

# Yuhang Zhou

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

## RESEARCH INTERESTS

Natural Language Processing, Large Language Model Fine-tuning, Social Media Mining

## 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*

- Coursework: Deep Learning for Vision, Data Mining, Database Management System

May 2020

GPA: 3.90/4.0

**Shanghai Jiao Tong University**, Shanghai, China

- *Bachelor of Science in Electrical and Computer Engineering*

- Coursework: Data Mining and Machine Learning, Honors Mathematics, Discrete Mathematics

Aug 2020

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.

- *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

**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 and proposed the emoji substitution method to resolve the effect.

## PUBLICATION

• **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*, 2024a

• **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*, 2024a

• 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**, Xuan Lu, Ge Gao, Qiaozhu Mei, and Wei Ai. Emoji promotes developer participation and issue resolution on github. In *Proceedings of the Eighteenth International AAAI Conference on Web and Social Media (ICWSM)*, 2024b

• **Zhou, Yuhang\***, Suraj Maharjan\*, and Beiye Liu. Scalable prompt generation for semi-supervised learning with language models. In *Findings of EACL 2023*, 2023

• Paiheng Xu\*, **Zhou, Yuhang\***, Bang An, Wei Ai, and Furong Huang. Gfairhint: Improving individual fairness for graph neural networks via fairness hint. In *Workshop on Trustworthy and Socially Responsible Machine Learning, NeurIPS 2022*, 2022

• **Zhou, Yuhang** and Wei Ai. #emoji: A study on the association between emojis and hashtags on twitter. In *Proceedings of the Eighteenth International AAAI Conference on Web and Social Media (ICWSM)*, 2022

• Oana Ignat, Santiago Castro, **Zhou, Yuhang**, Jiajun Bao, Dandan Shan, and Rada Mihalcea. When did it happen? duration-informed temporal localization of narrated actions in vlogs. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 2022

• Luke J DeRoos, **Zhou, Yuhang**, Wesley J Marrero, Elliot B Tapper, Christopher J Sonnenday, Mariel S Lavieri, David W Hutton, and Neehar D Parikh. Assessment of national organ donation rates and organ procurement organization metrics. *JAMA surgery*, 156(2):173–180, 2021

## PREPRINT

• **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. *arxiv*, 2024

• **Zhou, Yuhang**, Xuan Lu, and Wei Ai. From adoption to adaption: Tracing the diffusion of new emojis on twitter. *arXiv*, 2024a

• Xiyao Wang, Jiuhai Chen, Zhaoyang Wang, **Zhou, Yuhang**, Yiyang Zhou, Huaxiu Yao, Tianyi Zhou, Tom Goldstein, Parminder Bhatia, Furong Huang, et al. Enhancing visual-language modality alignment in large vision language models via self-improvement. *arXiv*, 2024b

• Xiaoyu Liu, Paiheng Xu, Junda Wu, Jiaxin Yuan, Yifan Yang, **Zhou, Yuhang**, Fuxiao Liu, Tianrui Guan, Haoliang Wang, Tong Yu, et al. Large language models and causal inference in collaboration: A comprehensive survey. *arXiv*, 2024

• **Zhou, Yuhang**, Paiheng Xu, Xiyao Wang, Xuan Lu, Ge Gao, and Wei Ai. Emojis decoded: Leveraging chatgpt for enhanced understanding in social media communications. *arXiv*, 2024b

## INTERNSHIP

**Amazon AGI**, New York, USA

- *Applied Scientist Intern, Manager: Giannis Karamanolakis* May 2024-Aug 2024
  - Designed a unified MoE (Mixture of Experts) model with various merging method to improve the performance.
  - Reproduced the results of state-of-the-art merging methods of large language models.

**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.
  - Documented Results into a paper and published them 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.
  - Documented Results into a paper and published them in *Findings of EACL*, 2023

**Alibaba Group**, Hangzhou, China

- *Research Intern, Manager: Tianyu Li* May 2021-Aug 2021
  - Designed a SQL framework to mine the consumer profiles of e-commerce merchants in the transaction data.
  - 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