TIANXIANG SUN

2005 Songhu Rd., Shanghai, China 200438

(+86) 198 2125 0368 ⋄ txsun1997.github.io ⋄ txsun19@fudan.edu.cn

EDUCATION

Fudan University *September 2019 - June 2024*
Ph.D., Computer Science Shanghai, China
School of Computer Science
Advisor: Prof. Xipeng Qiu & Prof. Xuanjing Huang

Xidian University *September 2015 - June 2019*
Bachelor of Engineering, Software Engineering Xi'an, China
School of Computer Science and Technology (Excellent Track)
GPA: 3.8/4.0 (top 1%), Outstanding Graduate

Shengli NO.1 Senior High School *September 2012 - June 2015*
First Prize of China High School Biology Olympiad Dongying, Shandong, China

INTERNSHIP

Shanghai AI Lab *March 2023 - January 2024*
Research Intern Shanghai, China

Alibaba DAMO Academy *December 2022 - March 2023*
Research Intern Hangzhou, China

Amazon Web Services *October 2019 - August 2020*
Applied Scientist Intern Shanghai, China
Amazon Shanghai AI Lab
Supervisor: Prof. Zheng Zhang

RESEARCH INTERESTS

My research interests lie in natural language processing and machine learning, specifically in:

- large language models and their optimization/inference-efficient algorithms
- knowledge-enhanced pre-training
- neural-based multitask learning for NLP

ACADEMIC SERVICES

Seminar Chair

CCL 2023

Conference Reviewer / Program Committee Member

ACL (2021, 2022, 2023, 2024)

EMNLP (2021, 2022)

COLING (2020, 2022)

COLM (2024)

ICML (2022, 2024)

ICLR (2023)

NeurIPS (2022, 2023)

AAAI (2021)

IJCAI (2021)

HONORS & AWARDS

ByteDance Scholarships (13 winners across China)	*October 2023*
National Scholarships	*October 2023*
WAIC Yunfan Award - Rising Star (15 winners across the world)	*July 2023*
Fudan Academic Star (top academic honor in Fudan University)	*May 2023*
Most Influential Paper Award of Sci. China Tech Sci.	*March 2022*
National Scholarship for Graduate Student	*October 2020*
Outstanding Graduate of Xidian University	*June 2019*
First Prize in China High School Biology Olympiad	*June 2014*

PUBLICATIONS

[Google scholar citation: 2700+, H-index: 17]

* denotes equal contribution. Publications are sorted by the order in author lists.

1. **Multitask Pre-Training of Modular Prompt for Chinese Few-Shot Learning**
   Tianxiang Sun*, Zhengfu He*, Qin Zhu, Xipeng Qiu, Xuanjing Huang
   ACL 2023

2. **BBTv2: Towards a Gradient-Free Future with Large Language Models**
   Tianxiang Sun, Zhengfu He, Hong Qian, Yunhua Zhou, Xuanjing Huang, Xipeng Qiu
   EMNLP 2022

3. **BERTScore is Unfair: On Social Bias in Language Model-Based Metrics for Text Generation**
   Tianxiang Sun, Junliang He, Xipeng Qiu, Xuanjing Huang
   EMNLP 2022

4. **Black-Box Tuning for Language-Model-as-a-Service**
   Tianxiang Sun, Yunfan Shao, Hong Qian, Xuanjing Huang, Xipeng Qiu
   ICML 2022 (Spotlight)

5. **Paradigm Shift in Natural Language Processing**
   Tianxiang Sun, Xiangyang Liu, Xipeng Qiu, Xuanjing Huang
   Machine Intelligence Research (Invited Paper)

6. **A Simple Hash-Based Early Exiting Approach For Language Understanding and Generation**
   Tianxiang Sun, Xiangyang Liu, Wei Zhu, Zhichao Geng, Lingling Wu, Yilong He, Yuan Ni, Guotong Xie, Xuanjing Huang, Xipeng Qiu
   Findings of ACL 2022

7. **CoLAKE: Contextualized Language and Knowledge Embedding**
   Tianxiang Sun, Yunfan Shao, Xipeng Qiu, Qipeng Guo, Yaru Hu, Xuanjing Huang, Zheng Zhang
   COLING 2020

8. **Learning Sparse Sharing Architectures for Multiple Tasks**
   Tianxiang Sun*, Yunfan Shao*, Xiaonan Li, Pengfei Liu, Xipeng Qiu, Xuanjing Huang
   AAAI 2020 (Oral)

9. **Early Exiting with Ensemble Internal Classifiers**
   Tianxiang Sun, Yunhua Zhou, Xiangyang Liu, Xinyu Zhang, Hao Jiang, Zhao Cao, Xuanjing Huang, Xipeng Qiu
   arXiv:2105.13792 2021

10. **Can AI Assistants Know What They Don't Know?**
    Qinyuan Cheng*, **Tianxiang Sun***, Xiangyang Liu, Wenwei Zhang, Zhangyue Yin, Shimin Li, Linyang Li, Kai Chen, Xipeng Qiu
    ICML 2024

11. **CodeIE: Large Code Generation Models are Better Few-Shot Information Extractors**
    Peng Li*, **Tianxiang Sun***, Qiong Tang, Hang Yan, Yuanbin Wu, Xuanjing Huang, Xipeng Qiu
    ACL 2023

