

XUEFENG LI

Sydney, Australia

ausmountain@gmail.com ◇ +61450306410 ◇ GitHub ◇ LinkedIn ◇ Google Scholar

Senior Research Engineer and Applied Scientist with 6+ years of experience spanning foundational ML research (weakly supervised learning, label noise) and applied agentic framework development (coding agents, model fine-tuning, and evaluation). Strong research background with publications and reviewing experience at top conferences (ICML, KDD). Recent work focuses on repository-level code modernization and rigorous benchmark and dataset creation for LLM evaluation across challenging software engineering tasks.

PROFESSIONAL EXPERIENCE

Modelcode.ai

Remote

Senior ML Research Engineer and Applied Scientist

Aug 2025 - Current

- Leading initiatives in repository-level code modernization, driving large-scale improvements in legacy codebases to align with general coding agents.
- Led research and development of agent tools for large-scale code modernization, enabling seamless coordination between AI agents and external tools. Contributing to an academic benchmark paper (RepoMod-Bench, accepted at KDD 2026) evaluating agent capabilities across 21 real-world projects totaling 1.6M lines of code.
- Expertise: Code Modernization with Coding Agents, Agent Evaluation/Benchmarking

NinjaTech AI

Sydney, Australia

Senior Applied Scientist

Jan 2024 - Aug 2025

- Lead AI agent prototype development and technical innovation, managing end-to-end delivery of production-ready prototypes from agent design to LLM fine-tuning.
- Calendar Agent: Developed an innovative ensemble method for improving date-time comprehension, compatible with the fine-tuned model. Built an end-to-end post-training pipeline including benchmark creation, synthetic dataset generation, and SFT+DPO training, boosting accuracy from 50% to 90% on internal benchmark.
- Fine-tuning: Fine-tuned Llama 405B using QLoRA on 8xH200 setup for company-specific tool calling across scheduling, image generation, and conversation tasks.
- Distillation: Built a distillation pipeline for DeepSeek-R1 by curating chain-of-thought training data from Claude Sonnet-3.5 on SWE-Bench and Aider benchmarks, improving coding performance on Aider Polyglot benchmark.
- Expertise: AI Agents, PEFT, Code Generation, Benchmark (GAIA, SWE-Bench Verified, Aider Polyglot and etc.)

Rokt

Sydney, Australia

Machine Learning Engineer

Jan 2023 - Jan 2024

- Spearheading the enhancement of Click-Through Rate (CTR) Model by integrating LLMs and multimodal models for improved prediction. Lead the project by designing, developing, and productionizing an end-to-end pipeline to fine-tune CLIP model, extract embeddings from the tuned Model, train the CTR prediction model and serve the embedding for online prediction
- Developed and implemented innovative machine learning algorithms and scalable data pipelines, addressing complex business challenges in collaboration with product, engineering, and data teams. This includes conducting exploratory data analysis (EDA), data pre-processing/engineering, offline experiments, and A/B testing
- Expertise: Recommendation Models, LLMs, Spark, Kubeflow, Tensorflow, AWS (EC2, S3 and etc.)

Amazon

Sydney, Australia

Applied Scientist, Internship

Apr 2022 - Dec 2022

- Virtual Try-On Project: Developed a novel framework for 3D novel view synthesis from a single image using Neural Radiance Fields (NeRF) and StyleGAN, which improved the inference speed by more than 100 times compared with the previous solution
- Authored a comprehensive scientific paper detailing the methodologies and breakthroughs of my proposed framework, contributing to the 3D computer vision and generative model applications
- Expertise: NeRF, 3D Computer Vision, StyleGAN, Diffusion Models, Pytorch, AWS (EC2, S3 and etc.)

University of Sydney

Sydney, Australia

Research Assistant

Feb 2019 - Dec 2021

- Focused on Weakly Supervised Learning with limited labeled data in image and text domains, developing algorithms for noisy and incomplete labels that led to top-tier conference publications
- Expertise: Weakly-supervised Learning, Label Noise, Matrix Factorisation, PyTorch

UNSW

Teaching Assistant

Sydney, Australia

Feb 2019 - Dec 2021

- Conducted teaching activities for Machine Learning, Neural Networks, Computer Vision, and Data Mining courses while assisting in course material preparation
- Expertise: Machine Learning, Deep Learning, Computer Vision, Data Mining

Toshiba

Software Engineer, Internship

Sydney, Australia

Jan 2018 - Jul 2018

- Automated testing processes and managed version control for printer driver development, creating Python test cases and collaborating with development teams on bug resolution
- Expertise: Python, Linux, Virtual Machine, Automated Testing

EDUCATION

UNSW Sydney

Ph.D. in Computer Science (Machine learning)

Sydney, Australia

Aug 2018 - Jun 2025

- Research Advisors: Alan Blair, Tongliang Liu
- Thesis: Learning Across the Supervision Spectrum: From Theory to Application

UNSW Sydney

B.Sc. (Hons) in Computer Science and Engineering

Sydney, Australia

Jul 2015 - Jul 2018

- First-Class Honours, UNSW CSE Performance Awards 4th Year (Top 10)
- Thesis: Low-rank Evolution Strategy

PUBLICATIONS

- Li, X, Liu, T, Han, B, Niu, G, and Sugiyama, M. "Provably end-to-end label-noise learning without anchor points." In International conference on machine learning, ICML 2021.
- Li, X, and Blair, A. "Eccentric regularization: minimizing hyperspherical energy without explicit projection." In International Joint Conference on Neural Networks, IJCNN 2022.
- Xue, T, Li, X, Azim, T, Smirnov, R, Yu, J, Sadrieh, A, and Pahlavan, B. "Multi-Programming Language Ensemble for Code Generation in Large Language Model." arXiv preprint arXiv:2409.04114, 2024.
- Xue, T, Li, X, Smirnov, R, Azim, T, Sadrieh, A, and Pahlavan, B. "Ninjalml: Fast, scalable and cost-effective rag using amazon sagemaker and aws trainium and inferentia2." arXiv preprint arXiv:2407.12057, 2024.
- Li, X, Ben-Israel, N, Raz, Y, Ahmed, B, Serebro, D, and Raux, A. "RepoMod-Bench: A Benchmark for Code Repository Modernization via Implementation-Agnostic Testing." *KDD 2026*.

VOLUNTARY ACTIVITIES

- Reviewer: International Conference on Machine Learning (ICML): 2022 - current
- Reviewer: Neural Information Processing Systems (NeurIPS): 2022 - current
- Reviewer: International Conference on Learning Representations (ICLR): 2022 - current

HONOURS & AWARDS

- 2018 - UNSW CSE Performance Awards 4th Year
- 2018 - UNSW Datathon First Place
- 2017 - Dean's Honours List

SKILLS

- Programming Languages: Python
- Data Science: Pandas, Numpy, SciPy, SQL, Spark
- Modelling: Pytorch, Tensorflow, Scikit-learn, HuggingFace Transformers, vLLM, FSDP
- Visualisation: w&b, Tensorboard
- Machine Learning Pipeline: Kubeflow, TFX, Docker
- AWS: EC2, S3, Athena, Redshift
- Version Control: Git
- Operating System: Linux, Mac OS