

Zhiwei Li

EMAIL: zwli23@m.fudan.edu.cn
GITHUB: <https://github.com/digbangbang>

Shanghai, China
PHONE: (86)191-2170-2023

EDUCATION	Fudan University <i>Master of Science (M.Sc)</i> <i>Statistics</i> <i>Thesis: EFFICIENT SPARSE NEURAL NETWORK OPTIMIZATION ALGORITHM</i> <i>Advised by Weizhong Zhang</i>	Sep, 2023 – Jul, 2025 <i>GPA: 3.18</i>
	Shanghai University of Finance and Economics <i>Bachelor of Science (B.Sc)</i> <i>Statistics</i> <i>Thesis: ANALYSIS OF LOGISTIC REGRESSION ON PERFECT LINEAR SEPARABILITY</i> <i>Advised by Yang Bai</i>	Sep, 2019 – Jul, 2023 <i>GPA: 3.12</i>
RESEARCH	(1) Model (LLM) Compression & Acceleration (2) Efficient Federated Learning (3) Multimodal LLM	
PUBLICATION & PREPRINT	2024	
	Low Precision Local Training is Enough for Federated Learning <i>38th Conference on Neural Information Processing Systems</i> https://openreview.net/pdf?id=vvpewjtnvm <u>Li, Z.</u> , Li, Y., Lin, B., Jin, Z., Zhang, W.	
	Computationally and Memory-Efficient Pruning for Deep Neural Networks <i>Submitted to CVPR2025</i> Not submitted to Arxiv <u>Li, Z.</u> , Liao, Y., Wu, B., Zhou, Y., Shi, X., Jiang, D., Li, Y., Zhang, W.	
	Enhance Consistency Models via Optimizing Training Data's Marginal Distribution <i>Submitted to ICML2025</i> Not submitted to Arxiv Zhou, Y., Wu, B., <u>Li, Z.</u> , Liao, Y., Ma, Q., Wei, Z., Gao, Y., Jin, Z., Zhang, W.	
INTERNSHIP	Meituan <i>Multimodal LLM</i> <i>Researcher Supervisor: Zequn Jie</i> Reproduce MetaCLIP, build Chinese Meta Data (50K) via balanced query distribution. Fine-tune Qwen2-vl for annotation, surpassing GPT-4o on specialized data.	Jun, 2024 – Oct, 2024
	Taoli <i>Machine Learning</i> <i>Researcher Supervisor: Hui Xiao</i> Build Gaussian HMM strategy for IC/IH/IF futures using 60-min K-line indicators. Boost HMM with XGB for state estimation, likelihood improvement, slide updates.	Jun, 2023 – Sep, 2023
RESEARCH EXPERIENCE	Analysis of Logistic Regression On Perfect Linear Separability <i>Excellent Undergraduate Thesis</i> Proved logistic regression fails on linearly separable data, validated via simulation. Showed regularization prevents parameter norm growth, enabling solutions.	
	Adversarial Diffusion Distillation Acceleration Reduce sampling steps via progressive distillation, cutting reverse process to 10 steps. Apply FreeU and adversarial training to improve quality and reduce steps to 4.	

SKILLS

Languages: Chinese(native), English(fluent)

Tools & Framework: PyTorch, Transformer, Deepspeed