LingJie Chen

221 Handan Rd., Zibin Academy, Shanghai, China

lingjiechen127@gmail.com ⋄ 17317322856

EDUCATION

Fudan University   Shanghai, China

Bachelor of Science in Data Science   September 2021 - June 2025

- GPA: 3.74 / 4.0 (Top 5)
- Core Course: Deep Learning, Advanced Linear Algebra, Statistics

RESEARCH INTEREST

- **Large Language Models Generation Evaluation**

- **Mechanistic Interpretibility**

- **Trustworthy LLM**

RESEARCH EXPERIENCE

BERT-Based QA Task Improvement

- Executed QA tasks using BERT, enhancing model accuracy.
- Refined and reformatted datasets with Python, ensuring data integrity and usability.
- Developed a BERT-centric model architecture in PyTorch, leading to a 6% increase in SOTA performance for smaller models, nearing the efficacy of renowned Chinese-based models (Qwen, Baichuan).
- Conducted ablation studies to evaluate critical components and identify optimal BERT models for task-specific applications.
- Writing essay in ACL's format

Large Language Model Evaluation for Understanding Puns

- Deployed and optimized various prominent open-source LLMs (e.g., vLLM, Fastchat) on servers for enhanced efficiency.
- Engineered and implemented a testing framework to experimentally assess LLMs' understanding of puns.
- Applied statistical methods to analyze results, quantifying model comprehension while eliminating biases.
- Improved prompt finetuning, boosting SOTA performance metrics from 72% to 83%.

Multilingual Model's Transfer Mechanism Exploration

- Familiarized with and modified the Sparse Autoencoder framework to conduct probing.
- Deepened understanding of mechanism interpretation and the rationale behind Sparse Autoencoders.
- Utilized self-synthesized data to probe the model's 'thinking state' when processing different languages.
- Conducted pioneering work to explore the inner stages of multilingual models, uncovering interesting and meaningful internal processes.

PUBLICATIONS

A good pun is its own reword: Can Large Language Models Understand Puns?

EMNLP 2024 Under review

A Mechanistic View of Intrinsic Multilingualism in Large Language Models

EMNLP 2024 Under review

Technical Skills

- Proficient in Python, C++. Familiar with Linux and Microsoft Office

- Rich experience in ML related packages and PyTorch

- Rich experience in LLM deployment and experiments

- Solid foundation in Statistics and Algebra

- Familiar with essay writing and Latex formatting

Academic Activities

University of California, Berkeley   California, USA

Statistics   August 2023 - January 2024

- GPA: 3.93 / 4.0

- Core Course: Natural Language Processing, Reinforcement Learning, Time Series

Selected Honors And award

- School of Data Science - Model Student.

- Fudan University's Scholarship (Multiple times).

- National Undergraduate Mathematics Competition - Third Prize.

- Sou-Bin Scholarship (Multiple times).

- Shanghai English Speech Competition - First Prize.

Language Skills

- Fluent oral English and excellent writing skills.

- CET-4/CET-6: Excellent.

- IELTS: 8.0 (Listening: 9.0, Reading: 9.0, Writing: 7.5, Speaking: 7.0).

- TOEFL: 109 (Listening: 30, Reading: 30, Writing: 21, Speaking: 28).

Additional

- Hobby: Running, Photography, Basketball.