

Guanli Liu

✉ guanli1@unimelb.edu.au · ☎ (+61) 450075953 · in [Guanli Liu](#) · 🌐 <https://github.com/Liuguanli>

👤 Profile

Data scientist with a PhD in AI4DB, specializing in probabilistic modeling, GenAI-assisted analytics, and production-grade experimentation. I translate messy behavioral data into ranking, risk, and recommendation models for teams in Australia and the US.

⚙️ Analytics & AI Skills

- **Modeling:** Causal inference, gradient boosting, sequence models, RL for tuning, LLM/RAG prompt engineering
- **Tooling:** Python (pandas, PyTorch, TensorFlow, scikit-learn), BigQuery, dbt, SQL, Spark, Airflow
- **Experimentation:** A/B testing, Bayesian optimization, feature stores, ML monitoring, bias/fairness analysis
- **GenAI:** GPT-4/Claude APIs, embeddings, retrieval pipelines, safety filters, synthetic data generation
- **Visualization & Storytelling:** Looker, Tableau, Plotly, executive dashboards, product requirement writing

🎓 Education

PhD in Computer Science, <i>The University of Melbourne</i> , Australia	2019 – 2023
M.S. in Computer Technology, <i>Northeastern University</i> , China	2013 – 2015
B.Eng. in Software Engineering, <i>Northeastern University</i> , China	2009 – 2013

🧑‍💻 Experience

Postdoctoral Research Fellow (AI4DB), University of Melbourne	2024 – Present
--	----------------

- Lead modeling for spatial index benchmarking: trained RL agents that beat hand-tuned baselines by 18% throughput while reducing maintenance cost.
- Built GenAI copilots for query engineers that summarize execution plans and suggest optimizer hints using GPT-4 + domain finetuning.
- Mentor MS students on experimentation hygiene, feature documentation, and reproducibility pipelines.

Data Scientist, nftDb 🌐, Remote/Australia	2023 – 2024
--	-------------

- Modeled anomalous wallet behaviors using PageRank-like graph scores + LightGBM; recall +22% vs. heuristic baselines over 1.4B transactions.
- Designed dbt modeling layers + Looker dashboards that exposed wash-trading KPIs to product and compliance stakeholders weekly.
- Shipped GenAI wallet briefing notes (GPT-4 + custom embeddings) that auto-generate collector summaries, adopted by BD and community teams.

Research Assistant / Data Scientist, University of Melbourne	2019 – 2023
---	-------------

- Developed a coding-style checker (Python, AST analysis + GenAI repair suggestions) improving student error detection by 35%.
- Instrumented reading-interruption study pipelines: Node.js backend, GCP infra, usage analytics, and Mid-journey/GPT content workflows.

Software Engineer, Baidu, China	2015 – 2017
--	-------------

- Partnered with data scientists to add telemetry hooks and experimentation switches across the InfoFlow messaging stack (Java, Android).
- Implemented on-device ranking features and logging that powered personalization models for 40K+ enterprise users.

Publications & Impact

- **VLDB 2024/2025:** Efficient cost modeling of space-filling curves powering adaptive storage layouts.
- **ICDE 2023:** Efficiently learning spatial indexes; toolkit open-sourced and reused by multiple labs.
- **ICDEW 2023:** Model reuse/fine-tuning for learned indexes, highlighting rapid experimentation workflows.