Leitian Tao

Mobile: +1-608-332-5670

Email: leitiantao@cs.wisc.edu
Link: taoleitian.github.io

RESEARCH INTERESTS

Post-Training of Large Language Model, Preference Learning, AI Agent, AI safety, Multi-modality

EDUCATION

- **University of Wisconsin–Madison** — Madison, Wisconsin, United States
  Ph.D. student in Computer Science — *August 2023 - Present*
  Research interest: *Large Language Model Alignment, AI Safety, AI agent, Reinforcement Learning.*

- **Wuhan University** — Wuhan, Hubei, China
  Undergraduate in Information Engineering — *August 2019 - June 2023*
  Advisor: *Prof. Zhenzhong Chen*

EXPERIENCE

- **Adobe Research** — San Jose, United States
  Research Scientist Intern, mentored by Dr. Xiang Chen. — *May. 2024 - Present*
  - **LLM for code generation**: Developed CodeLutra, a framework that improves code generation in smaller LLMs by using iterative preference learning to refine outputs from both successful and failed attempts.

- **University of Wisconsin Madison** — Madison, United States
  Research Assistant, under the supervision of Prof. Sharon Li. — *Dec. 2023 - Present*
  - **Weak-to-strong alignment**: Concludes that weak LLM can actually play a good supervisory role in the alignment of the student model. The secret is a strong teacher.
  - **Out-of-distribution detection**: Proposed Non-Parametric Outlier Synthesis (NPOS), a novel framework for generating OOD data without distributional assumptions, significantly improving OOD detection and outperforming existing methods

- **Shanghai Artificial Intelligence Lab** — Shanghai, China
  Research Intern, under the supervision of Dr. Peng Gao. — *October 2022 - March 2023*
  - **Multi-modality model**: Explore the effects and in-depth reasons of different fine-tuning methods on multimodal models

- **Wuhan University** — Wuhan, China
  Research Assistant, under the supervision of Prof. Zhenzhong Chen. — *July 2020 - March 2022*
  - **Scene graph generation**: Design a semantically relevant learning framework to better accomplish multimodal learning.

PREPRINT

- **Your Weak LLM is Secretly a Strong Teacher for Alignment**:
  ***Leitian Tao**, Yixuan Li*

- **CodeLutra: Boosting LLM Code Generation via Preference-Guided Refinement**:
  ***Leitian Tao**, Xiang Chen, Tong Yu, Tung Mai, Ryan Rossi, Yixuan Li, Yixuan Li,Saayan Mitra*

- **How Reliable Is Human Feedback For Aligning Large Language Models?**:
  *Min-Hsuan Yeh, **Leitian Tao**, Jeffrey Wang, Xuefeng Du, Yixuan Li*

PUBLICATIONS

- **Non-Parametric Outlier Synthesis**: ICLR 2023
  ***Leitian Tao**, Xuefeng Du, Xiaojin Zhu, Yixuan Li.*

- **Predicate Correlation Learning for Scene Graph Generation**: IEEE TIP
  ***Leitian Tao**, Li Mi, Nannan Li, Xianhang Cheng, Yaosi Hu, Zhenzhong Chen*

- **Activate and Reject: Towards Safe Domain Generalization under Category Shift**: ICCV 2023
  *Chaoqi Chen, Luyao Tang, **Leitian Tao** , Hong-Yu Zhou , Yue Huang , Xiaoguang Han , Yizhou Yu.*

PROFESSIONAL SERVICES

- **Conference Reviewer**: ICML , ICLR, NeurIPS, ICME
- **Journal Reviewer**: IJCV

TEACHING

- **Teaching Assistant at University of Wisconsin Madison**: Data Science Programming II (CS-320)

SKILLS SUMMARY

- **Languages**: Python, C++, JavaScript, SQL
- **Frameworks**: Pytorch, Sklearn, TensorFlow, Flask