

Zijie Wang

+1-404-6831781 | zijiewang@asu.edu | [linkedin.com/in/zijie-wang-23a133185](https://www.linkedin.com/in/zijie-wang-23a133185)

EDUCATION

Arizona State University

Doctor of Philosophy, Computer Science

- Advisor: Dr. Eduardo Blanco

Tempe, AZ, USA

Jan. 2022 – Present

Arizona State University

Master of Science, Computer Science

- With thesis track; Advisor: Dr. Jia Zou
- Thesis: Examining Data Integration with Schema Changes Based on Cell-level Mapping Using Deep Learning Models

Tempe, AZ, USA

Sept. 2019 – Dec. 2021

Beijing Information Science & Technology University

Bachelor of Engineering, Computer Science and Technology

- Graduate Project: Recognizing Verification Image Using Convolutional Neural Network

Beijing, China

Sept. 2015 – June 2019

EXPERIENCE

Research Assistant

Arizona State University

- Supervisor: Dr. Eduardo Blanco
- Research Area: Natural Language Processing

Jan. 2022 – Present

Tempe, AZ, USA

Research Assistant

Arizona State University

- Supervisor: Dr. Jia Zou
- Investigating applying deep learning to database systems

May 2020 – June 2021

Tempe, AZ, USA

RESEARCH TOPICS

Understand Indirect Answers to Yes-No Questions

- Ongoing research project
- Investigate the method of determining underlying indirect answers to yes-no questions
- Propose a novel Language Model pretraining objectives using context)

Jan. 2022 – Present

Integration of Unmanaged Data Using Deep Learning

- Investigated the problem of using deep learning to automatically integrate data with schema evolution
- Designed the experiment with three different deep learning models (Bi-LSTM, Transformer and BERT)
- Tested on various real-world scenarios (COVID-19 dataset and machine log dataset)
- **One paper published**

June 2020 – June 2021

Finding Related Models for Serving Based on Dynamic LSH

- Investigated a method to efficiently evaluate the adaptivity of models among different datasets without duplicate training
- Proposed several new measurement methods that outperformed the state-of-the-art similarity based metrics
- **One paper published**

July 2020 – Oct. 2020

PROJECTS

Referring Expression Comprehension on Heterogeneous Datasets Sept. 2020 – Dec. 2020

- Language & Platform: Python; TensorFlow, PyTorch
- Localizing target objects in images that also described in natural language
- Developed functions to automatically generate synthetic dataset
- Reproduced two state-of-the-art models (VL-BERT and ViLBERT) and inferenced on the synthetic dataset
- Contribution: Collected dataset samples and developed the experiments on VL-BERT model

Domain Adaptation for Fake News Detection Oct. 2019 – Dec. 2019

- Language & Platform: Python; TensorFlow
- Learned a domain-independent model that can be applied to different datasets
- Model transferring from domain dataset (gossipcop) to target dataset (politifact)
- Implemented an Auto-Encoder model for pretraining and a LSTM model for classification tasks
- Contribution: Worked on feature extraction and model pretraining

Verification Image Recognition based on Convolutional Neural Network Feb. 2019 – May 2019

- Language & Platform: Python, Java; Keras
- Designed a system to recognize a type of verification image (CAPTCHA)
- Developed algorithms to preprocess (denoise, grayscale and binaryzation) and segment image
- Implemented Convolutional Neural Network to classify segmented images with a single letter or number
- Contribution: An individual project; **Winner of the best graduate project award**

AWARDS

Outstanding Graduate, Beijing City (5 in 300)	June 2019
Best Graduate Project (20 in 300)	June 2019
Academic Excellent Scholarship (20 in 300)	Nov. 2016, 2017, 2018

SKILLS

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

Database & Machine Learning: Spark, MySQL, PostgreSQL; Tensorflow, PyTorch, Keras

Platform: Ubuntu, Android, AWS EC2

RELATED COURSES

Course taken at ASU: Social Media Mining (CSE 472), Natural Language Processing (CSE 576), Artificial Intelligence (CSE 571), Statistical Machine Learning (CSE 575), Data Visualization (CSE 578), Data-Intensive Systems for Machine Learning (CSE 598)

PUBLICATIONS

1. **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
2. Lixi Zhou, Arindam Jain, **Zijie Wang**, Amitabh Das, Yingzhen Yang, and Jia Zou. Benchmark of dnn model search at deployment time. *arXiv preprint arXiv:2206.00188*, 2022 (To appear in SSDBM 2022)

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