YUCHEN WANG

wangyuchen333@stu.pku.edu.cn ◇ https://wangyuchen333.github.io/

EDUCATION

Peking University, *Bachelor of Science in Intelligence Science* 2021 - 2025(expected)
GPA: 87 (Yr1 84.7, Yr2 88.4, Yr3 87.4) Rank: Top 20% (6/29)

SELECTED AWARDS

- **Zhi Class Scholarship** (Top 1%) 2023, 2024
- **Outstanding Academic Achievement Award** (Top 5%) 2023
- **First Prize in Provincial Chinese Mathematical Olympiad(CMO)** 2020

ACADEMIC SERVICES

Advisor: Prof. Liwei Wang, Peking University 2023.9 – 2024.4

- **Explored Model Merging for Large Models**
  Inspired by the paper *Evolutionary Optimization of Model Merging Recipes* (arXiv:2403.13187), I implemented and experimented with several model merging methods, including SLERP, Task Arithmetic, TIES, and DARE-TIES, utilizing the `mergekit` library (GitHub: mergekit). Despite combining multiple approaches, the lack of merging code in the official repository limited further exploration.

Advisor: Prof. Difan Zou, The University of Hong Kong 2024.7 – Present

- **Decoding the Mechanisms of In-Context Learning (ICL)**
  My research investigates why in-context learning (ICL) is effective in large models. I believe that current ICL studies do not fully address this phenomenon, as there is a gap in understanding its application to smaller, non-generative transformers due to the lack of pretrained knowledge embedded in their weights. Recent research suggests that ICL scaling enables models to retrieve knowledge from pretraining, improves robustness to prompt ordering, and sometimes outperforms fine-tuning. My goal is to identify the key transformer components that enable ICL in large models, explain why it achieves fine-tuning-level performance, and ultimately contribute to improving existing transformer architectures or designing new frameworks for generative models.
- **Examining Collapse Errors in Diffusion Models (Expected to be published in December)**
  We identify and investigate a critical issue in ODE-based sampling for diffusion models, referred to as the *collapse error*, where samples tend to collapse locally. This phenomenon is observed across various datasets. Our analysis reveals that this error occurs early in the sampling process, with its effects progressively accumulating and amplifying throughout the entire ODE sampling.

SKILLS

Programming Languages: Python, PyTorch, C, C++, LaTeX
Standard test scores: TOEFL 104 (S21), CET6 590