Satya Akhil Galla

+1 8575404550 | akhilg@bu.edu | linkedin.com/in/satya-galla | github.com/satyagalla

Summary

AI Engineer and Researcher with an MS in Artificial Intelligence (Boston University), specializing in Generative AI and Computer Vision for **Healthcare** and **Enterprise** applications. Expert in architecting **Hybrid RAG systems**, implementing high-fidelity medical information retrieval, and fine-tuning Transformers (Llama-3, Phi-3). Published author in IEEE with a strong foundation in deploying scalable, hallucination-free AI solutions for complex domains.

Education

Boston University **Jan 2026**
Master of Science in Artificial Intelligence *GPA: 3.7/4.0*

IIIT Sri City **May 2024**
Bachelor of Technology in Computer Science *GPA: 3.5/4.0*

Technical Skills

Generative AI: Hybrid RAG, Cross-Encoders, Medical NLP, Llama-3, Phi-3, Agents (LangGraph), Vector DBs
Engineering: Docker, FastAPI, AWS (EC2/S3), BM25 Retrieval, TCP/IPC Sockets, Git, CI/CD, SQL, Linux
Data Science: XGBoost, Time-Series Forecasting, GCNs, Regex Parsing, Pandas, NumPy, Scikit-learn, OpenCV

Experience

Boston University **Sep 2025 – Jan 2026**
Graduate Teaching Assistant – CS 365 Foundations of Data Science *USA*
- Lead technical labs on Python optimization and statistical modeling, mentoring 50+ students in data science fundamentals.
- Facilitate weekly code reviews and debugging sessions to reinforce best practices in algorithmic implementation.

Aaizel International Tech **Feb 2024 – May 2024**
Machine Learning Engineer Intern *India*
- Created the domain's **first remote sensing scene graph dataset** and architected a multimodal fusion pipeline (IFCNN) merging IR+RGB satellite imagery.
- Fine-tuned a **RelTR Transformer** for geospatial relationship extraction and deployed a **Gradio** interface for instant visualization.

Terrafic Inc **June 2023 – Nov 2023**
Computer Vision Intern *India*
- Implemented Super-Resolution (SR) algorithms (4x/6x) to enhance object detection on low-resolution defense mapping data.
- Integrated **Segment Anything Model (SAM)** to automate feature extraction, delivering a production-ready MVP that secured initial client pilots.

ISRO (Indian Space Research Organization) **Dec 2022 – May 2023**
Machine Learning Engineer Intern *India*
- Authored **IEEE WHISPERS 2023** paper presenting the first hyperspectral classification framework for Chandrayaan-2 IIRS data, utilizing a **Latent GCN** approach.
- Engineered a **Spectral-Spatial Non-Linearity** graph formulation to handle irregular topography, achieving **90.1% accuracy** on benchmarks (Chandrayaan-1 M3).

Key Projects

Advanced Hybrid RAG for Healthcare | *Llama-3, Hybrid Search, Re-ranking* | GitHub **Sep 2024 – Dec 2024**
- Engineered an enterprise-grade **Hybrid RAG pipeline** combining Dense Vector Search (semantic) with **BM25 (keyword)** retrieval to capture precise medical nomenclature often missed by standard embeddings.
- Implemented a **Cross-Encoder Re-ranking** layer to strictly filter retrieved contexts, ensuring only high-relevance clinical data reaches the generation model.
- Achieved an **87% SBERT alignment score**, effectively boosting semantic fidelity by **38%** over the un-augmented Llama-3 baseline and eliminating critical hallucinations.

Multimodal Skin Cancer Detection System | *PyTorch, Medical Imaging* | GitHub **Sep 2024 – Dec 2024**
- Tackled 0.1% class imbalance in medical datasets using synthetic oversampling and developed a voting ensemble fusing CNNs and tabular models.
- Achieved an **ROC-AUC of 0.96**, surpassing single-modal baselines by **15 points** in final validation, demonstrating robust diagnostic capability.

Predictive Analytics for MBTA Transit | *XGBoost, Business Intelligence* | GitHub **Feb 2024 – May 2024**
- Engineered a unified ETL pipeline aggregating **9 years** (2015-2024) of multi-modal data to train **XGBoost** (Reliability) and **Random Forest** (Ridership) models.
- Uncovered a system-wide **33.35% post-pandemic ridership drop** and identified critical service disparities in 30-60% AMI corridors, providing actionable insights for transit optimization.

Cozy Companion (CoCo): Neuro-Symbolic AI Agent | *Phi-3, System Design* | GitHub **Dec 2025 – Jan 2026**
- Engineered a **Subsumption Architecture** decoupling a 60Hz reactive loop from asynchronous **Phi-3 (4-bit)** LLM inference via TCP sockets to achieve **<16ms latency**.
- Integrated a **Leaky Integrate-and-Fire (LIF)** model to estimate user flow state from interaction patterns.

Select Publications

Efficient Graph Formulation and Latent Space Integration for Lunar Hyperspectral Image Classification – *IEEE WHISPERS 2023*
- Proposed a Latent Graph Convolutional Network (GCN) utilizing an Autoencoder for dimensionality reduction, achieving **90.1% accuracy** on benchmarks by engineering a spectral-spatial non-linearity function.

Enhancing Hyperspectral Classification through GCNs with Adaptive Graph Construction – *NASA NESF 2023 (Poster)*
- Presented a comparative analysis of adaptive graph constructions for **Chandrayaan-2 IIRS** data, identifying non-linear spectral-spatial connections as optimal for handling irregular topography.