Zijie Wang

zijiewang@arizona.edu | wang-zijie.github.io/

Education

University of Arizona — Tucson, AZ, USA
Doctor of Philosophy, Computer Science — *Aug. 2022 – Present*
- Advisor: Dr. Eduardo Blanco

Arizona State University (Transferred with advisor) — Tempe, AZ, USA
Doctor of Philosophy, Computer Science — *Jan. 2022 – Aug. 2022*
- Advisor: Dr. Eduardo Blanco

Arizona State University — Tempe, AZ, USA
Master of Science, Computer Science — *Aug. 2019 – Dec. 2021*
- Thesis: Examining Data Integration with Schema Changes Based on Cell-level Mapping Using Deep Learning Models

Beijing Information Science & Technology University — Beijing, China
Bachelor of Engineering, Computer Science and Technology — *Sept. 2015 – June 2019*

Research Interests

Natural language understanding, natural language generation, open-domain question answering, multilingual understanding, artificial intelligence, and machine learning.

Publications

1. **Zijie Wang**, Md Mosharaf Hossain, Shivam Mathur, Terry Cruz Melo, Kadir Bulut Ozler, Keun Hee Park, Jacob Quintero, MohammadHossein Rezaei, Shreya Nupur Shakya, Md Nayem Uddin, and Eduardo Blanco. Interpreting indirect answers to yes-no questions in multiple languages. In *Findings of the Association for Computational Linguistics: **EMNLP 2023***, Singapore, December 2023. Association for Computational Linguistics (Accepted, to appear)

2. Lixi Zhou, Arindam Jain, **Zijie Wang**, Amitabh Das, Yingzhen Yang, and Jia Zou. Benchmark of dnn model search at deployment time. In *Proceedings of the 34th International Conference on Scientific and Statistical Database Management*, **SSDBM '22**, New York, NY, USA, 2022. Association for Computing Machinery

3. **Zijie Wang**, Lixi Zhou, and Jia Zou. Integration of fast-evolving data sources using a deep learning approach. In ***Software Foundations for Data Interoperability and Large Scale Graph Data Analytics***, pages 172–186. Springer, 2020

Preprints

1. **Zijie Wang**, Lixi Zhou, Amitabh Das, Valay Dave, Zhanpeng Jin, and Jia Zou. Survive the schema changes: Integration of unmanaged data using deep learning. *arXiv preprint arXiv:2010.07586*, 2020

Experience

Research Assistant Aug. 2022 – Present
University of Arizona *Tucson, AZ, USA*
- Supervisor: Dr. Eduardo Blanco

Research Assistant May 2020 - Aug. 2021
Arizona State University *Tempe, AZ, USA*
- Supervisor: Dr. Jia Zou

Research Projects

Detecting and Answering Questions with False Assumptions June 2022 - Present
- Ongoing research project. Existing language models lack the ability to detect factually wrong assumptions in questions and generate completely wrong answers. We develop the largest dataset including questions with false assumptions. We show that fine-tuning LMs with our dataset is beneficial in detecting questions with false assumptions and generating more reasonable answers to these questions.
- **Expected one paper submission to a top conference**

Understanding Indirect Answers to Yes-No Questions Jan. 2022 - June 2023
- People often skip polar keywords when answering yes-no questions. We investigate interpreting indirect answers to yes-no questions in a multi-domain (three domains) and multilingual setting (nine languages). We show that domain adaptation and cross-lingual learning are beneficial.
- **First-author paper accepted to EMNLP 2023 Findings [1]**
- **One paper submitted to a top conference**

Integration of Unmanaged Data Using Deep Learning June 2020 - June 2021
- Data schema change is the main obstacle to the automation of the data integration process. We propose an automatic data integration pipeline based on deep learning models (Bi-LSTM and BERT). We show our pipeline is effective for various real-world scenarios.
- **First-author paper published [3]**

Awards

Outstanding undergraduate student, Beijing City (5 in 300) June 2019

Best Graduation Project (20 in 300) May 2019

Academic Excellent Scholarship (20 in 300) Nov. 2016, 2017, 2018

Skills

Programming Language: Python, Java, Perl, C++, Scala, SQL

Machine Learning & NLP Tools: PyTorch, TensorFlow, Keras; NLTK, spaCy.

Deep Learning Models: (Bi-)LSTM, (m)BERT, (XLM-)RoBERTa, BART, T5, GPT-X, Alpaca, LLaMa, etc.

Services

- Conference review: ACL 2023, EMNLP 2023, AACL 2023
- Journal review: ACM TALLIP (May 2022)