Lingjie (Jason) Chen

+86 173-173-22856 | ljchen21@m.fudan.edu.cn | https://lingjiechen2.github.io/

Summary

Research Interests: My research primarily focuses on using **Graph Network** to optimize the recommendation system.

Relevant Courses: Numerical Linear Algebra(A), Statistical Machine Learning (A), Time Series(A-), Convex Optimization(A), Computer Architecture(A), Algorithms and Data Structures(A), Database(A-).

Education

Fudan University Sep. 2021 - Jun. 2025 (expected)
B.S. in Data Science *Shanghai, China*

- Major GPA: **3.84/4.0** (ranking **top 5** in the department); Overall GPA: 3.62/4.0

University of California, Berkeley Aug. 2023 - Jan. 2024
Exchange Student, Statistics *California, USA*

Publication

- **[C3] WAPITI: A Watermark for Finetuned Open-Source LLMs.**
  Lingjie Chen*, Ruizhong Qiu*, Siyu Yuan, Zhining Liu, Tianxin Wei, Hyunsik Yoo, Zhichen Zeng, Deqing Yang, Hanghang Tong
  Under review for ICML 2025. [Paper]
- **[C2] Llama Scope: Extracting Millions of Features from Llama-3.1-8B with Sparse Autoencoders.**
  Zhengfu He, Wentao Shu, Xuyang Ge, Lingjie Chen, Junxuan Wang, Yunhua Zhou, Frances Liu, Qipeng Guo, Xuanjing Huang, Zuxuan Wu, Yu-Gang Jiang, Xipeng Qiu
  Available on Arxiv [Paper]
- **[C1] "A good pun is its own reword": Can Large Language Models Understand Puns?**
  Zhijun Xu, Siyu Yuan, Lingjie Chen, Deqing Yang
  EMNLP 2024. [Paper]

IDEA Lab, University of Illinois Urbana-Champaign Apr. 2024 – Present
Research topics: Watermark, Statistical Analysis, Model Intervention *Illinois, USA*
Advisor: Prof. **Hanghang Tong**

- **Watermarking Fine-tuned Large Language Models[C3]**
  - Identified and validated the incompatibility between existing watermarking techniques and fine-tuned models.
  - Proposed a training-free, parameter-based watermarking method with thorough theoretical derivation.
  - Designed experiments to demonstrate the effectiveness and generalizability of our method.
  - Performed an in-depth analysis of our method, offering insights into its effectiveness.

OpenMoss, Fudan University Jan. 2024 – Present
Research topics: Software Engineering, LLM Interpretability *Shanghai, China*
Advisor: Prof. **Xipeng Qiu**

- **Exploration of intrinsic and transferred multilingualism[C2]**
  - Synthesized custom datasets to investigate the model's 'thinking state' during multilingual processing.
  - Designed cross-SAE patching experiments to examine the relationships within the feature space of LLMs.
  - Explored the internal mechanisms of multilingual models, revealing meaningful internal processes.

Shanghai Key Laboratory of Data Science Dec. 2022 – Dec. 2023
Research topics: Evaluation Methodology, Dataset *Shanghai, China*
Advisor: Prof. **Deqing Yang**

- **Evaluation of Large Language Models for Pun Understanding[C1]**
  - Conducted a systematic evaluation of eight different LLMs' capabilities in three pun-related tasks.
  - Designed and implemented novel pipelines for pun explanation and generation.
  - Improved the state-of-the-art performance of LLMs in pun understanding from 72% to 83%.

Project Portfolio (Selected)

Sparse AutoEncoder Framework Jan. 2024 – Present
Founder & Developer. [Code] *Shanghai, China*

- Provide a general codebase for conducting dictionary-learning-based mechanistic interpretability research
- Provides tools for analyzing and visualizing the learned dictionaries.

BERT-based Chinese QA English Oct. 2023 – Dec. 2023
Founder & Developer. [Code] *California, USA*

- Evalute the BERT-based model's performance on Chinese QA and provide a comparison with SOTA LLMs.

Attendance Checking Miniprogram Mar. 2023 – Jun. 2023
Founder & Developer. [Code] *Shanghai, China*

- Developed a WeChat Mini Program that enables attendance checks for both the teacher and student sides.

Academic Services

Reviewer International Conference on Learning Representations **(ICLR)**, 2024, 2025
Reviewer Annual Meeting of the Association for Computational Linguistics **(ACL)**, 2025
Reviewer Empirical Methods in Natural Language Processing **(EMNLP)**, 2024

Honors & Awards (Selected)

Fudan University Scholarship (Top 10%) **2021-2024**
Sou-Bin Scholarship (Top-performing students in Shanghai) **2019-2024**
Second Prize, CUMCM **2024**

Skills

Languages: Mandarin(Native speaker), English(TOFEL L30 R30 W25 S28)
Programming: Python, C/C++, LaTeX, MATLAB, Linux, R, SQL, Bash
Frameworks: Pytorch, Numpy, Anaconda, MySQL, Git, OpenCV