Zijun (Brighton) Liu

San Diego, CA | liuzijun6688@gmail.com | 530-220-8681
github.com/BrightonLiu-zZ | brightonliu-zz.github.io | linkedin.com/in/zijun-brighton-liu-2b819b331

Seeking a Summer 2027 ML internship - sequence models | LLM post-training | production ML

Education

University of California, San Diego - B.S. Astrophysics, Data Science minor | *Jun 2027* | **GPA 3.92**
Available for a Summer 2027 internship, Jun - Sep 2027.

Projects

Self-supervised learning on stellar light curves | *First author, with Y. Ma and Prof. C. Theissen* | *2026 - present* | code

- Pre-trained a **self-supervised** conv encoder (GRU latent prediction, **PyTorch**) on unlabelled TESS light curves.
- Embeddings + 25 engineered features improved probes on **7 of 10 tasks**, beating end-to-end Conv1D and MLP.

KataGo x LLM - LLM post-training & edge deployment | *Project lead, volunteer team* | *2025 - present* | code | demo

- Post-trained **Qwen3-8B** (LoRA, **SFT -> GRPO**, HF TRL) on 302k engine-verified examples to play and explain Go.
- Deployed it **4-bit on an 8 GB GPU** (llama.cpp, ~42.5 tok/s); served base + 2 LoRA adapters from one **vLLM** instance.
- Built the evaluation (board-perception probe, refereed engine ladder) and stabilized **GRPO** against reward hacking.

Experience

Rothenberg Wealth Strategies - ML Research Intern | *Jan - Mar 2026, Jun - Sep 2026* | San Diego

- Trained a **LightGBM** model and shipped it as an unattended **production service** with a FastAPI dashboard.
- Built a multi-worker experiment runner (**1,000+ runs per sweep**) for feature ablation and walk-forward validation.
- Automated daily data ingestion, feature build and inference, with duplicate-safe API calls and daily reconciliation.

XView LLC - AI Engineer Intern | *Apr - Jun 2026* | Remote

- Built a 5-agent **RAG** pipeline (CrewAI, Chroma); prompt A/B and chunking evaluation cut factual errors by **~60%**.
- Set up the evaluation loop that brought the human revision rate of generated drafts to **~30%**.

Himiway Intelligent Technology USA - Data Science Intern | *Sep - Dec 2025* | Hybrid

- Trained and deployed an **XGBoost** model for store-location and inventory decisions on a SQL/Python ETL codebase.

Publications

Liu, Z., Ma, Y., Theissen, C. *What does self-supervision add to engineered features for stellar light curves?*
Accepted at the ML4PS workshop @ NeurIPS 2026.

Skills

ML: PyTorch, HF TRL, SFT / GRPO / LoRA, self-supervised learning, LightGBM, XGBoost, scikit-learn, RAG
Engineering: Python, C++, SQL | vLLM, llama.cpp, FastAPI, Docker, Linux, AWS, Git, Weights & Biases