkit containing [5 Models $\times$ 2 Datasets] to be open-sourced.

Sequence-to-Sequence for Joint Extraction of Relations and Entities

Team Leader, Researcher @ CSUST AI Lab

Mar 2019 - Seq 2019

- Figured out a linear algebra bug causing underfitting of training set in an ACL2018 paper.
- Based on theoretical analysis, added only one more non-linear layer to fix the bug.
- Yielded 14 and 31 (F1) absolute improvement over baseline on NYT and WebNLG dataset respectively.

SKILLS

Programming	Python, Haskell, C++.
Writting	LaTeX, Markdown, HTML.
Data	Data cleaning, data exploration, data visualization.
AI	Deep learning, natural language processing, computer vision.

TEACHING EXPERIENCE

- CMPSC 448 Machine Learning, Fall 2022, Penn State University
- CSE 597 Deep Learning for NLP, Spring 2022, Penn State University

COMMUNITY

- Reviewer of JMLC'21, COLING'22 and EMNLP'22.
- Gained 300+ stars for open-sourced projects on GitHub.
- Co-founded the first machine learning lab in CSUST.

AWARDS

- Best Demo Paper, NAACL'21 — Jun 2021
- Best Thesis Award, CSUST — Jun 2018
- Dean's Scholarship for 2015-2017 Academic Years, CSUST — 2015 - 2017