SREEHARI RAYANNAGARI

rsreehari.rsh@gmail.com | Portfolio | linkedin/sreehari | github/sreehari

Education

Masters in Computer Science and Engineering *Aug 2024 – Dec 2025*
University at Buffalo, The State University of New York | CGPA: **3.85/4.00**
Bachelor of Engineering in Computer Science and Engineering *Aug 2019 – Jun 2023*
Visvesvaraya Technological University, Belagavi, India | CGPA: 8.54/10.00

Skills

Languages: Java, Swift, C, C++, Python, JavaScript, TypeScript
Frameworks: React, Node.js, Express.js, Django
Machine Learning & AI: PyTorch, Scikit-learn, XGBoost, Hugging Face Transformers (LLMs)
Databases & Big Data: MySQL, MongoDB, Apache Spark, MapReduce, FAISS, ChromaDB
Tools: Git, Docker, Xcode, Visual Studio, IntelliJ
Academic expertise: Database Systems, Operating Systems, Computer Networks, AI/ML

Experience

Associate Engineer | HARMAN *Aug 2023 – Aug 2024*

- Developed and shipped **3 production iOS apps** in **Swift/Objective-C** for a leading automotive client, enabling smartphone-car connectivity; delivered **5+ features** including post-update in-app review flow and fixed **15+ major bugs**.
- Resolved critical production incidents (global update & connectivity failures) by coordinating US/China teams; avoided downtime for **1000+ users** by replacing a deprecated API and restoring service during a live outage.
- Modernized frameworks and strengthened cross-team collaboration through **Agile, Git, and unit-testing** practices, driving a **20% reduction in bug reports** and accelerating release velocity.

Summer Intern | Take it smart *Aug 2022 – Sep 2022*

- Built an **e-commerce platform** with secure authentication, product catalog browsing, cart management, and multi-method checkout, supporting dozens of concurrent users and enabling a seamless shopping experience.
- Designed and deployed responsive front-end (**HTML, CSS3, JS, Bootstrap**) integrated with **Django back-end and SQLite database** (5+ relational tables), scalable to **1K+ user/product records.**
- Implemented admin dashboards for **product, order, and coupon management**, and validated workflows via **unit, integration, and end-to-end tests**, ensuring >95% reliability across checkout and payment flows.

Projects

LLM Fine-Tuning & RAG for AAC Communication | Python, Flask/React *Jan 2025 – May 2025*

- Fine-tuned **FLAN-T5** on **25K EmpatheticDialogues** and **58K GoEmotions** samples using Hugging Face Seq2SeqTrainer, achieving measurable gains across **BLEU (+4x)**, **METEOR (+2x)**, and **ROUGE (+2x)** vs. zero-shot baselines.
- Engineered a **Retrieval-Augmented Generation (RAG)** pipeline (**100K+ retrievals**) with FAISS & ChromaDB, indexing simulated AAC user biographies; integrated emotion/context classifiers to reduce hallucinations by **30%** and boost personalization.
- Constructed a full-stack **AAC prototype** with Flask REST API + React UI (real-time interactions, accessibility features); final responses generated via **Mistral API (1K tokens, temp=0.7)** for expressive first-person communication.

Real-Time Music Streaming Platform | MongoDB, Express.js, React, Node.js *Jan 2025 - May 2025*

- Implemented a real-time **MERN stack music streaming platform** with MongoDB (1K+ records), Express.js (500+ concurrent requests/sec), React (UI <50ms), and Node.js (scalable server-side logic).
- Crafted scalable music management & playback with indexed CRUD, instant track switching (∼10ms latency), interactive volume control (0–100), and real-time progress tracking.
- Enhanced user experience with **JWT authentication** (AES-256, 1-hour tokens), **Spotify API integration**, admin dashboard, real-time chat, and custom analytics for engagement insights.

Systems & Networks Projects | OS, Networks, Computer Architecture *Jan 2025 – May 2025*

- Programmed **Pintos OS kernel**, implementing priority-based scheduling (MLFQ, priority donation), 21/28 syscalls (functional) and 28/34 (robust) tests, plus user program support with argument parsing & stack setup.
- Executed a **multi-client TCP chat system** (C++) handling 12+ commands with command multiplexing & offline queuing; simulated multi-layer networks in **ns-3**, performing TCP congestion analysis with Gnuplot visualizations.
- Deployed a **16-bit single-cycle processor** on Basys-3 FPGA (Verilog), integrating 8 components and 10+ R/I/J-type instructions, achieving **100% opcode functionality**, validated via custom testbenches and live waveform/LED verification.

Optimal Urban Tree Planting (Data-Driven Analysis) | Python, PySpark, ML *Aug 2024 - Dec 2024*

- Processed **133K+ urban forestry records (28 features)** using PySpark; cleaned data (IQR outlier removal, null handling, standardization) and performed exploratory analysis.
- Applied **K-Means clustering (k=5, Silhouette=0.95)**, **Linear Regression (R^2≈0.85)**, and **Decision Tree classification (91% accuracy)** to segment regions, predict eco-benefits, and rank zones by ecological value.
- Orchestrated MapReduce aggregations across **11 districts** and enhanced Folium-based **interactive geospatial maps**, surfacing the **top 10 low-benefit clusters** and prioritizing sites with **$1000+ yearly eco-benefits**.