





YIJIA XIAO

 [yijia-xiao.github.io](https://github.com/yijia-xiao) |  [linkedin.com/in/yijia-xiao](https://www.linkedin.com/in/yijia-xiao) |  Google Scholar |  yijia.xiao@cs.ucla.edu
Los Angeles, CA, USA

OBJECTIVE

Seeking to apply my expertise in **Multimodal Large Language Models** and **LLM Agents** to:

- I. Multi-Modal Alignment:** Fuse domain-specific data (e.g., science, finance) with text/image modalities using LLMs.
- II. First Principles:** Derive complex properties from primary data like financial time series and biological sequences.

EDUCATION

- **University of California, Los Angeles** 2022 – June 2027
Ph.D. Student in Computer Science Los Angeles, CA, USA
 - Advisor: Professor [Wei Wang](#)
- **Tsinghua University** 2018 – 2022
Bachelor of Computer Science and Technology Beijing, China
 - Advisor: Professor [Jie Tang](#)

EXPERIENCE

- **Amazon Web Services** June 2024 – Sep. 2024
Applied Scientist Intern Seattle, WA, USA
 - Researched the applications of Large Language Model Agents in automated code deployment.
 - Developed an LLM multi-agents framework to automate GitHub repositories deployment by enabling instruction understanding, shell script drafting, tool execution (e.g., Linux Shell, Perplexity Search), system log analysis, and issue database retrieval using Retrieval-Augmented Generation (RAG).
 - First-Authored a paper based on internship work, accepted at NAACL Main Track (Publication [C.2]).
 - Received excellent feedback rating from mentor and manager, resulting in a **return offer** for next year's internship.
- **NEC Labs America** July 2023 – Sep. 2023
Research Intern Princeton, NJ, USA
 - Conducted research on privacy-preserving fine-tuning of Large Language Models, proposing a lightweight tuning scheme that balances utility and privacy.
 - Published the work at the EMNLP 2024 main conference (Publication [C.3]).
 - Applied the method to medical scenarios, leading to a U.S. patent application filed on behalf of the company.
- **Tsinghua University** Dec. 2020 – Aug. 2021
Project Leader Beijing, China
 - Led and Pre-trained the world's 2nd-largest Protein LLM **ProteinLM** (*Wen Su*), significantly improving contact prediction accuracy from 36% to 75%. [Wikipedia Page](#) for my work, part of WuDao LLM.
 - Developed ProteinLM, a protein language model pretraining framework based on NVIDIA's Megatron-LM.
 - Managed a cluster of 504 NVIDIA Tesla V100 (32G) GPUs to pre-train *Wen Su*.
 - Model request form statistics show that ProteinLM has been used by researchers from over 20 institutions, including Harvard, University of Toronto, and NVIDIA, etc.
- **Sequoia Capital China** July 2022 – Sep. 2022
Seed Group Intern Beijing, China
 - Focused on investments in computer infrastructure, renewable energy (photovoltaic, hydrogen), and AI4Science.
 - Contributed to over 20 investment deals by assisting in market analysis, technical roadmap reviews, due diligence, team roadshows, term sheet drafting, and more.
- **Toyota Technological Institute at Chicago (TTIC)** June 2021 – June 2022
Research Assistant Chicago, IL, USA
 - Led the project *Unsupervised Contact Prediction with Large MSA Language Model*.
 - Proposed *fragment training* to balance the width and depth of Multiple Sequence Alignments (MSA), allowing pre-training of a 1-billion-parameter MSA model (the world's largest) using only 10% of the data required by previous state-of-the-art methods.
 - Achieved a 3.5% improvement over state-of-the-art results and exceeded Meta AI's performance using significantly less training data.
 - Demonstrated that increasing model scale enhances the capture of long-range dependencies in sequences.
- **Tencent Inc.** Sep. 2021 – May 2022
Research Intern Shenzhen, China
 - Developed models for protein multiple sequence alignments and localized protein database services.
 - Published research on deep-learning-powered protein-ligand docking affinity prediction under data-scarce scenarios in *Briefings in Bioinformatics*.

