Guangji Bai

201 Dowman Drive, Atlanta GA, 30322, USA | 202-957-7486 | guangji.bai@emory.edu | Google Scholar | LinkedIn

Research Statement

I am a fifth-year Ph.D. student at CS Department, Emory University working with Prof. Liang Zhao. I am generally interested in designing **efficient** and **generalizable** machine learning algorithms. Specifically, my current research topics include but are not limited to **1.** Domain/knowledge transfer, such as multi-task learning, domain adaptation, and domain generalization. **2.** Efficient machine learning for large-scale problems, such as model compression and acceleration of Large Language Models (LLMs), distributed training algorithms of deep neural networks. **3.** Memory-efficient continual/lifelong learning with experience replay and neuroscience inspiration.

Education

Emory University — **Atlanta, GA**
Ph.D. in Computer Science (GPA 3.95/4) — 2020.8-Present

The George Washington University — **Washington D.C.**
M.S. in Statistics — 2018.9-2020.5

Fudan University — **Shanghai, China**
B.S. in Mathematics — 2014.9-2018.6

Internship

Argonne National Laboratory — **Lemont, IL.**
Mathematics and Computer Science Division — 2024.5-2024.8

- I worked on how to integrate model pruning of LLMs under the privacy-preserved federated learning setting.
- I proposed an adaptive LLM pruning algorithm with personalization, tailored for the federated learning setting. The project is in process, but the pruning algorithm is available on arXiv here.

NEC Laboratory America — **Princeton, NJ.**
Data Science and System Security Team — 2023.5-2023.8

- I worked on developing machine learning algorithms for domain adaptation on time series data.
- We generalized the prompt tuning techniques from NLP to time series domain and leveraged the prompts to learn domain- specific and domain-invariant representation. Our work has been accepted by KDD 2024.

Selected Publications

- **Guangji Bai**, *et al*. "Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models." *Under review of CSUR* (more than **40** citations since 2024 on Google Scholar)
- **Guangji Bai**, Yijiang Li, Chen Ling, Kibaek Kim, Liang Zhao. "*SparseLLM*: Towards Global Pruning for Pre-trained Language Models." *(NeurIPS 2024)*
- **Guangji Bai***, Chen Ling*, Liang Zhao. "Temporal Domain Generalization with Drift-Aware Dynamic Neural Networks". *(ICLR 2023, **Oral, top 1% among all papers**).*
- **Guangji Bai,** Chen Ling, Yuyang Gao, Liang Zhao. "Saliency-Augmented Memory Completion for Continual Learning." *SIAM International Conference on Data Mining (SDM 2023)*
- **Guangji Bai**, Liang Zhao. "Saliency-Regularized Deep Multi-Task Learning." *(KDD 2022)*
- Junxiang Wang*, **Guangji Bai***, Wei Cheng, Zhengzhang Chen, Liang Zhao, Haifeng Chen. "Prompt-based Domain Discrimination for Multi-source Time Series Domain Adaptation." *(KDD 2024)*
- Zishan Gu, Ke Zhang, **Guangji Bai**, Liang Chen, Liang Zhao, Carl Yang. "Dynamic Activation of Clients and Parameters for Federated Learning over Heterogeneous Graphs.". *(ICDE 2023)*

**Equal contribution*. For a comprehensive list of my publication, please refer to my homepage.

Professional Services, Grants and Awards

- PC member for AISTATS (23'24'), NeurIPS (22'23'24'), ICLR (24'), AAAI (24'), ICML (24'), etc.
- Primary writer for the **NSF NAIRR** 240189 grant ($15k) on parallel and distributed training of LLMs on graphs.
- Travel Awards: KDD 2022, CIKM 2022, ICLR 2023, SDM 2023.

Skills

- Programming: Python, PyTorch, TensorFlow, MATLAB
- English-Proficiency
- Chinese – Native proficiency