

JIZHOU GUO

Homepage • sjtu18640985163@sjtu.edu.cn • LinkedIn • Google Scholar • DBLP

Research Interests: Large Language Models, Natural Language Processing.

EDUCATION

Shanghai Jiao Tong University • Shanghai, China	Aug 2022 – Present
<i>Bachelor of Science</i> • Zhiyuan College (Honor, Top 10%) • Mathematics and Applied Mathematics • GPA: 3.8	
Selected Courses: Topic Course (Deep Learning) (A+, 100), Topic Course (Applied Mathematics) (A+, 100), Data Structure (Honor) (A+), Introduction to Computer Science (A+), Numerical Methods for ODE & PDEs (A+), Differential Geometry (A+), Introduction to Physics (Honor) (A+), Problem-solving and Practice (A+), Foundations of Data Science (A), Mathematical Analysis (Honor) (A), Ordinary Differential Equations (A), Partial Differential Equations (A), etc.	

PUBLICATIONS

* denotes equal contribution

Mining Intrinsic Rewards from LLM Hidden States for Efficient Best-of-N Sampling

Jizhou Guo, Zhaomin Wu, Hanchen Yang, Philip S. Yu

KDD 2026 [arXiv]

Model-based Large Language Model Customization as Service

Zhaomin Wu*, **Jizhou Guo***, Junyi Hou, Bingsheng He, Lixin Fan, Qiang Yang

EMNLP 2025 main [arXiv]

PSG-Agent: Personality-Aware Safety Guardrail for LLM-based Agents

Yaozu Wu*, **Jizhou Guo***, Dongyuan Li*, Henry Peng Zou, Wei-Chieh Huang, Yankai Chen, Zhen Wang, Weizhi Zhang, Yangning Li, Meng Zhang, Renhe Jiang, Philip S. Yu

Under review [arXiv]

Calibrating Reasoning in Language Models with Internal Consistency

Zhihui Xie, **Jizhou Guo**, Tong Yu, Shuai Li

NeurIPS 2024 [arXiv] [poster] [code]

Cross-Stimulus Transfer Learning: Enhancing Emotion Recognition from Visual-Auditory to Olfactory Perception

Jiaqi Wang*, Zhengting Chen*, Keyan Huang, Yifan Wu, Dian Zhang, **Jizhou Guo**, Xinglan Liu, Dan Peng, Baoliang Lu, Weilong Zheng

EMBC 2025 [PubMed]

LLM DNA: Tracing Model Evolution via Functional Representations

Zhaomin Wu, Haodong Zhao, Ziyang Wang, **Jizhou Guo**, Qian Wang, Bingsheng He

Under review [arXiv]

DeepResearchGuard: Deep Research with Open Domain Evaluation and Multi-Stage Guardrails for Safety

Wei-Chieh Huang, Henry Peng Zou, Yaozu Wu, Weizhi Zhang, Yankai Chen, Dongyuan Li, Yangning Li, Angelo Zangari, **Jizhou Guo**, Liancheng Fang, Langzhou He, Chunyu Miao, Renhe Jiang, Philip S. Yu

Under review [arXiv]

RECODE-H: A Benchmark for Research Code Development with Interactive Human Feedback

RECODE Team

Under review [arXiv]

RESEARCH EXPERIENCE

BDSC Lab - University of Illinois Chicago	Jun 2025 – Aug 2025
Advisor: Prof. Philip S. Yu	

- Developed a highly parameter-efficient reward model leveraging the LLM hidden states, which systematically outperforms baselines with less than 0.005% of the parameters of baselines.
- Proposed PSG-Agent and a benchmark for personalized guardrail of LLM-based agents.

Xtra Group - National University of Singapore	Jun 2024 – Aug 2024
Advisor: Prof. Bingsheng He	

- Proposed a novel framework that facilitates LLM customization as a service for domain-specific applications. It achieved substantial performance improvements in domain-specific tasks, boosting accuracy by up to 26% while preserving privacy.