- [C.1] Yijia Xiao, Edward Sun, Di Luo, Wei Wang. (2025). **TradingAgents: Multi-Agents LLM Financial Trading Framework**. Accepted at *Multi-Agent AI in the Real World, AAAI 2025*, <https://tradingagents-ai.github.io/>; Financial Expert's *Summary on LinkedIn*.
- [C.2] Yijia Xiao, Runhui Wang, Luyang Kong, Davor Golac, Wei Wang. (2024). **GSR-Bench: Benchmarking Deployment in GitHub Science Repositories with Large Language Model Agents**. Accepted at *NAACL 2025*.
- [C.3] Yijia Xiao, Yiqiao Jin, Yushi Bai, Yue Wu, Xianjun Yang, Xiao Luo, Wenchao Yu, Xujiang Zhao, Yanchi Liu, Quanquan Gu, Haifeng Chen, Wei Wang, Wei Cheng. (2023). **PrivacyMind: Large Language Models Can Be Good Privacy Protection Learners**. Accepted at *Empirical Methods in Natural Language Processing (EMNLP) 2024*.
- [S.2] Yijia Xiao, Edward Sun, Yiqiao Jin, Qifan Wang, Wei Wang. (2024). **ProteinGPT: Multimodal LLM for Protein Property Prediction and Structure Understanding**. Submitted to *KDD 2025*.
- [S.3] Yijia Xiao, Edward Sun, Wei Wang. (2024). **LogicVista: Multimodal LLM Logical Reasoning Benchmark in Visual Contexts**. Submitted to *CVPR 2025*.
- [C.4] Yijia Xiao, Edward Sun, Yiqiao Jin, Wei Wang. (2024). **RNA-GPT: Multimodal Generative System for RNA Sequence Understanding**. In *Machine Learning for Structural Biology Workshop, NeurIPS 2024*.
- [C.5] Tianyu Liu, Yijia Xiao, Xiao Luo, Hua Xu, W. Jim Zheng, Hongyu Zhao. (2023). **Geneverse: A Collection of Open-source Multimodal Large Language Models for Genomic and Proteomic Research**. In *Empirical Methods in Natural Language Processing (EMNLP) 2024*.
- [C.6] Yiqiao Jin, Qinlin Zhao, Yiyang Wang, Hao Chen, Kaijie Zhu, Yijia Xiao, Jindong Wang. (2024). **AgentReview: Exploring Peer Review Dynamics with LLM Agents**. Accepted at *EMNLP 2024*.
- [C.7] Yijia Xiao, Jiezhong Qiu, Ziang Li, Chang-Yu Hsieh, Jie Tang. (2021). **Modeling Protein Using Large-scale Pretrain Language Model**. In *Pretrain Workshop, KDD 2021*.
- [J.1] Zi-Yi Yang, Zhao-Feng Ye, Yi-Jia Xiao, Chang-Yu Hsieh, Sheng-Yu Zhang. (2022). **SPLDExtraTrees: Robust Machine Learning Approach for Predicting Kinase Inhibitor Resistance**. In *Briefings in Bioinformatics*.
- [S.4] Yijia Xiao, Edward Sun, Dylan Steinecke, Alexander Russell Pelletier, Yushi Bai, Peipei Ping, Wei Wang. (2023). **Know2BIO: A Comprehensive Dual-View Benchmark for Evolving Biomedical Knowledge Graphs**.
- [C.8] Mingyu Ma, Yijia Xiao, Anthony Cuturrufo, Xiaoxuan Wang, Vijay Nori, Wei Wang. (2024). **Memorize and Rank: Elevating Large Language Models for Clinical Diagnosis Prediction**. In *AAAI 2025*.
- [C.9] Rakesh Bal, Yijia Xiao, Wei Wang. (2023). **PGraphDTA: Improving Drug Target Interaction Prediction using Protein Language Models and Contact Maps**. In *NeurIPS AI for Science Workshop 2023*.
- [S.5] Yijia Xiao, Fang Sun, Xiao Luo, Yizhou Sun, Wei Wang. (2024). **RLMC: RNA Language Model with Contrastive Learning**. Submitted to *International Conference on Intelligent Systems for Molecular Biology*.
- [P.1] Wei Cheng, Wenchao Yu, Yanchi Liu, Xujiang Zhao, Haifeng Chen, and Yijia Xiao. (2023). **Privacy Protection Tuning for LLMs in Medical Decision Making**. Patent Application No. Docket 23060.
- [C.10] Yushi Bai, Jiahao Ying, Yixin Cao, et al., Yijia Xiao, Haozhe Lyu, Jiayin Zhang, Juanzi Li, Lei Hou. (2023). **Benchmarking Foundation Models with Language-Model-as-an-Examiner**. In *NeurIPS 2023*.

HONORS AND AWARDS

- **Research Excellence Scholarship**, Tsinghua University, 2021. Ranked top 2%, 3rd out of 230 students.
- **Science and Technology Innovation Scholarship**, Tsinghua University, 2021. Ranked top 2%, 3rd out of 230 students.
- **Technology Innovation Scholarship**, Tsinghua University, 2020. Ranked top 5%, 6th out of 230 students.
- **Silver Medal**, International Collegiate Programming Contest (ICPC) Asia East Continent Final, 2020.
- **Gold Medal**, International Collegiate Programming Contest (ICPC) Asia Regional Contest, 2020.
- **First Prize**, Chinese Collegiate Physics Olympiad, 2019.
- **National Bronze Medal**, Chinese Physics Olympiad, 2017.
- **National Bronze Medal**, China Western Mathematics Invitational, 2017.

LEADERSHIP EXPERIENCE

- | | |
|--|-------------|
| • Business Development Team <i>Chinese Entrepreneurs Organization</i>
Engaged in business development activities at Stanford, CA | 2022 – 2024 |
| • Organizer <i>Tsinghua Future Internet Computing Club</i>
Organized events and workshops related to future internet computing | 2020 – 2022 |
| • Member <i>Tsinghua Entrepreneurship Training Program</i>
Selected as one of 30 members per year in Tsinghua University | 2020 – 2022 |

PROFESSIONAL MEMBERSHIPS

- **Reviewer**, NeurIPS 2023, ICML 2023, ICLR 2023, AISTATS 2024, ACM Multimedia, PKDD, PLOS ONE, etc.