Rui Li

email: rui_li@mail.ustc.edu.cn, ruili4edu@gmail.com Personal Website Tel: (+86) 18857726868

EDUCATION

University of California, Berkeley, CA Aug. 2023 - Dec. 2023

Visiting Student

- GPA (overall): 4.0/4.0
- Core Courses: CS288 Natural Language Processing, CS189 Machine Learning

University of Science and Technology of China (USTC), Hefei, China Sep. 2021 - Jul. 2025

Bachelor of Engineering in Data Science (School of the Gifted Young)

- GPA (overall): **4.21/4.3** Ranking: **1/54** Weighted average score: **95.9**
- Core Courses: Deep Learning, Big Data Algorithm

HONORS AND AWARDS

Berkeley Global Access Program Scholarship (Top 1%), UC Berkeley	2023
National Scholarship (Top 1%), USTC	2022, 2023
Outstanding Freshman Scholarship, Gold Award, USTC	2021

PUBLICATION

- **Are Human-generated Demonstrations Necessary for In-context Learning?**
  Rui Li, Guoyin Wang, Jiwei Li preprint arXiv: 2309.14681
  (Submitted to ICLR2024 https://openreview.net/forum?id=frRDT6EOhg Ratings: 8666)
- Similarity-based Neighbor Selection for Graph LLMs Submitted to ICML2024
  Rui Li, Jiwei Li, Jiawei Han, Guoyin Wang **To be released on Arxiv around Feb 5th**

RESEARCH EXPERIENCE

Research Intern, Advisor: Dr. Jiawei Han, UIUC. Nov. 2023 - Jan. 2024

- **Similarity-based Neighbor Selection for Graph LLMs** Nov. 2023 - Jan. 2024
  - Demonstrate that LLMs could achieve performance on par with vanilla GNNs on node classification tasks. LLMs exhibit initial proficiency in handling node classification on graph-structured data.
  - Proposed Similarity based Neighbor Selection (SNS in short) to mitigate challenges such as over-squashing and heterophilous issues on traditional prompting methods

Research Assistant, Advisor: Dr. Jiwei Li, Zhejiang University May. 2022 - Dec. 2023

- **Are Human-generated Demonstrations Necessary for In-context Learning?**
  May. 2023 - Aug. 2023
  - Proposed SEC, a simple, resource-efficient and broadly applicable prompting strategy without the need for external training data and human intervention, which is promising to be a more comprehensive and stable paradigm for evaluating LLMs
  - SEC achieves the strongest zero-shot performance on a variety of tasks.
  - Conducted extensive experiments on ChatGPT in both answer-only and Chain-of-Thought (CoT) scenario to reveal the new properties of model-generated few-shot examples and provide some insights into the mechanism of in-context learning
  - Addressed some of the issues associated with traditional in-context learning methods, such as the lack of manually annotated data and the instability of the performance
  - Proposed a comprehensive pipeline to clean and extract desired information from the output of the LLMs
- **KNN Search Enhance the Few-shot Capability of LLMs in Summarization** Jan. 2023 - Apr. 2023
  - Proposed a prompting framework optimizing the selection of Few-shot examples for prompting LLMs like GPT using k Nearest Neighbor search

- Utilized divide and conquer, two steps of summary as well as truncation to effectively solve the context limit of LLMs while dealing with long text (during the course of this research, the model with the largest context was text-davinci-003, with a maximum of 4097 tokens)

■ **BRIO-GNN: Text Summarization Based on Global Corpus via GNN** May. 2022 - Dec. 2022

- Utilized Graph Neural Networks (GNNs) to address the challenge that traditional transformer-based fine-tuned summarization models cannot directly reference the training corpus

SKILLS

Experience in NLP:

Proficient in mainstream NLP frameworks, notably transformers and attuned to some of the latest advancements in NLP research

Experience with LLM:

Have a deep understanding of prompt engineering and various methods to exploit the ability of LLMs, such as post-editing, majority voting, Chain-of Thought Prompting, and Least-to-Most Prompting

Programming Skills:

Adept in C, C++, SQL, Python, deep learning frameworks such as PyTorch, familiar with Linux, Matlab, CSS

TOEFL Test: 103 **Writing:** 27

GRE General Test: Verbal Reasoning 162, Quantitative reasoning 170