Edric Jeffrey Sam

+917305528953 | edricjsam@gmail.com | LinkedIn | Github | Portfolio | Coimbatore, India

SKILLS

- **Programming:** Python, R, JavaScript
- **Machine Learning & AI:** Computer Vision, NLP, BERT, RAG, Generative AI, Reinforcement Learning (PPO), YOLO, LoRA/QLoRA, Agentic AI
- **Databases:** MongoDB, MySQL, PostgreSQL, Neo4j, FAISS, ChromaDB, Pinecone
- **Backend & API:** FastAPI, Django, Node.js, Express.js, Rest APIs, WebSockets, ngrok SSE Streaming.
- **Distributed Systems & MLOps:** Docker, Celery, Redis, Asynchronous Processing, LLM Inference Pipelines, Containerization, Model Monitoring
- **Cloud & Deployment:** Render, Vercel, Azure
- **ML Frameworks & Libraries:** Langchain, LangGraph, Gen AI SDK, scikit-learn, Tensorflow, Pytorch, Hugging Face, Gymnasium, Stable-Baselines3, Nomic Atlas, Transformers, DeepEval
- **Dev Tools & OS:** Git, Linux (Fedora)

WORK EXPERIENCE

Infotact Solutions | Gen AI Engineer Intern | Coimbatore, India **March 2026 - Present**

- Led a team of 4 by assigning tasks, coordinating workflows, and overseeing end-to-end execution of GenAI projects.
- Built advanced RAG systems using Pinecone, Cohere Rerank-3, Nomic Atlas for adaptive retrieval, sub-1.5s TTFT for grounded AI responses
- Developed Agentic AI workflows using LangGraph and LangSmith with writer-critic orchestration, human-in-the-loop validation, and SSE Streaming

Lysa Solutions | AI Engineer Intern | Coimbatore, India **Oct 2025 - Jan 2026**

- Built a production-grade AI-powered grading system combining OCR, LLM-based evaluation, and result visualization.
- Designed a structured LLM grading pipeline with schema validation, criteria-wise scoring, and concept-level feedback.
- Implemented a centralized LLM Gateway for secure access, retries, token usage tracking, and cost monitoring.
- Developed asynchronous AI workflows using Celery, Redis, Docker, and PostgreSQL; deployed on Microsoft Azure with DevOps collaboration.

OneYes Infotech Solutions | Data Analyst Intern | Chennai, India **May 2025 - Jun 2025**

- Built a Streamlit pipeline to automate K-Means, GMM, and Agglomerative clustering, saving time on data analysis.
- Visualized model performance with Silhouette and Davies-Bouldin scores, enabling quick comparison of clustering methods and guiding model selection
- Enabled storage of user analysis history and generation of downloadable reports through MongoDB.

EDUCATION

Kumaraguru College of Liberal Arts and Science | Bharathiyar University **July 2023 – May 2026**

Bachelor of Science, Data Science **(GPA: 8.13)**

- **Achievements:** Perfect Score in Inferential Statistics, Mahatma Gandhi Merit Scholarship, CSR Project Chair - Rotaract Club of KCLAS (2024 - 25), Brigade Lead of Ignite '25, Best Student Innovation Award

PROJECTS

LLMOps Reliability Platform for Hallucination Detection & Monitoring **March 2026**

- Fine-tuned Qwen3-8B using QLoRA (4-bit quantization) on ~8K samples to develop a hallucination detection model.
- Achieved 83.6% accuracy and 0.86 macro F1-score on a 2K-sample evaluation set using structured metrics (Precision, Recall, Confusion Matrix).
- Built an end-to-end LLM reliability platform integrating Groq inference APIs, FastAPI backend, and React frontend for real-time hallucination detection.

All-in-One RAG Assessment & Adaptive Learning Engine | ***Best Final Year Project Award*** **Dec 2025 - Feb 2026**

- Built a RAG system to generate question papers & answer keys in university template, reducing creation time from ~3 hrs to ~1 min.
- Developed LLM-powered rubric grading for answer scripts and an analytics dashboard, reducing evaluation from ~15 min to ~1 min per paper.
- Engineered an adaptive learning engine with RL (PPO), BKT, and Neo4j; 30–33% better topic selection vs. baseline.
- Migrated MongoDB to PostgreSQL 16 (asyncpg), containerized full stack with Docker Compose (4 services, health checks, persistent volumes).

CareerCompass AI | Kaggle + Google Deepmind Hackathon **Dec 2025**

- Built a multimodal GenAI career intelligence platform using Gemini 3, TTS & STT for profile analysis, interview simulation, and automated asset generation.
- Designed a hybrid GenAI architecture with strict JSON schema enforcement, Google Search grounding for real-time skill gap analysis.
- Converted raw inputs (PDFs, GitHub repos, audio) into deployable outputs including optimized resume bullets, cover letters, and HTML portfolios.

PUBLICATIONS

- **IEEE (Scopus Indexed)**

Emerging Trends & Analytical Perspectives in the Development of New Smartphones.

- **IJIRT (UGC)**

All-in-One RAG Assessment Engine: Dynamic Creation, Automated Evaluation, and University-Centric Output.