Siyuan Li (黎思源)

Tel: (+86) 13815397394 | Email: 1793706453@qq.com/liaoyanqing666@gmail.com | Wechat: abcdefghi314159
Projects Website: https://github.com/liaoyanqing666 | Personal Website: https://siyuanli.tech/

EDUCATION

University of Georgia, USA 08/2025(01/2026)-06/2030
Ph.D. in Computer Science

Shanghai University, China 09/2021-06/2025
B.E. in Artificial Intelligence **GPA: 3.81/4.0 (93.60/100), Ranking: 1/52**
Key courses: Calculus(94), Linear Algebra(100), Object Oriented Programs(94), Probability and Statistics(95), Data Structure(97), Pattern Recognition(90), Computer Vision(91), Operations and Optimization(88), Data Mining and Knowledge Processing(94), Mathematical Logic(95), Principles and Techniques of Large Language Models(95), Principle and Algorithm of Artificial Intelligence(93)

RESEARCH & INTERNSHIP

Research Intern in Guangming Laboratory | Complex Medical LLM | Leader 09/2025-01/2026
Advisor: Dr. Wenhao Jiang

- Constructed a high-quality medical Instruction Tuning dataset by performing in-depth cleaning and structuring of large-scale private clinical records, and designing diverse QA templates to ensure comprehensive data coverage.
- Developed a fully automated Supervised Fine-Tuning (SFT) pipeline, streamlining the end-to-end workflow from heterogeneous data loading and dynamic prompt construction to efficient model training.
- Proposed an answer-guided self-distillation mechanism to generate Chain-of-Thought (CoT) reasoning based on correct answers. Incorporating these rationales into SFT guided the model to adopt optimal reasoning paths, significantly enhancing logical reasoning and accuracy in complex medical scenarios.
- Applied Reinforcement Learning for continuous policy optimization, further reinforcing the model's multi-step reasoning abilities and improving the accuracy of clinical decision-making.

Intern in Huawei | Department of Ascend Computing Inference Development | AI (NLP) engineer 12/2024-03/2025
Location: Huawei Lianqiu Lake R&D Center, Shanghai

- Migrate vLLM to Ascend NPU platform (vllm-ascend), responsible for unit testing and adaptation of some operators.
- Adapt speculative decoding on vllm-ascend.
- Reported and fixed a bug (issue #11978 and PR #11979) of vLLM.

Research in SHU Brain-like Computing Center Lab | AI for Recognizing Preference | Leader 04/2024-09/2024
Advisor: Prof. Huiran Zhang (Shanghai University)

- Proposed a novel ERP composite formula for analyzing human preferences.
- Achieved effective classification of preferences using AI methods combined with the developed formula.
- Authored a manuscript as the first author titled "The Study of Human Preference Based on Integrated Analysis of N1 and LPP Components", available on arXiv.

Research Intern in Westlake University | Rule Discovery in Physical Data/Video | Leader 07/2023-06/2024
Advisor: Prof. Tailin Wu (Westlake University), Prof. Sebastian Musslick (Brown University)

- Developed a transformer-based model and programmed it to experiment with symbolic regression tasks.
- Extended symbolic regression from mathematical expressions to the video domain by building a multimodal model.
- Explored the discovery of physical system patterns from videos to empower scientific discovery tasks.

Research in Shanghai University | Video Frame Interpolation with PVT | Leader 04/2023-06/2023
Advisor: Prof. Hang Yu (Shanghai University)

- Proposed a novel encoder-decoder video frame interpolation model leveraging PVT v2 as the encoder and a UNet-like decoder with deconvolution and residual concatenation.
- Achieved an SSIM of 0.9879 on Vimeo90K Dataset, surpassing state-of-the-art methods.

PROJECTS (SELECTED BY LEARNING PATH)

Project repositories: **https://github.com/liaoyanqing666** (Received **200+ stars** and **100+ forks** in total)

Fine-Tuning of Multimodal Medical Large Models Integrating the RAG Mechanism (Graduation Project) 05/2025

- Designed and implemented a medical content generation system combining Retrieval-Augmented Generation (RAG) and multimodal large language model (MLLM) fine-tuning.
- Developed a multimodal RAG framework supporting joint image-text input, featuring multiple retrieval paradigms such as joint embedding, label-guided retrieval, and image-text pair binding.
- Fine-tuned the Qwen2.5-VL model in two stages using Chinese medical QA and image-text datasets, yielding the Qwen2.5-VL-Med model with domain-specific reasoning capabilities.
- Built a modular web-based interactive system supporting local/cloud API deployment, multimodal input, streaming response, and history tracking.

Cross-Modal Pretrained Model Alignment 07/2024

- Proposed and implemented a method to quickly align pre-trained models from different modalities.
- Designed a twin neural network similarity module to align pre-trained models with varying embedding dimensions.
- Achieved rapid model alignment between text and image modalities with minimal training on a standard image classification dataset, rather than requiring a large "image-description" dataset typical for models like CLIP.
- Experimentally demonstrated the project's ability to align quickly with minimal GPU requirements and satisfactory performance.

Violent Video Classification Based on CNN-LSTM 02/2023

- Developed a network model based on CNN for video frame feature extraction and LSTM for sequential frame feature computation.
- Compared the classification performance of KNN and ANN after freezing the feature extraction model parameters.
- Achieved 92% accuracy on a public dataset, comparable to results from another study using a non-public dataset.

Handwriting Recognition System Based on Siamese Neural Networks 11/2022

- Independently designed and coded a system utilizing VGG16 for signature feature extraction.
- Achieved 100% accuracy on the CEDAR dataset using Siamese neural networks for classification.
- Developed frontend-backend interaction programs enabling the utilization of training results on web platforms.

PAPERS

Siyuan Li, Xiangze Meng, Yijian Yang, Yiwen Xu, Yunfei Wang, Chenghu Qiu, Hanyi Jiang, Pin Wu, Shengbo Chen, Xiao Wei, Hao Wang, Lan Ni, Huiran Zhang. The Study of Human Preference Based on Integrated Analysis of N1 and LPP Components[J]. arXiv preprint arXiv:2505.19879, 2025.

Yuhang Guo, **Siyuan Li (Co-first author)**, Jinxuan Wu. Research advanced in offline handwritten signature verification[J]. Applied and Computational Engineering, 2023, 6: 1236-1244.

AWARDS (SELECTED)

ASC Student Supercomputer Challenge **National Second Prize** 02/2024

ICPC (International Collegiate Programming Contest) Asia Regional Contest (Hefei) **Bronze Medal** 11/2023

ICPC (International Collegiate Programming Contest) Asia Regional Contest (Nanjing) **Bronze Medal** 11/2022

Group A of C/C++ Division of Blue Bridge Programming Cup **National Third Prize** 06/2023

Group A of C/C++ Division in Shanghai Division of Blue Bridge Programming Cup **First Prize** 04/2023

CCPC (China Collegiate Programming Contest) Shanghai Programming Contest **Silver Medal** 10/2022

SKILLS & QUALIFICATIONS

Programming language: Python (Advanced), C++ (Proficient), Matlab (Familiar), Html/Css/JavaScript(Familiar)

Software: Git&Github, Office Word, LATEX, Markdown, Remote SSH of VSCode

AI-related skills: Pytorch (Advanced), Transformers (Proficient Understanding), LLM (General Understanding), vLLM