

Gahn Wuwong

Software Engineer Intern @ HelixDB YC25

(858) 666-5628 | gahnxd@gmail.com | linkedin/gahn-wuwong | gwuwong.com

EDUCATION

University of California San Diego

B.S. Data Science & Business Analytics Minor, 3.98 GPA, Provost Honors

San Diego, CA

2022 - 2026

EXPERIENCE

Software Engineer Intern

June 2025 - Sept 2025

HelixDB YC25

- Enhanced Model Context Protocol (MCP) infrastructure in Rust and built an MCP server in the Python SDK, enabling Pydantic type-safe tool-calling for AI agents to traverse the database, accelerating development at scale.
- Engineered a high-throughput pipeline to ingest 288k Hacker News posts and 1.3M comments with parallel processing and vector embedding workflows to power a large-scale production knowledge graph.
- Designed a hybrid retrieval pipeline on AWS Lambda with semantic search, graph traversal, and hierarchical multi-agent orchestration to enhance reasoning depth, retrieval accuracy, and speed.
- Implemented a high-performance codebase indexing system in Rust, leveraging Tree-sitter for recursive parsing and concurrent API requests, enabling better embedding throughput, ingestion scalability, and agentic retrievability.

Data Science Intern

June 2024 - Sept 2024

Kasikorn Business Technology Group

- Engineered an automated forecasting platform, leveraging statsmodels and Scikit-learn, with interactive dashboards, driving data-driven decision-making for SMEs and reducing front-end development costs by 20%.
- Developed a modular pipeline in Python to handle diverse forecasting workloads and evolving business needs.
- Implemented advanced model evaluation and deployment workflows, selecting top-performing machine learning models to deliver highly accurate, production-grade forecasts.

Instructional Assistant

Jan 2024 - April 2025

Hahciroğlu Data Science Institute

- Instructed and mentored 300+ students in data structures, algorithms, and Java/Python programming, strengthening problem-solving skills and technical proficiency.
- Optimized and maintained automated grading systems in Python and Java; streamlined course operations, graded exams, and addressed student inquiries.

PROJECTS

Professor Network Search | *Agentic Workflows, LangGraph, Next.js, RAG, FastAPI*

April 2025 - May 2025

- Built an AI-powered academic discovery platform with a Next.js frontend (Vercel) and FastAPI backend (Digital Ocean), leveraging Gemini LLMs and RAG with Qdrant vector stores to match students with professors.
- Engineered a modular backend pipeline with Supabase Postgres for chat history, Cloudflare R2 for storage, and an Exa-powered fallback search engine to ensure reliability and low-latency performance.
- Implemented a robust data ingestion and retrieval workflow for professor data using Crawl4AI and Browser Use, integrating embeddings for high-accuracy responses and research summarization.

Inventory Forecast Case Study (Deloitte) | *ML, Timeseries Forecasting, Data Analytics*

April 2024 - June 2024

- Built machine learning forecasting models to predict monthly sales trends and optimize inventory levels across multiple product categories, driving supply chain efficiency, cost reduction, and data-driven decision-making.
- Performed EDA to identify trends and high-impact products, driving feature engineering, higher predictive accuracy, and actionable inventory insights.

SKILLS

Programming Languages: Python, Java, SQL, TypeScript, JavaScript, Rust

Frameworks & Tools: PyTorch, Pandas, Pydantic, LangGraph, Postgres, Scikit-learn, Statsmodels, Next.js, FastAPI

Cloud & Infrastructure: Docker, AWS Lambda, Cloudflare R2, Supabase, Qdrant, Vercel, Digital Ocean

Domains: Machine Learning (ML), Timeseries Forecasting, Retrieval-Augmented Generation (RAG), Agentic Workflows, Data Analytics, Vector & Graph Databases, Full-Stack Development