

ZHENZHUO LI

altairpaca@gmail.com · altairpaca.github.io

EDUCATION

The Hong Kong University of Science and Technology (Guangzhou)

Sep 2023 – Expected Jun 2027

BSc in Data Science and Big Data Technology

- CGA: 3.807/4.3; Major CGA: 3.837/4.3; Dean's List for three semesters
- Relevant Coursework: Advanced Algorithms and Data Structures (A+), Introduction to AI (A+), Database Management Systems (A+), Linear Algebra (A+), Design and Analysis of Algorithms (A), Machine Learning (A-), Applied Statistics (A-), Cloud Computing and Big Data Systems (A-)

Osaka University

Fall 2025

Exchange Student

RESEARCH INTERESTS

Time-Series Representation Learning; Reliable Machine Learning under Distribution Shift; Transfer and Adaptation for Non-stationary Sequential Data; Online Learning and Sequential Decision-Making.

PUBLICATIONS

- Yuanjian Xu, Jianing Hao, Anxian Liu, **Zhenzhuo Li**, Shichang Meng, Siyuan Yuan, and Guang Zhang, "LENS: Large Pre-trained Transformer for Exploring Financial Time Series Regularities," in *Proceedings of the 6th ACM International Conference on AI in Finance (ICAIF 2025)*, pp. 771–778, 2025. DOI: 10.1145/3768292.3770349.

RESEARCH EXPERIENCE

LENS: Large Pre-trained Transformer for Financial Time Series

Jun 2024 – Oct 2024

Undergraduate Research Assistant / Co-author

Supervisor: Prof. Guang Zhang, HKUST(GZ)

- Developed parallel and memory-efficient preprocessing pipelines for terabyte-scale financial time-series data, supporting Transformer pretraining through chunk-based cleaning, aggregation, and model-input construction.
- Achieved up to $13\times$ faster data cleaning and $6\times$ faster statistical aggregation in internal benchmarks by restructuring data access and parallel processing.
- Contributed data-processing and experimental infrastructure to the resulting paper published at ACM ICAIF 2025.

SpatialReflect: Spatially Disentangled Reflection for Controllable Generation

Jun 2025 – May 2026

First-author Independent Research

Supervisor: Prof. Zeke Xie, HKUST(GZ)

- Formulated failure modes in multi-region diffusion-based generation and developed a training-free framework for spatially controlled synthesis.
- Designed controlled experiments and completed a full first-author manuscript, gaining end-to-end experience in problem formulation, method development, implementation, evaluation, and scientific writing.

Final Year Project: Reliable Multi-Agent Financial Forecasting (Working Title)

Jul 2026 – Present

Undergraduate Researcher

Supervisor: Prof. Hao Xue, HKUST(GZ)

- Developing a hybrid forecasting framework that uses LLM agents to transform unstructured information into structured evidence for downstream statistical and economic decision models.
- Studying evidence aggregation, reliability evaluation, and calibration in multi-agent forecasting under noisy and dynamically changing information.

INDUSTRY RESEARCH EXPERIENCE

Xuanyuan Private Fund Management (Guangdong) Co., Ltd.

Jul 2026 – Present

Quantitative Research Intern | Guangzhou, China

- Conducted end-to-end research on time-series factor evaluation and market-timing models using Level-2 equity-market data, covering data construction, model development, signal evaluation, and decision-oriented analysis.
- Built scalable pipelines for multi-year, market-wide minute-level panels, including temporal and cross-asset alignment, memory-mapped storage, parallel I/O, sample filtering, and model-ready dataset construction.
- Developed leakage-aware evaluation procedures for intraday and cross-day signals, including information-coefficient analysis, temporal-stability diagnostics, entry-timing evaluation, and sample-selection audits.
- Investigated tree-based and neural sequence models, including LightGBM, XGBoost, and GRU, for market-exposure and intraday-entry decisions under non-stationary conditions.

Shanghai Yunyan Private Fund Management Co., Ltd.

May 2026 – Jul 2026

AI Investment Research Intern

- Conducted AI-assisted research on semiconductor and AI-infrastructure companies by synthesizing company disclosures, industry research, expert commentary, and open-source information.
- Designed LLM-agent workflows for information screening and multi-source verification under noisy and rapidly changing information.

SELECTED AWARDS & HONORS

- ASC World University Supercomputer Competition 2026 — **First Prize**, Team Member
- HKUST(GZ) TradeMaster Cup Quantitative Trading Competition 2026 — **5th Place**, Team Leader
- INFH Lizhi Innovation Talent Scholarship, 2025
- MCM/ICM Mathematical Contest in Modeling 2025 — Honorable Mention
- ASC World University Supercomputer Competition 2025 — Second Prize, Team Member
- Dean's List — Three Semesters
- Admission Scholarship, First Class, HKUST(GZ), 2023

RESEARCH & TECHNICAL SKILLS

Agentic AI Systems & Research Automation: Practical experience deploying and operating coding and research-agent environments; agent-harness design and multi-agent orchestration built on mainstream patterns — ReAct-style reasoning loops, plan-and-execute and orchestrator-worker structures, evaluator-optimizer and self-reflection workflows, tool use and Model Context Protocol (MCP) integration, multi-model routing and fallback, context and memory management, automated evaluation, and failure recovery

HPC & Model-Training Optimization: Python, C++, Rust, SQL; Linux, Git, Slurm, CUDA, and MPI; profiling and optimization of model-training and data pipelines, including data loading, parallel I/O, memory usage, cluster execution, experiment automation, and reproducible benchmarking

Deep, Representation & Transfer Learning: PyTorch; Transformer, GRU, and diffusion models; custom data pipelines, masking and normalization, pretrained-model reuse and adaptation, ablation studies, hyperparameter tuning, and training diagnostics

Time-Series & Reliable Machine Learning: LightGBM, XGBoost, and scikit-learn; multivariate sequence modeling, leakage-aware dataset and label construction, walk-forward evaluation, distribution-shift and stability analysis, positive controls, and decision-oriented validation

Reinforcement Learning & Sequential Decision-Making: Source-level implementations of bandit, tabular, and deep RL algorithms, including UCB, Q-learning, DQN, REINFORCE, A2C, and DPG; MDP and bandit formulation, state-action-reward design, Bellman methods, policy gradients, actor-critic training, experience replay, target networks, importance sampling, and literature familiarity with distributional and risk-sensitive RL

LANGUAGES

Chinese: Native; English: Proficient, undergraduate instruction in English; Japanese: JLPT N1