Yuhang Zhou

email: tonyzhou@umd.edu
734-882-8851

RESEARCH INTERESTS Natural Language Processing, Large Language Model Fine-tuning, Computational Social Science

EDUCATION

University of Maryland, College Park, the U.S.

- *Ph.D. in Information Studies*
  - Coursework: Natural Language Understanding, Computational Linguistics — GPA: 3.98/4.0

University of Michigan, Ann Arbor, the U.S.

- *Bachelor of Science in Data Science* — May 2020
  - Coursework: Deep Learning for Vision, Data Mining, Database Management System — GPA: 3.90/4.0

Shanghai Jiao Tong University, Shanghai, China

- *Bachelor of Science in Electrical and Computer Engineering* — Aug 2020
  - Coursework: Data Mining and Machine Learning, Honors Mathematics, Discrete Mathematics — GPA: 3.56/4.0

RESEARCH

University of Maryland

- *Concept Spurious Correlation in LLMs, advised by Prof. Furong Huang*
  - Explored the spurious correlation at the concept level for large language model fine-tuning or in-context learning.
  - Utilized the advanced LLMs to generate the "counterfactual" sentences to resolve the spurious correlation.
  - Data augmentation method can bring 1.44% relative improvement on classification benchmarks and effectively mitigate performance bias.
- *LLM Long-tailed Knowledge Distillation, advised by Prof. Furong Huang*
  - Developed a framework with active learning and data augmentation to resolve the challenges of LLM knowledge distillation on long-tailed datasets.
  - Proposed framework can bring about a relative improvement of 6. 81% on long-tailed datasets.
- *Image Sequence Benchmarks, advised by Prof. Furong Huang*
  - Constructed a benchmark for image sequences with varying lengths and assessed the performance of current vision-language models on developed benchmarks.
  - Analyzed the unsuccessful cases and identified three failure reasons for current vision-language models.

University of Maryland, CLIP Lab

- *Emoji Diffusion on the Social Networks, advised by Prof. Wei Ai*
  - Explored the diffusion patterns of newly created emojis on social media on the frequency and semantic aspects.
  - Discovered the negative effect of new emojis on the sentiment classification task of language models.
  - Proposed a emoji substitution method to resolve out-of-vocabulary challenges and lead to a 6.92% relative improvement on tweets with new emojis.

SELECTED PUBLICATION

* denotes equal contribution. See full publications in Google Scholar

- **Zhou, Yuhang***, Jing Zhu*, Paiheng Xu, Xiaoyu Liu, Xiyao Wang, Danai Koutra, Wei Ai, and Furong Huang. Multi-stage balanced distillation: Addressing long-tail challenges in sequence-level knowledge distillation. *Findings of EMNLP 2024*, 2024
- **Zhou, Yuhang**, Paiheng Xu, Xiaoyu Liu, Bang An, Wei Ai, and Furong Huang. Explore spurious correlations at the concept level in language models for text classification. *ACL 2024*, 2024
- **Zhou, Yuhang** and Wei Ai. Teaching-assistant-in-the-loop: Improving knowledge distillation from imperfect teacher models in low-budget scenarios. *Findings of ACL 2024*, 2024
- Xiyao Wang, **Zhou, Yuhang**, Xiaoyu Liu, Hongjin Lu, Yuancheng Xu, Feihong He, Jaehong Yoon, Taixi Lu, Gedas Bertasius, Mohit Bansal, et al. Mementos: A comprehensive benchmark for multimodal large language model reasoning over image sequences. *ACL 2024*, 2024
- Jing Zhu*, **Zhou, Yuhang***, Vassilis N Ioannidis, Shengyi Qian, Wei Ai, Xiang Song, and Danai Koutra. Pitfalls in link prediction with graph neural networks: Understanding the impact of target-link inclusion & better practices. In *ACM International Conference on Web Search and Data Mining (WSDM)*, 2024
- **Zhou, Yuhang***, Suraj Maharjan*, and Beiye Liu. Scalable prompt generation for semi-supervised learning with language models. In *Findings of EACL 2023*, 2023

INTERNSHIP **Amazon AGI** , New York, USA

- *Applied Scientist Intern, Manager: Giannis Karamanolakis* May 2024-Aug 2024
  - Designed a novel MoE (Mixture of Experts) architecture to incorporate heterogeneous expert models.
  - Proposed multiple routing heuristics to improve MoE performance when merging without fine-tuning.

Amazon Alexa AI, Boston, USA

- *Applied Scientist Intern, Manager: Yuguang Yue* May 2023-Aug 2023
  - Designed a unified large language model knowledge distillation framework with student model signals to improve the student model fine-tuning performance.
  - Conducted extensive experiments on datasets and multiple public models with the corresponding promoting methods.
  - Proposed two-stage framework brings a relative improvement of up to 20.79% compared to random fine-tuning (paper published in *Findings of ACL, 2024*)

Amazon Alexa AI, New York, USA

- *Applied Scientist Intern, Manager: Beiye Liu* May 2022-Aug 2022
  - Designed a new prompt-tuning semi-supervised learning pipeline for large language models without manual prompts and verbalizers to promote the scalable text classification.
  - Reproduced the state-of-the-art few-shot prompt-tuning methods as the experiment baseline results.
  - Proposed method brings a relative improvement of 2.52% over the previous method with manual prompts and verbalizers (paper published in *Findings of EACL, 2023*)

Alibaba Group, Hangzhou, China

- *Research Intern, Manager: Tianyu Li* May 2021-Aug 2021
  - Applied Siamese BERT-Networks to retrieve items of Taobao, given particular style descriptions.

SKILLS

- Language: Proficient in Python, SQL, LaTeX, C++, C, R, Java, JavaScript