John Hopcroft Center for Computer Science - Shanghai Jiao Tong University	Oct 2023 – May 2024
Advisor: Prof. Shuai Li and Dr. Tong Yu	

- Worked on training-free LLM calibration approach by leveraging hidden states.

Zhiyuan Innovative Research Center - Shanghai Jiao Tong University	Dec 2022 - Jan 2024
Advisor: Prof. Bao-Liang Lu and Prof. Wei-Long Zheng [Certificate of Completion]	

- Implemented a Transformer-based Domain-Adversarial Neural Network (DANN) to transfer knowledge from visual to olfactory EEG data for emotion classification.

Quantitative Biology Summer School - Center for Life Sciences, Peking University

Jul 2023

- Chosen from 50 candidates nationwide.

Tencent Spark Project - Tencent Corporation

Aug 2022

- Chosen from 50 high-school students with talents nationwide.
- Engineered a robust palm liveness detection system and successfully blocking palm images displayed on screens.

SELECTED PROJECTS

Deep Reinforcement Learning: Insights from AlphaGo

Spring 2024

Team leader, advised by [Prof. Dan Hu](#) (**Scored 100**)

Spectral vs. Fourier Neural Operators in Parametric PDE Modeling: Analysis and Experiments

Spring 2025

Individual project, advised by [Prof. Lei Zhang](#) (**Scored 100**) [\[PDF\]](#) [\[code\]](#)

Vertical Federated Learning Challenge

Spring 2024

Individual project, advised by [Prof. Bingsheng He](#) [\[code\]](#)

LENS: Learning Ensemble Confidence from Neural States for Multi-LLM Answer Integration

Autumn 2024

Individual project, advised by [Prof. Xiaoqun Zhang](#) (Scored 95) [\[PDF\]](#)

SELECTED AWARDS

Click [here](#) to view all certificates

Contest Prizes

- Gold Award** and **First Runner-up** in the National College Students' Career Planning Contest (Shanghai), Jan 2025
- Third Prize in Mathematics competition of Chinese College Students (Shanghai), Dec 2023
- First Prize** in Shanghai Collegiate Programming Contest, Sep 2023 (Ranked 2nd in Shanghai)
- Gold Medal** in Astar Programming Contest (Shanghai region), Aug 2023 (Ranked 2nd in Shanghai)
- Gold Medal** in 2023 China Collegiate Programming Contest (**CCPC**) National Invitational Contest (Hunan), May 2023
- Gold Medal** in 2023 International Collegiate Programming Contest (**ICPC**) Xi'an Invitational Contest, May 2023
- Gold Medal** in 2022 International Collegiate Programming Contest (**ICPC**) Asia Hangzhou Regional Contest, Dec 2022 (Ranked 8th nationwide)
- Gold Medal** in 2022 China Collegiate Programming Contest (**CCPC**) (Shanghai region), Sep 2022
- Silver Medal** in National Olympiad in Informatics (**NOI**), Jul 2021
- Ranked 22nd nationwide** in National Olympiad in Informatics (**NOI**) Online Senior Group, Mar 2021

Honors

- Zhiyuan **First-Class** Overseas Research Scholarship
- Merit Student of SJTU (twice)
- Academic Excellence Scholarship, SJTU (Top 10%, ranked 2nd overall, twice)
- Zhiyuan Honors Scholarship (three times)

SERVICES

Invited as reviewer: [ICLR 2026](#), [EMBC 2025](#), [AI4MATH @ ICML 2025](#)

SKILLS

- Programming languages: Python, C/C++, Matlab.
- Language: Chinese (Native Speaker), English (Proficient, TOEFL 110, TOEFL MyBest Score 113).
- Expertise & Hobbies: Piano (Amateur Level 10), Singing (Amateur Level 9), Music Theory (Amateur Level 5).