Zihan Wu

PhD Candidate, Electrical Engineering, City University of Hong Kong

✉ wzh4464@gmail.com | ☎ (+852) 9810 6427 | 🌐 https://scholar.zihanng.shop | in zihan-wu-ustc

SUMMARY

Ph.D. candidate with dual B.Sc. in Physics and Mathematics, specializing in machine learning on feature selection and data valuation. Expert in reinforcement learning for trading systems, statistical modeling, and high-performance computing. Demonstrated success in cryptocurrency trading (Sharpe 1.42) using self-developed RL algorithms. Published in top-tier venues (CIKM, AISTATS) with focus on data valuation and unsupervised machine learning. Passionate about developing innovative quantitative strategies combining rigorous mathematical foundations with cutting-edge AI techniques.

EDUCATION

City University of Hong Kong   **Hong Kong SAR, China**
Ph.D. in Electrical Engineering   2020 – 2025

University of Science and Technology of China   **Hefei, China**
Dual B.Sc. Degrees in Physics and Mathematics & Applied Mathematics, School of the Gifted Young   2015 – 2020

EXPERIENCE

Huawei Technologies, AI Research Intern   **Hong Kong**
Optimized LLM prompts for automated code review using RAG and context engineering. Enabled accurate, low-latency analysis by dynamically injecting relevant codebase context.   Aug 2025 – Present

PROJECTS

Python, PyTorch, PPO, Rust, Binance API   BTC Trader: Reinforcement Learning for Crypto Trading (Independent)
2023 – 2025

- Developed reinforcement learning models for decision-making under uncertainty in volatile crypto markets
- Implemented PPO-based agents for price prediction and signal generation with comprehensive risk management
- Designed novel algorithms for data efficiency in ML, achieving 4.08% return with Sharpe ratio 1.42
- Backtested on 3 years of minute-level data across 15 trading pairs
- Developed custom feature engineering based on co-clustering
- Built low-latency trading infrastructure enabling real-time reinforcement learning deployment

C++, Distributed Systems, Queuing Theory, Performance Optimization   X-Shard: High-Frequency Distributed Transaction Engine
IEEE TPDS 2024   2023 – 2024

- Developed distributed transaction processing system achieving 37% latency reduction
- Applied queuing theory (M/M/1 Jackson networks) to model system performance, proved with linear throughput scaling
- Designed graph partitioning algorithm (Metis-based) reducing cross-shard transactions by 80%
- Implemented optimistic concurrency control enabling parallel transaction processing at 3k+ TPS with *ms* commit latency
- Built threshold signature scheme reducing network overhead from $O(n^2)$ to $O(n)$, essential for multi-venue order routing

C++, Rust, MPI   **High-Performance Co-Clustering System**
IEEE SMC 2024 & IEEE TIM   2020 – Present

- Co-clustering (biclustering) simultaneously groups rows (e.g., assets/trades) and columns (e.g., factors/time windows) to find coherent blocks; enables automatic feature discovery for finance data mining (market regimes, asset–factor segments, anomaly patterns)
- Developed scalable statistical co-clustering algorithms for large datasets (800K+ corpus)
- Achieved 83% runtime reduction via algorithmic optimization
- Applied to crypto market: identified 8 regime clusters (risk-on/off, momentum/mean-reversion)
- Feature discovery: automatically extracted 47 alpha factors from 200+ raw market indicators

Python, PyTorch, Data Mining, DCAI   **LiveVal: Novel Algorithms for Data Efficiency in ML**
CIKM 2025   2023 – Present

- Architected LiveVal, a novel time-aware data valuation framework using adaptive reference point mechanism and dual-queue system for real-time ML analysis
  - Developed foundational AI research contributions achieving 180x speedup and 92% memory reduction over LOO methods
  - Demonstrated scalability on large models, enabling efficient data valuation for large-scale ML enablement

Python, PyTorch, LLM Fine-tuning, Model Optimization   LMEraser: Efficient Unlearning System for Large Models
AISTATS 2025   2023 – 2024
- Built an exact unlearning system via prompt tuning to cut data removal costs for large models (86M ViT)
- Developed novel algorithms for efficient model updates, slashing data removal computation by >100x vs. retraining
- Enabled privacy-preserving AI through efficient model governance and ethical AI practices

SELECTED PUBLICATIONS

- LiveVal: Time-Aware Data Valuation for Real-Time ML Systems, CIKM 2025 (Core A)
- LMEraser: Large Model Unlearning via Adaptive Prompt Tuning, AISTATS 2025 (Core A)
- X-shard: Optimistic cross-shard transaction processing for sharding-based blockchains, IEEE Transactions on Parallel and Distributed Systems (TPDS) 2024 (CCF-A)
- Scalable Co-Clustering for Large-Scale Data through Dynamic Partitioning and Hierarchical Merging, IEEE SMC 2024 (CCF-C)
- Ellipse Detection via Global Arc Compatibilities and Adaptive Co-Clustering for Real-World Measurement Systems, IEEE TIM 2025 (Q1)
- A convex-hull based method with manifold projections for detecting cell protrusions, Computers in Biology and Medicine 2024 (Q1)

SKILLS

AI Research: Reinforcement Learning, Unsupervised/Self-Supervised Learning, Optimization, Generative Models, Data Valuation, LLM Fine-tuning, Prompt Tuning
Quant Finance: Time-series, Stochastic Processes, Risk Management
Option Pricing Math: Itô Calculus, Stochastic Differential Equations, Girsanov Theorem, Martingale Pricing, Black–Scholes–Merton, Greeks, Monte Carlo, Finite-Difference PDE
Programming: Python, C++, Rust, CUDA
ML Frameworks: PyTorch, TensorFlow, Scikit-learn
Systems: Distributed Computing, HPC, Real-time Processing, Docker
Systems Engineering: Concurrency Control, Consensus Protocols (PBFT), Network Optimization, Latency Profiling, Throughput Analysis
Mathematical Modeling: Queuing Theory (Jackson Networks), Graph Theory (Balanced Partitioning), Performance Analysis, Asymptotic Complexity

AWARDS

Hong Kong PhD Fellowship: *Top 5% acceptance rate*   2020–2024
National Encouragement Scholarship: *Top 3% nationwide*   2017–2018
Patent: Physical Activity Assessment System: *Patent ID: HK30081186*   2023