Anningzhe Gao

Phone：(+86)156 5278 9516 Email：gaoanningzhe@sribd.cn DOB：1994/08/29

Education

Tsinghua University – B.S. – Math and Applied Math 2012/08-2016/06

Rank(4/91) GPA 93/100 Major GPA 95/100

Vice chair, Association of science and technology of department of math

University of California, Berkeley – PhD Candidate - Mathematics 2016/08-2021/05

GPA 4.0/4.0

Advisor：Martin Olsson

Dissertation：Some results in essential dimension and arithmetic of elliptic curves and abelian varieties

Working Experience

SRIBD

2023/06-present

Associate Researcher

Tencent

2021/07-2023/06

Senior Algorithm Researcher

Publications and Preprints

[1] **Anningzhe Gao**, Essential dimension of moduli stack of polarized K3 surfaces , Proceedings of the American Mathematical Society, 2020

[2] Yu, Fei, **Anningzhe Gao(Co-corresponding author)**, and Benyou Wang. "Outcome-supervised verifiers for planning in mathematical reasoning." arXiv preprint arXiv:2311.09724 (2023). To appear in Findings of NAACL 2024

[3] Ge, W., Chen, S., Chen, G., Chen, J., Chen, Z., Yan, S., Zhu, C., Lin, Z., Xie, W., Wang, X. and **Gao, A.**, 2023. Mllm-bench, evaluating multi-modal llms using gpt-4v. arXiv preprint arXiv:2311.13951.

[4] Chen, J., Wang, X., **Gao, A. (Co-corresponding author)**, Jiang, F., Chen, S., Zhang, H., ... & Wang, B. (2023). Huatuogpt-ii, one-stage training for medical adaption of llms. arXiv preprint arXiv:2311.09774.

[5] Wang, Xidong, Nuo Chen, Junyin Chen, Yan Hu, Yidong Wang, Xiangbo Wu, **Anningzhe Gao**, Xiang Wan, Haizhou Li, and Benyou Wang. "Apollo: Lightweight Multilingual Medical LLMs towards Democratizing Medical AI to 6B People." arXiv preprint arXiv:2403.03640 (2024).

[6] **Anningzhe Gao**, Fourier-Mukai partners of abelian varieties, arxiv 1908.03308

[7] **Anningzhe Gao**, Tate conjecture and finiteness of abelian varieties over finite fields, arxiv 1808.04783

[8] **Anningzhe Gao**, The bijection between exceptional subcategories and non-crossing partitions, arxiv 1601.07672

[9] **Anningzhe Gao**, The period-index problem for elliptic curves and the essential dimension of Picard stacks, arxiv 2002.11814

Projects

1. HuatuoGPT (Co-PI):
    1. Purpose: Build a Large-scale Language Model in the medical field
    2. Method: We build a medical database consisting of Pubmed, Wudao, articles and other medical materials. Then we apply pretraining, supervised finetuning, RLHF to improve our model's behavior. The procedure of data deduplicating, domain-specific data identification and data cleaning is applied to get hybrid train data. During training, we apply dynamic sampling, medical knowledge injection and RLAIF to improve the diversity and safety of our model. Now the model achieved the highest score in the CMB (Chinese Medical

Benchmark) test.