12. **DiffusionBERT: Improving Generative Masked Language Models with Diffusion Models**
    Zhengfu He*, **Tianxiang Sun***, Kuanning Wang, Xuanjing Huang, Xipeng Qiu
    ACL 2023

13. **Towards Efficient NLP: A Standard Evaluation and A Strong Baseline**
    Xiangyang Liu*, **Tianxiang Sun***, Junliang He, Lingling Wu, Xinyu Zhang, Hao Jiang, Zhao Cao, Xuanjing Huang, Xipeng Qiu
    NAACL 2022 (Oral)

14. **Late Prompt Tuning: A Late Prompt Could Be Better Than Many Prompts**
    Xiangyang Liu, **Tianxiang Sun**, Xuanjing Huang, Xipeng Qiu
    Findings of EMNLP 2022

15. **Pre-trained Models for Natural Language Processing: A Survey**
    Xipeng Qiu, **Tianxiang Sun**, Yige Xu, Yunfan Shao, Ning Dai, Xuanjing Huang
    SCIENCE CHINA Technological Sciences (Invited Paper, most influential paper award)

16. **Agent Alignment in Evolving Social Norms**
    Shimin Li, **Tianxiang Sun**, Xipeng Qiu
    arXiv:2401.04620 2024

17. **Evaluating Hallucinations in Chinese Large Language Models**
    Qinyuan Cheng, **Tianxiang Sun**, Wenwei Zhang, Siyin Wang, Xiangyang Liu, Mozhi Zhang, Junliang He, Mianqiu Huang, Zhangyue Yin, Kai Chen, Xipeng Qiu
    arXiv:2310.03368 2023

18. **Improving Contrastive Learning of Sentence Embeddings from AI Feedback**
    Qinyuan Cheng, Xiaogui Yang, **Tianxiang Sun**, Linyang Li, Xipeng Qiu
    Findings of ACL 2023

19. **Accelerating BERT Inference for Sequence Labeling via Early-Exit**
    Xiaonan Li, Yunfan Shao, **Tianxiang Sun**, Hang Yan, Xipeng Qiu, Xuanjing Huang
    ACL 2021

20. **Does Syntax Matter? A Strong Baseline for Aspect-Based Sentiment Analysis with**

RoBERTa
Junqi Dai, Hang Yan, **Tianxiang Sun**, Pengfei Liu, Xipeng Qiu
NAACL 2021

21. **Data Mixing Laws: Optimizing Data Mixtures by Predicting Language Modeling Performance**
Jiasheng Ye, Peiju Liu, **Tianxiang Sun**, Yunhua Zhou, Jun Zhan, Xipeng Qiu
arXiv:2403.16952 2024

22. **Flames: Benchmarking value alignment of chinese large language models**
Kexin Huang, Xiangyang Liu, Qianyu Guo, **Tianxiang Sun**, Jiawei Sun, Yaru Wang, Zeyang Zhou, Yixu Wang, Yan Teng, Xipeng Qiu, Yingchun Wang, Dahua Lin
NAACL 2024

23. **Dictionary Learning Improves Patch-Free Circuit Discovery in Mechanistic Interpretability: A Case Study on Othello-GPT**
Zhengfu He, Xuyang Ge, Qiong Tang, **Tianxiang Sun**, Qinyuan Cheng, Xipeng Qiu *arXiv:2402.12201* 2024

24. **DenoSent: A Denoising Objective for Self-Supervised Sentence Representation Learning**
Xinghao Wang, Junliang He, Pengyu Wang, Yunhua Zhou, **Tianxiang Sun**, Xipeng Qiu
AAAI 2024

25. **Llatrieval: Llm-verified retrieval for verifiable generation**
Xiaonan Li, Changtai Zhu, Linyang Li, Zhangyue Yin, **Tianxiang Sun**, Xipeng Qiu
NAACL 2024

26. **Turn Waste into Worth: Rectifying Top-k Router of MoE**
Zhiyuan Zeng, Qipeng Guo, Zhaoye Fei, Zhangyue Yin, Yunhua Zhou, Linyang Li, **Tianxiang Sun**, Hang Yan, Dahua Lin, Xipeng Qiu *arXiv:2402.12399* 2024

27. **AnyGPT: Unified Multimodal LLM with Discrete Sequence Modeling** Jun Zhan, Junqi Dai, Jiasheng Ye, Yunhua Zhou, Dong Zhang, Zhigeng Liu, Xin Zhang, Ruibin Yuan, Ge Zhang, Linyang Li, Hang Yan, Jie Fu, Tao Gui, **Tianxiang Sun**, Yugang Jiang, Xipeng Qiu
arXiv:2402.12226 2024

28. **Secrets of RLHF in Large Language Models Part I: PPO**
Rui Zheng, Shihan Dou, Songyang Gao, Yuan Hua, Wei Shen, Binghai Wang, Yan Liu, Senjie Jin, Qin Liu, Yuhao Zhou, Limao Xiong, Lu Chen, Zhiheng Xi, Nuo Xu, Wenbin Lai, Minghao Zhu, Cheng Chang, Zhangyue Yin, Rongxiang Weng, Wensen Cheng, Haoran Huang, **Tianxiang Sun**, Hang Yan, Tao Gui, Qi Zhang, Xipeng Qiu, Xuanjing Huang
arXiv:2307.04964 2023