Raghu Talla

(972) 837 – 6406 | raghu.vt6@gmail.com | rtalla1.github.io | linkedin.com/in/raghutalla/

EDUCATION

Texas A&M University, *B.S. in Computer Engineering* Expected Graduation: May 2027

- **GPA:** 3.7
- **Relevant Coursework:** Computer Architecture and Systems, Digital Systems Design, Design and Development of Computer Applications, Data Structures and Algorithms, Discrete Math, Electrical Circuit Theory, Program Design and Concepts, Linear Algebra, Differential Equations, Vector Calculus

SKILLS

Languages: Python, C++, Java, TypeScript, Bash
Frameworks: React, Next.js, Node.js, Flask, OpenCL
Cloud/Infra: AWS, GitHub Actions, Supabase
Developer Tools: Git, Docker, Postman, GDB, Linear
Databases: PostgreSQL, ClickHouse
Systems & Platforms: Linux/Unix, OpenStack

EXPERIENCE

Scale AI, *GenAI Technical Advisor Intern* Nov 2024 - Dec 2025

- Enhanced reasoning and code-generation of LLMs deployed by leading tech companies, improving accuracy and efficiency on competitive programming tasks
- Designed and executed **RLHF** workflows (reward model calibration, ranking tasks) to refine LLM outputs for reasoning, research, and multi-hop problems
- Benchmarked state-of-the-art models on long-context and deep-research prompts, delivering insights that shaped enterprise deployment strategies

Aris Insights, *Software and Data Engineering Intern* May 2025 - Aug 2025

- Engineered **ClickHouse** BI analytics platform with error prevention for 7+ error codes, reducing query failures by **95%** and ensuring **100% uptime** through progressive complexity and fallback strategies
- Developed LLM-driven query engine infrastructure by experimenting with few-shot, XML, Markdown, and structured output prompts, ensuring production-grade reliability in analytics pipelines
- Built data generation framework simulating transaction patterns and lifecycle events with Python, NumPy, and statistical modeling; optimized batch processing with **PostgreSQL**, improving processing efficiency by **65%**
- Collaborated in an **Agile** workflow using Linear for sprint planning, issue tracking, and feature rollout

PROJECTS

OpenLedger: FinTech Transparency Platform (Github) Oct 2025

- Built full-stack prototype that scans fintech codebases and auto-generates verifiable privacy policies
- Integrated Gemini multi-agent workflows and **Supabase** telemetry for versioned disclosures
- Designed modular **Rest API** endpoints and **Next.js** interface for live policy generation and platform transparency

Multithreaded Networked System (GitHub) Nov 2025 - Dec 2025

- Implemented multiple **TCP** server processes in C++ using **POSIX** sockets, each managing a custom **thread pool** with mutexes and condition variables to handle parallel client requests
- Designed a request/response protocol and implemented signal-safe shutdown with EINTR-aware I/O handling

GPU-Accelerated Conway's Game of Life (GitHub) May 2025

- Developed a **C++/OpenCL** simulator from John Conway's Game of Life with parallelized cell updates
- Optimized kernel for GPU memory layout and parallel efficiency with **1.5-2.5x** speedups on **2K-4K** maps

Secure UART Bootloader & Bare-Metal Application (RISC-V) Jul 2025 – Present

- Created secure UART **bootloader** with flash programming, SHA-256/Ed25519 verification, and rollback protection
- Developed **bare-metal** application with custom linker/startup code and interrupt vectors via **JTAG** and **UART**