Mohnish Behera

704-778-9364 | mohnishb07@gmail.com | LinkedIn | GitHub

Education

University of California, Berkeley

Bachelor of Arts in Cognitive Science, Data Science

- **Activities and Clubs:** Datagood, Claude @ Cal, Cursor @ Cal, BASIS (Berkeley AI Safety Student Initiative) Paper Reading Group, Intro to Quant Finance

Experience

Software Engineer at DigiSignals Inc. **2026 – Present**

- Co-built an AI-compute market analyst agent for a client's executives (FastAPI, Claude on Azure AI Foundry, MCP, pgvector, Next.js): second-most merged pull requests on the team (118 of 443) and 800+ tests written, after building its Phase 0 demo features in 3 days (60 commits).
- Rebuilt production in the client's own Azure tenant and ran the cutover: wrote the Terraform for the stack (100+ managed resources) and cross-tenant copy tools that moved 1,195 Cosmos DB documents, about 5 GB of blobs and about 305K vector chunks.
- Added job, dependency and serving health checks and dead-man switches to the team's health monitor (170+ tests) after a roughly 18-hour chat outage went undetected and an audit found one crashed detector could silently drop 22 of 25 checks; replaced per-check alert emails with one daily digest.
- Fixed an eval gate that tolerated two regressions before blocking, and added a full golden-set run after an audit found only 10 of 38 cases were executing; added PR-validation and supply-chain pipelines (secret-pattern scan, Trivy, CycloneDX SBOMs) and resolved the 7 npm audit findings (6 high) the first run reported.
- Capped agent web fetches at 25K tokens, built an opt-in per-run spend limit for the nightly ingest job, and provisioned the new production Cosmos DB account on the free tier.
- Wrote most of a board-memo skill that turns board packs into .docx/PDF memos behind a provenance validator; built snapshot capture and workbook read-through, tested to regenerate identical CapEx figures with zero live calls; wrote the TSMC CapEx extractor.
- Built 7 of the 12 spam, spoofing and prompt-injection gates specified for the auto-ingested newsletter inbox and extended an eighth (DMARC reject, embedding topic filter, injection screen with an opt-in LLM judge); added fail-closed per-tool authorization to the MCP server (18 tools at the time).
- On the company's multi-tenant marketing intelligence platform, filed 34 bugs (1 critical) and, for SOC 2 readiness, added HMAC-signed tenant propagation and fixed spoofable client IPs, fail-open DB TLS and an SSRF redirect bypass.

Founding Engineer / Lead ML Engineer at Athira **2025 – 2026**

- Built the curriculum-planner stage (44 tests) of a four-stage content ingestion pipeline that converts training material into structured session objects for an AI tutor.
- Built cognitive-state and pedagogical feature pipelines (10 cognitive categories) with MongoDB persistence, plus a 63-test integration suite for cross-pipeline data flow.
- Validated the team's Azure ML Studio GPU setup (NVIDIA T4) with a PyTorch benchmark script.

Founding AI Engineer at Stealth AI Startup **June 2024 – January 2025**

- Built a prototype RAG tutoring assistant for UC Berkeley's IND ENG 173 (Stochastic Processes) over Sheldon Ross's *Introduction to Probability Models*, using LangChain, NVIDIA NIM embeddings, FAISS and Llama 3 70B.
- Built its Streamlit front end, connecting a question box to the retrieval chain.

Projects

Repair Loop | GitHub **2026**

- Built a multi-agent code-repair loop in Python: diagnoser, manager, fixer, critic and QA roles call hosted models through W&B Inference, each candidate patch passes a diff safety check and runs in a network-isolated, read-only Docker container on GitHub Actions, and failed attempts feed the next one; traced with W&B Weave, with per-run budgets and checkpoint/resume.
- Built a 185-case benchmark of small buggy Python repositories with hidden regression tests, 150+ of them tested so that a fix special-casing the public tests still fails; ran 273 verification jobs on GitHub Actions.

Technical Skills

Languages: Python, SQL, JavaScript, TypeScript, Java, HTML, Regex

AI & Integrations: MCP server development, LLM/agent applications, RAG, LangChain, REST APIs, webhooks, agent guardrails & evaluation

Tools & Platforms: Git, Docker, GitHub Actions, Azure (DevOps, AI Foundry, ML Studio, Cosmos DB), Terraform, W&B Weave, PostgreSQL/pgvector, MongoDB, FastAPI, Node.js, Express.js, React/Next.js, Streamlit