Zhen Zhang

+1 (805) 570 5808 | Homepage | Google Scholar | zhen_zhang@ucsb.edu

Education

University of California, Santa Barbara (UCSB) *Sept. 2023 – Present*

PhD in Computer Science, Advised by Prof. Zheng Zhang and Prof. Xin (Eric) Wang *California, USA*

- **Research Interests:** Training and Inference Efficiency, LLMs
- **Teaching Assistant:** CS 130A Data Structures and Algorithms, CS 8 Introduction to Computer Science

Tsinghua University *Sept. 2019 – June 2023*

B.Sc. in Physics *Beijing, China*

- **Awards & Honors:** Academic Excellence Scholarship (2021, 2022)

Publications

* indicates equal contribution

- **Zhen Zhang**, Yifan Yang, Kai Zhen, Nathan Susanj, Athanasios Mouchtaris, Siegfried Kunzmann, Zheng Zhang. "MaZO: Masked Zeroth-Order Optimization for Multi-Task Fine-Tuning of Large Language Models" **(Arxiv)**
- **Zhen Zhang**, Guanhua Zhang, Bairu Hou, Wenqi Fan, Qing Li, Sijia Liu, Yang Zhang, Shiyu Chang. "Certified Robustness for Large Language Models with Self-Denoising" **(NAACL 2024)**
- **Zhen Zhang**, Jialu Wang, Xin Eric Wang. "Parameter-Efficient Cross-lingual Transfer Learning of Vision and Language Models via Translation-based Alignment" **(EMNLP 2023)**
- Yujia Qin, Shengding Hu, ...,**Zhen Zhang**, ..., Tongshuang Wu, Heng Ji, Zhiyuan Liu, Maosong Sun. "Parameter-Efficient Cross-lingual Transfer Learning of Vision and Language Models via Translation-based Alignment" **(ACM Computing Surveys, CSUR)**
- Shengding Hu, Ning Ding, Weilin Zhao, Xingtai Lv, **Zhen Zhang**, Zhiyuan Liu, Maosong Sun. "OpenDelta: A Plug-and-play Library for Parameter-efficient Adaptation of Pre-trained Models" **(ACL 2023)**
- Shengding Hu*, **Zhen Zhang***, Ning Ding, Yadao Wang, Yasheng Wang, Zhiyuan Liu, Maosong Sun. "Sparse Structure Search for Parameter-Efficient Tuning" **(NeurIPS 2022)**

Research Experience

University of California, Santa Barbara *Feb. 2025 – Now*

Research Assistant, advised by Xin Eric Wang

- Explored novel methods to reduce resource consumption in reasoning models, focusing on improving efficiency by compressing token information. Developed innovative techniques to optimize token representations while maintaining model performance.

University of California, Santa Barbara *Oct. 2024 – Feb. 2025*

Research Assistant, advised by Zheng Zhang

- Developed a novel multi-task fine-tuning framework for LLMs under Zeroth-Order Optimization (MaZO), addressing gradient variance and task conflicts through parameter-level masking. Achieved state-of-the-art performance on LLMs and outperformed first-order multi-task learning methods.

University of California, Santa Barbara *June 2023 – Oct. 2023*

Research Assistant, advised by Shiyu Chang

- Proposed a self-denoising method leveraging LLMs' intrinsic capabilities, improving inference accuracy on noisy inputs while maintaining certified robustness. The method achieved superior robustness and 5% higher empirical accuracy on LLMs like Alpaca. This paper is published in NAACL 2024.

Tsinghua University, THUNLP Group *Feb. 2023 – June 2023*

Undergraduate Thesis, advised by Zhiyuan Liu

- Applied RLHF to a 10B Chinese pre-trained models (CPM-Bee), enhancing helpfulness and reducing harmfulness. Empower the model with instruction following ability while teaching it to refuse to answer harmful information. Evaluated the model on the HH-RLHF dataset and ZeroCLUE benchmarks (top 10).

University of California, Santa Cruz *July 2022 – Feb. 2023*

Research Intern, advised by Xin Eric Wang

- Designed a parameter-efficient framework to reduce language gaps in Multilingual-CLIP, achieving significant performance gain with training and memory efficiency. This paper is published in EMNLP 2023.

Tsinghua University, THUNLP Group *Jan. 2022 – Sept. 2022*

Research Assistant, advised by Zhiyuan Liu

- Participated in developing OpenDelta, a PEFT toolkit supporting diverse models and tasks. Published in ACL 2023 Demo Track.
- Proposed a NAS-based method to automatically search for a sparse PEFT combination, achieving 99% performance of full model fine-tuning with 0.01% additional parameters on T5-large. This paper is published in NeurIPS 2022.

Internship Experience

Microsoft STCA, WebXT-Search&Distribution *April 2024 – Sep. 2024*

MLE, advised by Gong Ming and Wenbiao Ding

- Optimized Bing Search models to generate TL;DR in RAG scenarios. Designed and implemented pipelines for high-quality training data generation. Applied DPO to empower LLMs to reject inappropriate or unanswerable user queries when search results were insufficient, ensuring response reliability and minimizing hallucination risks.

Skills & Other Awards

Program Languages: Python, LateX, C/C++, SQL, Mathematica

Others: PyTorch, Linux, Git

Conference Reviewer: ACL 2023, NAACL 2024

Awards: Silver Medal in the 35th Chinese Physics Olympiad