

# Prasham Shah

San Diego, CA | 626-734-8075 | prs008@ucsd.edu | linkedin.com/in/prashamshahh | www.prasham.ai

## Education

---

**University of California, San Diego**

*B.S. in Computer Science | GPA: 3.75 | Provost Honors*

**Sep 2023 – Jun 2027**

*San Diego, CA*

**Relevant Coursework:** Advanced Algorithms, Operating Systems, Deep Learning, NLP, Quantitative Finance

## Experience

---

**Founder & Software Engineer**

*Novalo*

**Aug 2025 – Present**

*San Diego, CA*

- Building an **AI-first financial intelligence product** for retail investors across distributed market data ingestion, **ClickHouse**-backed analytics, Hono/TypeScript APIs, Supabase state, and React/TanStack surfaces for options flow, equity research, charts, macro, positions, and notifications.
- Designed the AI agent layer with read-only market tools, source-freshness checks, ticker snapshots, event retrieval, chart rendering, and **MCP/DataPlat** integrations to ground answers in live data and caveats.
- Implemented backend systems for page-aware **AI memory/context**, watchlists, user activity, daemon-driven market events, relevance scoring, and portfolio/thesis/watchlist notification pipelines.

**Software Engineering Research Assistant**

*UCSD DigiHealth Lab*

**Jun 2025 – Present**

*San Diego, CA*

- Developed **secure full-stack infrastructure** for clinical AI research with authenticated workflows, validation layers, and audit-conscious handling of sensitive data.
- Deployed and performance-tuned backend services for **LLM-based triage tools** on dedicated servers, achieving **sub-1.2s median latency** under concurrent load.
- Served fine-tuned open-source models via **vLLM**, optimizing inference pipelines to reduce compute cost by **80%** versus API-based alternatives.

**Software Engineer & Technical Stakeholder**

*BuyFineDiamonds / Amada Diamonds*

**Sep 2024 – Sep 2025**

*London, UK / Remote*

- Built a luxury ecommerce storefront from scratch, owning catalog architecture, responsive shopping flows, search/discovery, product detail pages, checkout-adjacent integrations, and backend systems from zero to **\$500K ARR**.
- Integrated **AI-assisted shopping** for product discovery, recommendations, and high-intent search, helping customers navigate **100K+ diamond SKUs** with clearer intent matching.
- Engineered backend APIs, ranking/search logic, caching, and database/query optimization for fast, reliable high-value shopping flows across a large jewelry catalog.

## Projects

---

**OpenCL Accel. for Financial Time Series, CNN Workloads | C++, Python**

**Oct 2025 – Mar 2026**

- Built **GPU-accelerated OpenCL pipelines** for financial time-series feature generation and CNN-style workloads, parallelizing rolling-window statistics, convolution, and softmax across large batched datasets.
- Optimized kernels using **tiling, local memory, coarsening**, and reduced host-device transfer overhead, achieving up to **10x speedups** over CPU baselines on matrix-heavy workloads.
- Analyzed **arithmetic intensity, memory traffic**, and kernel bottlenecks using FLOP and memory-access models to guide performance tuning and architecture-aware optimization.

**Real-Time Analytics Platform | JavaScript, Node.js, PostgreSQL, Nginx, Docker**

**Jan 2026 – Mar 2026**

- Built an end-to-end analytics platform from scratch with a **client-side collector**, backend ingestion and reporting services, and dashboards for **traffic, event, and behavioral** analysis.
- Instrumented a live website to capture **page views, sessions, custom events, performance signals**, and request metadata, then processed and stored the data through a production-style backend pipeline.
- Deployed the system on real infrastructure with **Nginx**, domain routing, and server logging, demonstrating full ownership from collection and transport through storage and visualization.

## Technical Skills

---

**Languages:** Python, TypeScript, JavaScript, SQL, C, C++, OpenCL

**Product Engineering:** React, Next.js, TanStack Router, Hono, FastAPI, Node.js, REST APIs, WebSockets, Git, GitHub

**Data & Infrastructure:** PostgreSQL, ClickHouse, Redis, Supabase, Docker, Nginx, Linux, AWS, Vercel, background jobs, daemons

**AI Systems:** AI agents, MCP, tool-calling workflows, RAG, retrieval/search, model routing, AI memory/context, vLLM

**Testing & Core:** Vitest, Hardhat, DS&A, system design, distributed systems, caching, parallel computing