YUE YU

E-mail: yueyu@gatech.edu | Webpage: yueyu1030.github.io

Address: CODA Building 1312, 756 W Peachtree St NW, Atlanta, GA 30308, USA

EDUCATION

School of Computational Science and Engineering, Georgia Institute of Technology   **Atlanta, GA, USA**
Ph.D. in Computational Science and Engineering   *Aug. 2019 - Dec. 2024 (Expected)*

- Ph.D. Advisor: Dr. Chao Zhang;
- **Research Interest**: Augmented Language Models, Data-centric AI.
- **Core Courses**: Computational Data Analysis, Mathematical Foundation of Machine Learning, Algorithms, Convex Optimization, Modeling and Simulation, Deep Learning for Text Data.
- **Thesis Topic**: Towards Efficiently and Effectively Harnessing Large Pre-trained Models via Data-centric Lens.

Department of Electronic Engineering, Tsinghua University   **Beijing, China**
B.Eng. in Electronic Engineering   *Aug. 2015 - July 2019*

- Research Assistant in the Future Internet & Communication Lab advised by Dr. Yong Li;

INDUSTRY EXPERIENCE

Llama Team, Meta GenAI   **Bellevue, WA, USA**
Research Intern, Host: *Rui Hou*, Manager: *Melanie Kamdadur*   *May - Aug. 2024*

Applied Deep Learning Research Team, NVIDIA   **Santa Clara, CA, USA (Remote)**
Research Intern, Host: *Wei Ping*, Manager: *Mohammad Shoeybi*   *Jan. 2024 - May 2024*

Topic: Retrieval-augmented Instruction Finetuning for Large Language Models.

Publication: One paper under review.

News Understanding Team, Google Research (Now Google Deepmind)   **New York City, NY, USA**
Research Intern, Host: *Jiaming Shen*, Co-host: *Tianqi Liu, Zhen Qin*, Manager: *Jialu Liu*   *May 2023 - Aug. 2023*

Topic: Empowering Large Language Model In-context Learning with Free-text Rationales.

Publication: One conference paper in ACL 2024.

Productivity and Intelligence Group, Microsoft Research   **Redmond, WA, USA**
Research Intern, Mentor: *Chenyan Xiong*, Manager: *Arnold Overwijk*   *May 2021 - Aug. 2021*

Topic: Zero-shot Learning for Generlizable Dense Text Retrieval.

Publication: One conference paper in EMNLP 2022.

Analytics Center of Excellence, IQVIA   **Boston, MA, USA**
Machine Learning Research Intern, Mentor: *Cao (Danica) Xiao*   *May 2020 - Aug. 2020*

Topic: Multi-typed Drug Interaction Prediction via Knowledge Graph Summarization.

Publication: One journal paper in Bioinformatics 2021.

SELECTED PUBLICATIONS

(The full publication list can be found in this link, * stands for equal contribution):

1. **Yue Yu**, Wei Ping, Zihan Liu, Boxin Wang, Jiaxuan You, Chao Zhang, Mohammad Shoeybi, and Bryan Catanzaro. "RankRAG: Unifying Context Ranking with Retrieval-Augmented Generation in LLMs". In *arXiv*, 2024.

2. **Yue Yu**, Jiaming Shen, Tianqi Liu, Zhen Qin, Jing Nathan Yan, Jialu Liu, Chao Zhang, Michael Bendersky. "Explanation-aware Soft Ensemble Empowers Large Language Model In-context Learning". In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2024.

3. **Yue Yu***, Yuchen Zhuang*, Jieyu Zhang*, Yu Meng, Alexander Ratner, Ranjay Krishna, Jiaming Shen, Chao Zhang. "Large Language Model as Attributed Training Data Generator: A Tale of Diversity and Bias". In *Proceedings of the Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks (NeurIPS)*, 2023.

4. Yuchen Zhuang*, **Yue Yu***, Kuan Wang*, Haotian Sun, Chao Zhang. "ToolQA: A Dataset for LLM Question Answering with External Tools". In *Proceedings of the Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks (NeurIPS)*, 2023.

5. **Yue Yu**, Rongzhi Zhang, Ran Xu, Jieyu Zhang, Jiaming Shen and Chao Zhang. "Cold-Start Data Selection for Few-shot Language Model Fine-tuning: A Prompt-Based Uncertainty Propagation Approach." In *Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.

6. **Yue Yu**, Chenyan Xiong, Si Sun, Chao Zhang, and Arnold Overwijk. "COCO-DR: Combating the Distribution Shift in Zero-Shot Dense Retrieval with Contrastive and Distributionally Robust Learning." In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2022.

7. **Yue Yu**, Lingkai Kong, Jieyu Zhang, Rongzhi Zhang, and Chao Zhang. "AcTune: Uncertainty-Based Active Self-Training for Active Fine-Tuning of Pretrained Language Models." In *Proceedings of the 2022 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022.

8. **Yue Yu***, Simiao Zuo*, Haoming Jiang, Wendi Ren, Tuo Zhao and Chao Zhang, "Fine-Tuning Pre-trained Language Model with Weak Supervision: A Contrastive-Regularized Self-Training Approach", In *Proceedings of the 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2021.

9. **Yue Yu**, Yinghao Li, Jiaming Shen, Hao Feng, Jimeng Sun and Chao Zhang, "STEAM: Self-Supervised Taxonomy Expansion via Path-Based Multi-View Co-Training", In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2020.

10. Chen Liang*, **Yue Yu***, Haoming Jiang*, Siawpeng Er, Ruijia Wang, Tuo Zhao and Chao Zhang, "BOND: Bert-Assisted Open-Domain Named Entity Recognition with Distant Supervision", In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2020.

HONORS AND AWARDS

- Best Paper Award at Machine Learning for Health 2022   *Nov. 2022*
- ACM SIGKDD Student Registration Award   *Aug. 2020*
- Excellent Graduate, Tsinghua University & Beijing City (Top 2% over 3292 graduate students)   *July 2019*
- Comprehensive Scholarship, Tsinghua Univiersity (Top 1%)   *Oct. 2018*
- Award from Tsinghua University Initiative Scientific Research Program (5000 USD)   *May 2018*
- Comprehensive Scholarship, Tsinghua Univiersity (Top 5%)   *Oct. 2016, Oct. 2017*

PROFESSIONAL SKILLS

- Programming language: C++, Python, MATLAB, Latex.
- Deep learning frameworks: Keras, Pytorch.

SERVICES

- **Teaching Experience**: TA for CX4240: Introduction to Computational Data Analysis.   Spring 2020, 2021, 2024
- **Area Chair**: ACL 2024, EMNLP 2024.
- **Conference Program Committee**: NeurIPS 2022/2023/2024; ICML 2022/2024; ICLR 2024; ACL 2023; KDD 2023; IJCAI 2023; EMNLP 2022/2023.