Yanggan Gu

yanggangu@outlook.com | Google Scholar | Homepage | GitHub

Education

The Hong Kong Polytechnic University, Hong Kong, China. 2026/01 - Present
Ph.D. student in Computing. Advisor: Prof. Hongxia Yang

Soochow University [*Project 211*], Suzhou, China. 2022/09 - 2025/06
M.S. in Computer Science and Technology. GPA 3.7 (Top-5%)

Guangdong University of Technology, Guangzhou, China. 2018/09 - 2022/06
B.S. in Software Engineering. GPA 4.0 (Top-1%)

Research Experience

Model Fusion in LLMs, *Researcher* 2025/04 - Present
Department of Computing, Hong Kong Polytechnic University. Advisor: Prof. Hongxia Yang
Research Topic: Model Fusion, Alignment of LLMs, Reasoning in LLMs

Trustworthy LLMs, *Research Assistant* 2024/06 - 2025/03
AI Thrust, The Hong Kong University of Science and Technology (Guangzhou). Advisor: Prof. Xuming Hu
Research Topic: Alignment of LLMs, Factual Reasoning in LLMs, Knowledge Distillation

Syntactic Parsing, *Master's Researcher* 2022/09 - 2025/06
School of Computer Science and Technology, Soochow University. Advisor: Prof. Zhenghua Li
Research Topic: Linguistic Structure Prediction, Text2SQL (maintaining SeSQL dataset)

Selected Publications

FeatCal: Feature Calibration for Post-Merging Models [Code]
First Author **NeurIPS** 2026 Spotlight (CCF-A)
Calibrates merged models layer by layer with closed-form updates to reduce feature drift without gradient descent.

InfiFPO: Implicit Model Fusion via Preference Optimization in Large Language Models [Project]
First Author **NeurIPS** 2025 Spotlight (CCF-A)
Introduces sequence-level probability fusion through preference optimization and validates it on 11 benchmarks.

SMAT: Simple and Efficient Merge-Aware Training [Code] [HF]
Co-first Author **Preprint** 2026
Trains merge-aware experts by simulating Scale, Mask, and Perturb operations, with under 2% training-time overhead.

Model Merging Scaling Laws in Large Language Models [Code]
Co-first Author **ICML** 2026 (CCF-A)
Develops scaling laws that predict model-merging performance across model sizes, expert counts, and data regimes.

Capturing Nuanced Preferences: Preference-Aligned Distillation for Small Language Models [Code]
First Author **ACL** 2025 Findings (CCF-A)
Distills full preference distributions into small language models for more precise preference alignment.

StructFact: Reasoning Factual Knowledge from Structured Data with Large Language Models [Code]
Co-first Author **ACL** 2025 Findings (CCF-A)
Introduces StructFact, a 13,407-query benchmark for reasoning over heterogeneous structured knowledge.

From Parameters to Behaviors: A Survey of Model Fusion for Large Language Models
Co-first Author **EMNLP** 2026 Findings (CCF-B)
Provides a unified taxonomy of model fusion spanning parameter-, prediction-, and behavior-level methods.

Honors and Awards

- **1st place in the CCL-2023 Chinese AMR Parsing Evaluation Task.** [Report] 2023/08
- **Top 10 Innovative Figures of Guangdong University of Technology.** [Sohu News] 2021/12
- **National Encouragement Scholarship.** [Link] 2021/11

Academic Service

- **Reviewer**: ICML, ICLR, NeurIPS, ARR, and TMLR.