

# Nehanth Narendrula

310-598-9433 | [nehanthnarendrula@gmail.com](mailto:nehanthnarendrula@gmail.com) | [linkedin.com/in/nehanth](https://linkedin.com/in/nehanth) | [github.com/nehanth](https://github.com/nehanth)

## EDUCATION

The University of Texas at Dallas

Bachelor of Science in Computer Information Systems and Technology

December 2025

Dallas, TX

## EXPERIENCE

Red Hat

May 2025 – August 2025

Machine Learning Engineer Intern

Boston, MA

- Designed and implemented inline and remote post-training providers for **Direct Preference Optimization (DPO)** in **Llama Stack**, enabling scalable multi-GPU **distributed fine-tuning**.
- Proposed and built a post-training pipeline for **RAG systems** using **contrastive learning** and a **neural fusion model** to blend dense, sparse, and metadata signals, boosting **retrieval accuracy**.
- Researched and prototyped a decentralized LLM inference system by building a private peer-to-peer mesh network with **dynamic layer-wise model sharding**, enabling scalable LLM inference.

Red Hat

May 2024 – August 2024

Software Engineer Intern

Raleigh, NC

- Architected a **query classification system** for OpenShift Lightspeed's LLM layer by training and integrating an ML model to **filter** OpenShift-specific queries and prevent excessive compute usage.
- Led development of telemetry reporting and **S3-based data ingestion**, using NLP to generate **automated KPIs** and implementing CI/CD pipelines to support iterative product improvement.

Blackstone (LaunchPad)

June 2023 – May 2024

Software Engineer Intern

Dallas, TX

- Developed an automated content optimization system leveraging GPT-4 API to generate **context-aware website content**, increasing user engagement and overall site traffic by more than **15%**.
- Engineered a **data-capture module** enabling efficient lead generation and visitor data collection, boosting marketing operations effectiveness by 3% and streamlining business-intelligence processes.

## STARTUP EXPERIENCE

Spin (Concluded) | Python, FastAPI, OpenAI API, Supabase, Terraform, React.js

[www.tryspin.io](https://www.tryspin.io)

Founder & Founding Engineer

[Demo Video](#)

- Built Spin, a multi-cloud platform enabling natural language-driven **infrastructure provisioning**.
- Progress: Developed an LLM-driven agentic orchestration pipeline spanning AWS, GCP, and Azure.

DietAI (Acquired) | Python, Swift, SwiftUI, FastAPI, OpenAI Vision API

[www.dietapp.ai](https://www.dietapp.ai)

Founding Engineer & Advisor

[App Link](#)

- Joined and co-built DietAI, a mobile app utilizing **AI image classification** for calorie counting.
- Progress: Achieved product growth to \$100K MRR and exited at a \$1.6M valuation in 8 months.

## PROJECTS

DecentraHealth | Python, Flask, Node.js, OpenCV, TensorFlow, IPFS

[Demo Video](#)

- Bootstrapped DecentraHealth, a fully serverless decentralized telemedicine platform with IPFS-based video conferencing and **AI-powered diagnostics**, designed for reliable use in disaster zones.

Meta Glasses Enhancement | Python, JavaScript, Node.js, Browser Use API

[Blog Post](#)

- Discovered and created a system to enhance Meta Ray-Ban Glasses capabilities by enabling **voice commands** to communicate with a server for advanced task execution without an official SDK.

## TECHNICAL SKILLS

**Languages:** Python, JavaScript, Java, Go, Swift, Bash, SQL

**Frameworks & Libraries:** Llama Stack, FastAPI, Node.js, React.js, PyTorch, TensorFlow, TRL

**Developer Tools:** AWS, Git, Terraform, Supabase, Kubernetes, Docker, torchrun, GitHub Actions

**Interests/Hobbies:** Hackathons, Tech Conventions, Nintendo Gaming, Cafe Hopping, Rock Climbing