

Berwyn Berwyn

☎ 650-661-7197 ✉ berwyn@stanford.edu [in berwynb](https://www.linkedin.com/in/berwynb) [GitHub Berwyn19](https://github.com/Berwyn19) [🌐 berwynb.vercel.app](https://berwynb.vercel.app)

Education

Stanford University

Expected June 2028

Bachelor and Master of Science in Computer Science (GPA: 3.80 / 4.00)

Stanford, CA

- **Relevant Coursework:** Machine Learning for Algorithmic Trading, Time Series Analysis, Design and Analysis of Algorithms, Programming Abstractions (C++), Natural Language Processing with Deep Learning, Deep Learning for Computer Vision, Machine Learning, Reinforcement Learning

Experience

The Prompting Company

Jun 2026 – Present

Software Engineer

San Francisco, CA

- Built a Go CLI tool with 15+ subcommands for querying AI bot traffic analytics and managing prompts, with paginated results, client-side filtering, and multi-format output (JSON/CSV/table)
- Built an end-to-end AI visibility pipeline tracking brand surface rate across ChatGPT, Perplexity, and Gemini; orchestrated fan-out runs via Temporal workflows and aggregated Share-of-Voice and citation metrics through Tinybird, cutting report generation from 3 hours to under 5 minutes

Telkomsel

Jun 2025 – Aug 2025

Machine Learning Engineer

Jakarta, Indonesia

- Designed a hybrid model-serving architecture integrating in-house inference services with AWS compute and Google Analytics data pipelines, replacing Vertex AI and reducing annual infrastructure cost by **90%**
- Processed over **10 billion clickstream records** in BigQuery, joining heterogeneous datasets to construct reliable large-scale training data pipelines
- Built end-to-end machine learning pipelines using CatBoost including feature engineering, training, evaluation, and hyperparameter tuning for personalized search ranking

Stanford University, Department of Materials Science and Engineering

Jun 2024 – Aug 2024

Machine Learning Research Assistant

Stanford, CA

- Constructed a machine learning pipeline in Python (PyTorch) and C++ to demonstrate that lightweight, interpretable Moment Tensor Potentials (MTP, implemented in C++) can achieve **90%** of the accuracy of state-of-the-art equivariant neural networks (trained in PyTorch) for interlayer interactions using a split-layer dataset approach
- Co-authored a peer-reviewed paper introducing a novel split-model framework and holistic validation metrics, achieving **10x increase** in energy and force prediction accuracy over conventional models [Paper]

Projects

memflux | Go, gRPC, Protocol Buffers

- Built a distributed in-memory cache from scratch featuring TTL-based eviction, gRPC replication, and Raft-inspired leader election with automatic failover under 500ms
- Engineered concurrent store achieving 56M reads/sec single-threaded and 3.9M ops/sec under parallel mixed load with zero allocations per read
- Implemented two-phase sweep algorithm minimizing write-lock contention during TTL eviction and Raft leader election with randomized timeouts and term-based authority

Buzzly | Python, Firebase Storage, Docker, Cloud Functions, React, Tailwind, OpenAI API

- Engineered Python backend services and HTTP APIs for an AI-powered marketing content generator using the OpenAI Agents SDK, implementing custom orchestration logic, extensible handlers, and safety guardrails to support scalable automated generation.
- Designed and deployed a fault-tolerant, distributed asynchronous multimedia pipeline integrating external HTTP APIs (Heygen, DALL-E, Runway ML, Whisper), with robust retry logic and production deployment via Docker and GCP Cloud Functions.

Skills

Languages: Python, C++, C, Go, JavaScript, SQL, TypeScript

Technologies: Linux, React.js, Next.js, AWS, Docker, Firebase, PostgreSQL, PyTorch, Langchain, Tailwind, gRPC, Protocol Buffers, Kafka