Harish Manoharan

F-1 OPT (May 2026) | +1 (945) 216-8488 | harish_manoharan@outlook.com | linkedin.com/in/harishm17 | harishm17.github.io

EXPERIENCE

Software Engineer (Core Team)   Jun 2025 - Present
Purgo AI   San Francisco Bay Area

- Engineered a full-stack validation platform for Databricks, automating **IQ/OQ validation workflows** to reduce validation cycles from days to 20 minutes via orchestrated API testing and audit-ready PDF generation
- Designed and implemented an **LLM Evaluation Service**, automating 20+ reliability checks (hallucination rate, PII leakage, traceability) to verify model safety and support enterprise audit readiness
- Built a **graph-based retrieval layer** that indexes Databricks catalog schemas and table lineage, serving semantic context to the code generation agent to boost SQL/PySpark answer correctness by 12%
- Implemented an **agent orchestration state machine** (LangGraph Supervisor-Worker) that routes tasks between specialized agents (Research, Design, Coding) to reliably execute end-to-end data engineering workflows
- Architected an **embedded eSignature workflow** (Docusign API, OAuth2, webhooks), enabling in-platform signing and cutting approval cycle time by 50% through server-side envelope creation and event-driven status updates

Software Engineer Intern   May 2024 - Jul 2024
Digitus Business Solutions   Bangalore, India

- Engineered a **text-to-SQL agent** with dynamic schema discovery and execution-feedback retries, improving SQL execution accuracy from 52% to 76% on a 500-query evaluation set by automating joins across complex schemas
- Built a **multimodal RAG system** indexing internal knowledge (PPTs, videos) with layout-aware chunking and citation enforcement**,** reducing median time-to-resolution (TTR) from 20 min to 6 min
- Delivered a **human-in-the-loop contract tool** with risk detection, shortening reviews from 5 days to 2 days

Technology Consulting Intern - Data Analytics   May 2023 - Jul 2023
PricewaterhouseCoopers   Bangalore, India

- Built a **sales forecasting pipeline** using an XGBoost-SARIMA ensemble with feature engineering for seasonality and promotions, reducing MAPE by 15% over the baseline model for 50+ product SKUs
- Engineered an **aspect-based sentiment analysis pipeline** on 30k+ reviews to identify key satisfaction drivers

PROJECTS

Realtime Exam Prep Coach   Live | GitHub

- Built an AI exam-prep platform that converts course materials into topic-structured notes, quizzes, and graded practice exams using a **multi-stage RAG** pipeline with semantic chunking and hybrid retrieval
- Implemented a **real-time voice coach** using **WebRTC + OpenAI Realtime** with tool-driven drills, concept-only guardrails, automated scoring, and latency/quality telemetry; achieved p95 time-to-first-audio of 600 ms
- Deployed on GCP (Cloud Run/Cloud Tasks) with GitHub Actions CI/CD, Prisma + Postgres, and Google OAuth

Multi-agent Email Assistant   GitHub

- Built a **tool-orchestrated email and calendar copilot** that answers questions over inbox and schedule, drafts reply with personalized writing style, and manages calendar events through Google Workspace integration
- Built a **safety and reliability stack** with schema-validated tool outputs, PII redaction, policy checks, and an independent **LLM judge**; added robust fallbacks and observability for latency and tool failures

Unsupervised Behavior Discovery and Genetics of Mice - Master's Thesis (with The Jackson Laboratory)

- **Temporal Video segmentation** and **action recognition** using key point data from large mice video dataset
- Trained a **bidirectional RNN** encoder-decoder model and performed **k-means** clustering to segment videos

TECHNICAL SKILLS

Programming Languages: Python, C++, TypeScript/JavaScript, SQL
LLM Systems: RAG, AI Agents, LangGraph/LangChain, PyTorch, Langfuse (tracing), LLM evaluation, guardrails/safety
Cloud, Data and Backend: API design, AWS/GCP, Docker, FastAPI, Node.js, Postgres, PySpark, MongoDB, Redis

EDUCATION

M.S. in Computer Science   Expected May 2026
University of Texas at Dallas - Dean's Graduate Scholar   *GPA: 3.91/4.0*

Interdisciplinary Dual Degree: M.Tech, Data Science & B.S., Biological Sciences   May 2024
Indian Institute of Technology Madras - INSPIRE Scholar, Young Research Fellow   *GPA: 8.4/10.0*