Wang Xiaochen

Mobile: (+86)-155-1041-0388 · google scholar
Email: wangxiaochen@stu.pku.edu.cn · Blog: Shandong Province

Education

Peking University, School of Software and Microelectronics, *Master* 2022.09 - 2025.06

- MOE Key Laboratory of Computational Linguistics
- Advisor: Prof. Zhifang Sui

Beijing University of Posts and Telecommunications, College of Computer Science, *B.S.* 2018.09 - 2022.06

- Graph Data Mining and Machine Learning Lab
- Advisor: Prof. Cheng Yang

Publications

- **FSM: A Finite State Machine Based Zero-Shot Prompting Paradigm for Multi-Hop Question Answering**

  Xiaochen Wang, Junqing He, Zhe Yang, Yiru Wang, Xiangdi Meng, Kunhao Pan, Zhifang Sui *submitting to EMNLP*

- **Statistical Dataset Evaluation: Reliability, Difficulty, and Validity**

  Chengwen Wang[1], Qingxiu Dong[1], **Xiaochen Wang**, Haitao Wang, Zhifang Sui *Natural Language Engineering (minor revision)*

- **PeriodicLoRA: Breaking the Low-Rank Bottleneck in LoRA Optimization**

  Xiangdi Meng, Damai Dai, Weiyao Luo, Zhe Yang, Shaoxiang Wu, **Xiaochen Wang**, Peiyi Wang, Qingxiu Dong, Liang Chen, Zhifang Sui

  submitting to EMNLP

- **Abnormal Event Detection via Hypergraph Contrastive Learning**

  Bo Yan, Cheng Yang, Chuan Shi, Jiawei Liu, **Xiaochen Wang**

  Proceedings of the 2023 SIAM International Conference on Data Mining **[SDM 2023]**

Research Experiences

PeriodicLoRA: Breaking the Low-Rank Bottleneck in LoRA Optimization 2023.12 - 2024.03

- **Problems:** Rank adjustment during training is fixed and optimal rank value is uncertain. Besides, parameter scale for LoRA conflicts with memory usage.
- **Solutions:** We propose PeriodicLoRA (PLoRA), a method that periodically **unloads** the LoRA matrices back to the original backbone with a ratio. This process generates a higher rank update matrix by accumulating low rank matrices without extra memory usage.
- **Results:** Experimental findings demonstrate that PLoRA with low rank outperforms LoRA, improving training speed by 80%.

LoRA++, Adding an Implicit Momentum SGD Optimizer for LoRAs, in progress 2024.05 - present

- **Problems：** When tuning PLoRA, we discovered that the parameter update process can be viewed from the perspective of an optimizer, with the update interval treated as the batch size and PLoRA resembling momentum-SGD.
- **Solutions：** Inspiring from optimization, we incorporated dynamic learning rates and batch size to relieve overfitting. Besides, we imitate the AdamW to make the parameter updates smoother.
- **Results：** Our method, employing a lower-rank parameter scale, can outperform higher-rank LoRA and perform on par with full fine-tuning.

RouterLoRA, LoRA Meets MOE, exploring 2024.03 - 2024.04

- **Problems：** Using LoRA combined with MOE resolves the capability conflicts in multitask training by incorporating multiple LoRA modules as experts to isolate the parameter space. There are two different projects.
- **Solutions:** Instead of topk strategy, we fix foundational expert to share common knowledge or information avoiding knowledge redundancy in different experts.
- **Solutions:** Pre-trained LLMs is equipped with foundational knowledge and do not require finetuning for certain data. We design a dynamic selection (0-k) of experts, allowing the model to learn how to choose a number of experts for processing.

FSM: A Finite State Machine Based Zero-Shot Prompting Paradigm for Multi-Hop Question Answering 2023.12 - 2024.02

- **Problems:** Large Language models (LLMs) with few-shot methods underperform in multi-hop reasoning because of hallucination, error propagation and limited context length.
- **Solutions:** We design a zero-shot FSM framework based on Finite State Machine that explicitly supports the model in the phases of decomposition, retrieval, and verification. FSM tackles one task at a time, decides the next action based on the current results and states.
- **Results:** Our results have a more standardized format and outperform COT on challenging datasets like Musique.

Can Large Multimodal Models Uncover Deep Semantics Behind Comic Strip?, in progress 2024.07 - present

- **Problems：** Comic strips can be considered as important frames in video. We use them to evaluate the capability of video-LLM to understand the deep meanings of sequential multiple images.
- **Solutions：** We design three tasks: next frame prediction, Reorganize and deep meaning to evaluate.

Statistical Dataset Evaluation: Reliability, Difficulty, and Validity 2022.09 - 2023.01

- **Problems:** Existing datasets have exposed numerous issues, leading to biased models and unreliable evaluation, without human evaluation of dataset quality.
- **Solutions:** We seek to understand the statistical properties of datasets and address three fundamental dimensions: reliability, difficulty, and validity. Taking the Named Entity Recognition datasets as a case study, we introduce 9 statistical metrics.
- **Results:** We studied how the scores of datasets on statistical metrics impact model performance and advocate for assessing dataset quality or making targeted improvements to datasets before training or testing models.

Research Interests

- My research interests has been focusing on enhancing the reasoning capabilities of LLMs since 2023.
- I am also interested in multimodal and embodied intelligence because they can be implemented in real-world applications in the future.

Intern Experience

Shanghai AI LAB & SenseTime 2023.07 - 2023.10

Speech Group, Mentor: Baoxiang Li

International Digital Economy Academy 2023.11 - 2024.03

Cognitive Computing and Natural Language Research Center, Mentor: Junqing He and Kunhao Pan

Rewards

Beijing Outstanding Graduate	2022
National Motivational Scholarship	2019, 2020, 2021
Haohan Enterprise Scholarship	2019, 2020, 2021
Outstanding Student Leader	2021
Active Participant in Culture and Sports	2021
Merit Student	2020
Outstanding Communist Youth League Secretary	2019