Anirudh Kashyap Ramesh
Arlington, TX |+1 (682) 559-5269 | anirudh11011@gmail.com | LinkedIn | Portfolio

Objective

Seeking full-time or internship opportunities in Machine Learning or Generative AI to build smart and useful AI solutions.

Experience

Full Stack Software Developer **Avanseus, Bengaluru** **February 2024 – July 2024**

- Developed and maintained full-stack applications using React, Spring Boot, MongoDB, and RESTful APIs.
- Designed and implemented REST APIs for seamless data exchange, enabling CRUD operations and server-side pagination.
- Improved system functionality and user experience by integrating Axios for efficient API communication.

Education

Masters-Computer Science (GPA 4.0/4.0) **University of Texas at Arlington** **August 2024 - Present**

Data Analysis & Modelling techniques	Design & Analysis of Algorithms	Artificial Intelligence
Machine Learning	Data Mining	Database Systems

Bachelors- Information Science (GPA 3.7/4.0) **JSS Academy of Technical Education, Bengaluru** **August 2019 – May 2023**

Big Data Analytics	Database management	Machine Learning
Operating Systems	Object-Oriented Concepts	Software Engineering

Technical Skills

Programming Languages	: Python, JavaScript, C
Development Tools	: GitHub, Git, Docker, Kubernetes, SageMaker, MS Excel , VS Code
Operating Systems	: Windows, Mac OS, Linux
Web Technologies	: HTML, CSS, React, Node.js, AWS, Streamlit, Gradio, Flask
Machine Learning	: TensorFlow, PyTorch, Keras, LangChain, NumPy, Pandas, Matplotlib, Scikit-learn, CNN, RNN, NLP, MLflow, BeautifulSoup, Hugging Face
Predictive Modelling & Analysis	: Regression, Classification, Clustering, Anomaly Detection
Data Handling & Databases	: MySQL, MongoDB, SQL, Vector Databases

Projects

Multi-Agent Photography Assistant **June 2025**

- Developed a modular multi-agent AI system using LangChain's agent-supervisor architecture, with LLaMA-based language models analysing user scenarios to suggest DSLR configurations.
- Designed and coordinated independent agents for scenario interpretation and camera spec generation via a centralized LangChain supervisor, demonstrating intelligent task delegation.
- Integrated ChromaDB as a vector database and implemented a Retrieval-Augmented Generation (RAG) system to enhance contextual understanding and recommendation accuracy.
- Built a user-friendly interface using Gardio, and incorporated Whisper for real-time speech-to-text input, improving accessibility and interactivity.
- Leveraged the Model Context Protocol (MCP) for structured communication between agents and seamless external tool integration within the multi-agent ecosystem.

Deep Learning – Image Classification & Transfer Learning **April 2025**

- Trained and evaluated CNN and ResNet18 on Imagenette dataset using PyTorch Lightning, achieved 63.08% and 56.40% test accuracy with early stopping and model checkpointing for stable convergence.
- Enhanced model generalization by applying data augmentation (random flip, rotation, colour jitter) and dropout (0.5) to ResNet18, improving CIFAR-10 test accuracy to 82.06% and reducing validation loss to 0.0130.
- Applied transfer learning on ResNet18 pre-trained on Imagenette, achieving 83.4% validation accuracy and outperforming baseline models (train loss: 0.0508, val loss: 0.650 vs. 0.964).
- Integrated MLflow for automated experiment tracking, comparative performance visualization, and reproducible model development across training configurations.

UTA Chatbot – LLM-Powered University Info Assistant **March 2025**

- Built in 24 hours during a hackathon, this conversational assistant provides real-time answers about the University of Texas at Arlington using LLMs (LLaMA 3–70B), LangChain, and Prompt Engineering.
- Deployed an intuitive Streamlit user interface on Streamlit Cloud for a responsive chat experience.
- Implemented a Retrieval-Augmented Generation (RAG) pipeline using Google Custom Search, web scraping (BeautifulSoup), ChromaDB, and Sentence Transformers to retrieve and embed relevant content.
- Integrated Groq API to generate accurate context-aware responses using external search and vector similarity.
- Developed structured tests to assess LLM output accuracy, bias, and reliability using prompt variations and RAG-based context to detect hallucinations and edge-case failures.