

Yuhang Zhou

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

RESEARCH INTERESTS

Computational Social Science, Causal Inference, Natural Language Processing

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, CLIP Lab

- *Causal Effect of Emojis in GitHub Issues, advised by Prof. Wei Ai*

- Extracted emoji information in GitHub issues to detect emoji's effect on developer participation on GitHub.
- Applied propensity score matching method to control the confounding variables, e.g. issues length and topics.

- *Emoji Hashtag Association Analysis, advised by Prof. Wei Ai*

- Extracted the emojis and hashtags information from the twitter data using Pyspark framework.
- Encoded the emoji feature and clustered the hashtags by applying the hierarchical clustering algorithm.

University of Maryland

- *Individual Fairness Promotion in GNN, advised by Prof. Furong Huang*

- Designed representation learning model to promote individual fairness for multiple graph neural network models.
- Constructed a "individually fair" graph and trained the fairness representation learning model on it.

PUBLICATION

1. **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
2. **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
3. 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
4. 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
5. **Zhou, Yuhang**, Xuan Lu, Ge Gao, Qiaozhu Mei, and Wei Ai. Emoji promotes developer participation and issue resolution on github. In *ICWSM*, 2024
6. **Zhou, Yuhang***, Suraj Maharjan*, and Beiye Liu. Scalable prompt generation for semi-supervised learning with language models. In *Findings of EACL 2023*, 2023

7. **Zhou, Yuhang** and Wei Ai. #emoji: A study on the association between emojis and hashtags on twitter. In *ICWSM*, 2022
8. 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
9. 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

INTERNSHIP

Amazon Alexa AI, Cambridge, 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 various datasets and multiple public models with the corresponding promoting methods

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.

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, ~~La~~TeX, C++, C, R, Java, JavaScript