Yingjie Bai

yjbai819@gmail.com | ybai0439@uni.sydney.edu.au | yingjiebai.xyz

Data valuation, robot learning, and AI agents. I study how learning systems value, select, and trust information, drawing on economics and the social sciences. Honours student in computer science and financial economics at The University of Sydney.

Education

The University of Sydney (USYD) Feb 2023 - Nov 2026 (expected)

Bachelor of Advanced Computing (Honours), Sydney, Australia.
Double major: Computer Science & Financial Economics. Honours supervisor: Dr. Weiming Zhi.

University of California, Los Angeles (UCLA) Jun 2026 - Aug 2026
Selected to represent **The University of Sydney's United States Studies Centre** at UCLA Summer Sessions.
COMM 109 - Entrepreneurial Communication; SOCIOL 147A - Sociology of Crime.

National University of Singapore (NUS) Sep 2025 - Dec 2025
Exchange student. Recipient of The University of Sydney's Vice-Chancellor's Global Mobility Scholarship.

Accepted Publications

First author on all works. Related full and workshop versions are identified below.

[1] **Yingjie Bai**, Weiming Zhi. *When Does Data Value Reduce to Class Balance? A Coverage View of Per-Point Data Valuation.* **Transactions on Machine Learning Research (TMLR), 2026. Accepted.** Full version of [3]. OpenReview | Project page.

[2] **Yingjie Bai**, Weiming Zhi. *Risk-Aware Data Auditing for Resource-Efficient Learning under Distribution Shift - Data-CAPM: A Group-Level Factor Decomposition.* **KDD 2026 Workshop on Resource-Efficient Learning for Knowledge Discovery (RelKD 2026). Accepted.** Full preprint | Workshop paper.

[3] **Yingjie Bai.** *Whose Geometry? When Learner-Relative Data Selection Beats Input-Space Selection for Efficient Fine-Tuning.* **KDD RelKD 2026 workshop. Accepted.** Short version; extended in [1]. Workshop paper.

[4] **Yingjie Bai**, Weiming Zhi. *A Verifier Validated on a Proxy Can Reverse in Deployment: The Signed Effect of Audit Certainty.* **NeurIPS 2026 Workshop Who Verifies the Agents? Toward Reliable Agent Development. Accepted (Poster).**

[5] **Yingjie Bai**, Weiming Zhi. *The Memory of Failure Should Be Lossy: Wrongful Convictions in Shared Failure Banks for Autonomous Research.* **NeurIPS 2026 Verify-Agents Workshop. Accepted (Poster).**
Related workshop version: *The Optimal Memory of Failure Is Lossy: Wrongful Convictions in Shared Failure Banks for Autonomous Research.* **NeurIPS 2026 AI for Science Workshop. Accepted (Poster).** OpenReview.

Submitted Manuscripts

[6] **Yingjie Bai**, Weiming Zhi. *Audit Certainty Is a Signed Lever: Why a Verifier Calibrated on a Proxy Can Backfire in Deployment.* **Submitted to ICLR 2027.** Extended version of [4].

[7] **Yingjie Bai**, Weiming Zhi. *When Failure Banks Kill Good Ideas: Calibrated Negative Memory for Autonomous Research.* **Submitted to ICLR 2027.** Extended version of [5].

Research Experience

PAIR Lab, The University of Sydney Current
Honours student, supervised by Dr. Weiming Zhi. Research on data valuation and selection, robot learning, and verification and memory for AI agents.

Lin Shao Group, National University of Singapore Aug 2025 - May 2026
Research Intern. LLM-based task and motion planning for dual-arm and multi-arm robots; benchmark development for coordination and control in dynamic environments.

Songfang Huang Group, Peking University Nov 2024 - Feb 2025
Research Intern. Classical, graph, and agentic RAG for an energy-sector application with ENN Group, including local model inference, databases, and an interaction interface.

Kun Tang Group, Tsinghua University Jul 2023 - Sep 2023
Research Intern. Aligned multi-source data and extracted satellite-imagery features to analyze regional health disparities.

Selected Research Contributions

Data selection under limited budgets and distribution shift [1-3]

Characterized why fixed per-point rankings miss set-level coverage, and when class balance recovers it: labels must identify the learner's modes. Developed Data-CAPM to separate robust contribution from shortcut-driven value. On synthetic and Adult shortcut benchmarks, alpha-based acquisition improved robust-average balanced accuracy by **10.4 percentage points at k = 2**.

Demonstration selection for imitation learning

Built a reproducible Minari/D4RL evaluation framework for reliability-gated latent coverage. Initial PointMaze experiments improved paired rollout return by approximately **21-24** over random and quality-diversity baselines, while exposing a mismatch between validation action MSE and closed-loop performance. Ongoing work.

The economics of verification and memory [4-7]

Investigating how audit timing and mistaken failure judgments change agent behavior. Across 12 deployed LLMs, agents made **0/94** free-choice revisits of DEAD-marked ideas; evidence-graded lossy memory recovered **24-35%** more truth-validated discoveries than binary locks in the evaluated settings. Workshop versions accepted; extended manuscripts submitted to ICLR 2027.

Selected Projects & Engagement

Lumen Agora. Initiator of an open-source, non-profit academic exchange platform; a six-person early team working on research discovery and knowledge exchange.
SO-101 robot arm. Assembled and calibrated a LeRobot-based manipulator for teleoperated demonstration collection and imitation-learning experiments.
Kent Ridge Ministerial Forum (2025). Asked former Singaporean PM Lee Hsien Loong about the societal structure of an AI-driven future.
International Relations Department Lead, Sydney University CSSA (2024-2025). Organized English Corner events with 50+ participants and coordinated cultural outreach.

Academic Service

Reviewer, NeurIPS 2026 Workshop AI4Science.
Reviewer, NeurIPS 2026 Workshop Verify-Agents.

Skills

Research: data-centric ML, data valuation and selection, robustness, imitation learning, multi-agent systems, AI safety and agent verification; econometrics and quantitative finance.
Tools: Python, R, C, Java, Git, Linux (self-hosted Pi server), agent frameworks, LeRobot / SO-101.