

Harish Akula

harish.akula2000@outlook.com | +1 (857)-234-2862 | Boston, MA | LinkedIn: [harish213](#) | GitHub: [harish32k](#)

TECHNICAL SKILLS

Languages: Python, C++, Java, SQL, Bash (Shell Scripting)
Machine Learning: PyTorch, TensorFlow, LLMs, PEFT, Scikit-Learn, OpenCV, Hugging Face, Computer Vision
MLOps & Others: Google Cloud/GCP, AWS, Docker, MLflow, Databricks, Apache Spark, Linux, Git, Flask

WORK EXPERIENCE

Institute for Experiential AI at Northeastern University

Jul. 2024 – Present

Machine Learning Researcher

Boston, MA

- Engineered a Multi-Agent **Reinforcement Learning** framework (PPO) regulating international flight traffic during epidemics to balance economy and healthcare, reducing 90% of the infected travelers while retaining 80% trade.
- Achieved a **75% reduction** in computational resources, while scaling the RL framework from single-country to multi-country networks by reducing state-space dependency from quadratic to linear.
- Accelerated inference by **1,000,000x** (seconds to microseconds) by developing ensembled surrogate ML models that achieved 99.63% accuracy (state of the art) for physics simulations, eliminating a critical bottleneck.
- Automated an **MLOps pipeline** on Databricks using MLflow and Optuna, lowering manual hyperparameter tuning time by ~60% and enabling non-ML experts to validate models without managing underlying infrastructure.

Deloitte

Aug. 2022 – July 2023

Business Technology Analyst (Tax Services)

India

Recognition: Received "Spot Award" for critical contributions to ETL workflows.

- Engineered hybrid Alteryx/Python ETL pipelines to parse corrupted files and standardize tax data for enterprise ingestion, preventing failures and reducing manual effort by ~80 hours.
- Refactored complex SQL Server stored procedures to process millions of client records, resolving logic errors and query inefficiencies to ensure 100% data accuracy for downstream reporting.

PUBLICATIONS & PATENTS

- [Patent]** IoT-Based Smart Blind Assistive System | *Indian Patent App. No. 202541058572 (Published 2025, in examination)*
- [Paper]** Akula, H. et al. Assistive System for the Visually Impaired using Multiple Cameras | *IEEE ICCMCLA 2022* | [Link](#)
- [Paper]** Akula, H. et al. Dynamic Game Difficulty Adjustment based on Facial Emotion Recognition | *Springer ICDSMLA 2023* | [Link](#)

PROJECTS

AI - Powered Blind Assistive System (patent-filed, best-paper award) | Python, Docker, GCP, Android | [GitHub](#) | [Video Demo](#)

- Deployed Vision-Language models on Google Cloud for a blind assistance device, enabling real-time blind navigation assistance.
- Architected distributed AI microservices fusing multi-model outputs, linking mobile and edge devices to the cloud via a Flask API.
- Scaled transformer-model inference via Docker containers, offloading heavy compute from low-power edge-devices to the cloud, **reducing infrastructure costs by ~\$400 per user.**

Efficient Multimodal Fine-Tuning Framework | Python, PyTorch, Transformers, LoRA | [GitHub](#)

- Pioneered **the first** finetuning (PEFT) framework for large Vision-Language Models, achieving linear time complexity (Mamba).
- Designed light-weight neural-network adapters for alignment between text and images encodings, **surpassing transformer** baselines in segmentation accuracy by ~6.6% (IoU).
- Reduced trainable parameters to **only 5.8%**, maintaining competitive accuracy on complex, ambiguous natural language queries.

Abstract Generation with Small Language Models (SLMs) | Hugging Face, Transformers, Q-LoRA | [GitHub](#)

- Trained Small Language Models (Mistral-7B, Llama-2, Gemma) via Instruction Tuning and Q-LoRA on arXiv research papers, enabling long-document understanding for automated abstract generation.
- Engineered a text-ranking pipeline to pre-process long texts (10,000+ tokens), bypassing SLM context limits on constrained GPUs.

EDUCATION

Northeastern University

Dec 2025

Master of Science, Computer Science (GPA: 4.0/4.0)

Boston, MA

- Related coursework: Deep Learning, Natural Language Processing (NLP), Machine Learning, Algorithms

Chaitanya Bharathi Institute of Technology

June 2022

Bachelor of Engineering, Computer Science and Engineering (GPA: 8.74/10)

India

- Awards: Won the **"Best Paper Award"** for AI based Smart Blind Assistance System on the institute's 3rd Research Day. | [Link](#)
- Related coursework: Data Structures, Artificial Intelligence, Probability and Statistics