

Zhiyu Xie

Tsinghua University – Beijing, China

☎ +86 18030308900 • ✉ xiezhiyu32768@gmail.com

🌐 xiezhiyu01.github.io

Education

Tsinghua University

Beijing, China

2019.8 - Present

- Bachelor of Engineering in Computer Science and Technology
- GPA: 3.96/4.00 (Ranking: 3/202)
- Core Courses: Linear Algebra(A+), Probability and Statistics(A), Data Structures(A), Operating Systems(A), Introduction of Theory of Computation(A), Introduction to Artificial Intelligence(A)

Honors and Awards

- Academic Excellence Scholarship in Tsinghua University (Top 1%) 3 times in 2020,2021,2022
- Google Women Techmakers Scholarship (34 winners in China) 2020
- Freshman Scholarship in Tsinghua University (top 10 in Fujian Province) 2019
- FuGuang Scholarship in Fujian Province 2019
- The Bronze Medal in National Olympiad in Informatics (NOI) 2017
- The Bronze Medal in the 11th Asia-Pacific Informatics Olympiad (APIO) 2017

Internship

Model Distillation in Model Maker

TensorFlow Lite, Google Beijing

STEP Intern, Mentor: [Tian Lin](#)

2021.7 - 2021.9

- To further facilitate TensorFlow Lite on-device training, this work introduced model distillation method for model compression in *TensorFlow Lite Model Maker library*.
- Followed Object-oriented programming principles to implement model distillation for image, audio and text classification tasks in Model Maker Framework, enabling users to create a fine-tuned end-to-end model on a customized dataset in just **6 lines of code**.
- Carried out comprehensive experiments to show that the work can lead to a **90% reduction of parameters** while maintaining a competitive and sometimes better accuracy performance compared to the teacher network.

Research Experience

Prefix Generator for Low-resource Event Extraction

PlusLab, UCLA

Advisor: [Nanyun Peng](#)

2022.6 - Present

- **Incorporated useful external information**, such as syntax trees, into generative models, which have been important for event extraction tasks because of their flexibility and efficiency.
- Proposed **Prefix Generator**, which encodes external information (e.g. Abstract Meaning Representation Graph, Optimus robust representation) by pre-training and mapping representations into prefixes in encoder-decoder models.
- Our method proved to be generally applicable and especially **effective in low-resource settings**.
- This work will submit to *ACL 2023*.

Domain Relabeling for Subpopulation Shift

IRIS Lab, Stanford

Advisor: [Huaxiu Yao](#)

2022.6 - Present

- Analyzed spurious correlations caused by poor quality domain labels in current approaches to address subpopulation shifts between training and testing distributions.
- Proposed integrating the **candidate features in datasets' metadata table** to get higher-quality domain labels.
- Built **reinforcement learning framework** that utilizes downstream task performance feedback of each metadata feature to optimize the domain labels iteratively.
- Our method leads to **improved worst-group performance** in real-world datasets covering fields such as **healthcare and weather forecasting**.

Reviewing Test Protocols of Distantly Supervised Relation Extraction

THUNLP, Tsinghua

Advisor: [Zhiyuan Liu](#)

2020.9 - 2021.3

- Examined two popular relation extraction datasets (NYT10 and Wiki20) for **annotation errors due to distant supervision** methods adopted.
- Proposed an **improved relation ontology** and adopted data-cleaning, constructed manually-annotated test sets for NYT10 and Wiki20, **correcting 53% wrong labels** in NYT10.
- Analyzed performance differences of competitive models on manually-annotated and distantly supervised datasets.
- Our conclusion sheds light on the importance of a more accurate evaluation for relation extraction research.

Publication

Manual Evaluation Matters: Reviewing Test Protocols of Distantly Supervised Relation Extraction

- Tianyu Gao, Xu Han, Keyue Qiu, Yuzhuo Bai, **Zhiyu Xie**, Yankai Lin, Zhiyuan Liu, Peng Li, Maosong Sun, Jie Zhou.
- In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*. **ACL Findings 2021**.

Selected Projects

🌀 RISC-V CPU

Fall 2021

- Designed and implemented a 5-stage pipelined RISC-V CPU in Verilog.
- Enabled interrupt and exception handling, virtual memory system, Translation Lookaside Buffer (TLB), branch prediction, and Video Graphics Array (VGA) interface.

🌀 Education Knowledge Graph APP

Summer 2020

- Developed an android application providing high-quality education services based on knowledge graphs.
- Enabled users to search, learn, bookmark, and share entries in various categories.

🌀 Ray Tracing

Spring 2020

- Implemented ray tracing and progressive photon mapping algorithm.
- Added soft shadow, anti-aliasing, depth of field, and texture to the algorithm.
- Accelerated the calculation by algorithm (e.g. KD tree, octree) and parallel computing.

Activities

Volunteer in *Program Buddy*

Fall 2021, Fall 2022

- Volunteered to assist in teaching undergraduate students learning C++ and Python.

Organizer of *Trigger Salon*

Fall 2020

- Planned and organized the salon to **encourage the diversified development** of undergraduate students and discover their inner drive.
- The salon was widely acclaimed with **over 1,500 feed reads**, and continues to operate today.

Lecturer in *Amoy Youth Elite Summit*

Summer 2020

- Selected 70 outstanding high school students in Fujian Province to attend the summit.
- Gave lectures on **set theory and computer science** in small class sizes, organized **QA sessions** after class.

Skills

- Programming Language: C++, Python, Java, HTML+CSS, RISC-V, Verilog
- Language Skills: Mandarin (Native), English (Fluent, TOEFL: 112(S25), GRE: V158+Q170+4.0)