FANGCONG YIN

fangcongyin@utexas.edu
https://fangcong-yin-2.github.io

EDUCATION

The University of Texas at Austin 2023 - 2028 (Expected)
Ph.D. in Computer Science GPA: *3.92*

Cornell University 2021 - 2023
Bachelor of Science in Information Science, System, and Technology GPA: *4.03*

University of Notre Dame 2019 - 2021
Bachelor of Science in Computer Science GPA: *3.97*

PUBLICATIONS

1. **Fangcong Yin**, Xi Ye, and Greg Durrett. 2024. LoFiT: Localized Fine-tuning on LLM Representations. *Proceedings of the 38th Conference on Advances in Neural Information Processing Systems (NeurIPS 2024).*

2. Zayne Sprague, **Fangcong Yin**, Juan Diego Rodriguez, Dongwei Jiang, Manya Wadhwa, Prasann Singhal, Xinyu Zhao, Xi Ye, Kyle Mahowald, and Greg Durrett. 2024. To CoT or Not to CoT? Chain-of-thought Helps Mainly on Math and Symbolic Reasoning. *Preprint; In Submission to the 13th International Conference on Learning Representations (ICLR 2025).*

3. Xinyu Zhao, **Fangcong Yin**, and Greg Durrett. 2024. Understanding Synthetic Context Extension via Retrieval Heads. *In Submission to the 13th International Conference on Learning Representations (ICLR 2025).*

4. **Fangcong Yin** and Marten van Schijndel. 2023. Linguistic Compression in Single-Sentence Human-Written Summaries. *Findings of the 2023 Conference on Empirical Methods for Natural Language Processing (EMNLP 2023-Findings).*

RESEARCH EXPERIENCE

TAUR Lab, The University of Texas at Austin (Advisor: Greg Durrett) Austin, TX
Localized Fine-Tuning on LLM Representations Aug 2023 – May 2024

- Explore mechanistic interpretability methods to understand question answering and reasoning capabilities of large language models (LLMs)
- Develop localized, parameter-efficient fine-tuning methods that achieve comparable performance to LoRA with 200x fewer parameters **(Paper accepted to NeurIPS 2024)**

Investigate Limits of Chain-of-Thought Prompting May 2023 – Present

- Conduct controlled experiments to compare the performance of 14 LLMs between chain-of-thought prompting and tool augmentation to investigate the limits of chain-of-thought on 20 reasoning tasks (**Paper submitted to ICLR 2025**)

Understand Synthetic Data for Extending Context Window of LLMs July 2023 – Present

- Experiment with synthetic data construction approaches to fine-tune short-context LLMs to adapt to long-context retrieval and reasoning tasks
- Utilize interpretability methods to predict the transferability of fine-tuning with synthetic data for long-context tasks via attention heads (**Paper submitted to ICLR 2025**)

Cornell NLP Group, Cornell University Ithaca, NY

Linguistic Compression in Summarization (Advisor: Marten van Schijndel)Apr 2022 – June 2023

- Probe LLMs to explore the difference between human-written and model-generated summaries in terms of linguistic compression **(Paper accepted to EMNLP 23-Findings)**

Movie Summarization Benchmark (Advisor: Claire Cardie) Sep 2021 – June 2023

- Create a multi-reference scene-to-scene fine-grained movie summarization dataset for long-form summarization

Human Language Technology Center of Excellence, Johns Hopkins University Remote

Visiting Researcher May 2022 – August 2022

- Attend Summer Camp for Applied Language Exploration (SCALE) 2022 workshop on authorship identification

Data Mining Towards Decision Making Lab, University of Notre Dame Notre Dame, IN

Undergraduate Research Assistant (Advisor: Meng Jiang) Aug 2020 – Aug 2021

- Experiment with diverse question generation for document retrieval augmentation
- Build benchmark datasets to evaluate the diversity of natural language generation

RWTH Aachen University Aachen, Germany (Remote)

Research Assistant (Advisor: Elma Kerz & Daniel Wiechmann) May 2020 – Aug 2020

- Leverage linear mixed effects models to correlate text readability features with eye-tracking measures during human reading

GRANTS & AWARDS

- Honorable Mention, Computing Research Association's (CRA) Outstanding Undergraduate Researcher Award for 2023 Dec 2022
- Cornell Engineering Learning Initiatives Undergraduate Research Grant. *Linguistic Influences on Automatic Summarization Strategies.* ($ 2500) Fall 2022
- Wood Excellence Engineering Edu Research Award. *Linguistic Influences on Automatic Summarization Strategies.* ($ 3000) Summer 2022
- Cornell Engineering Learning Initiatives Undergraduate Research Grant. *MovieRecap Dataset Creation and Evaluation.* ($ 2100) Spring 2022

- College of Engineering Dean's List, Cornell University Fall 2021 – Fall 2022
- College of Engineering Dean's List, University of Notre Dame Fall 2019 – Spring 2021

SERVICES

Conference Reviewer: ACL Rolling Review (April 2024), EMNLP 2024, ICLR 2025
Workshop Reviewer: NeurIPS 2024 Workshop on Foundation Model Interventions
PhD Admission Reviewer: The University of Texas at Austin (2023)

TEACHING EXPERIENCE

Cornell University Ithaca, NY
Teaching Assistant, Language and Information (CS 4300) Spring 2023
Teaching Assistant, Natural Language Processing (CS 4740) Fall 2022
Teaching Consultant, Object-Oriented Programming and Data Structures (CS 2110) Spring 2022

WORK EXPERIENCE

Innovation and Automation Lab, Marmon Holdings, Inc. Chicago, IL
Intern Software Developer, Innovation and Automation Lab Dec 2020 – Aug 2021

- Developed a web application that automatically extracts hand measurements from images
- Customized object detection model to detect beacon lights from real-time videos in a manufacturing plant
- Created a full-stack native mobile application to detect scan medical trackers and upload their broadcasting data to database for a medical Internet of Things project

SKILLS

Programming Languages: Python, C++, Java, JavaScript, Matlab
Frameworks: PyTorch, Scikit-Learn, Angular, React
Languages: English, Mandarin Chinese, Japanese, German