KIRAN BABU ATHINA

Austin, TX · (979) 721-3721 · kiranathina8@gmail.com · linkedin.com/in/athinakiranbabu · github.com/FriendyKiran · kiranbabuathina.com

SUMMARY

Data Scientist and Machine Learning Engineer with 4 years of experience turning large, complex datasets into statistical models, predictive systems, and applied ML/GenAI solutions across telecom, finance, healthcare, agriculture, and scientific research. Experience spans big data pipelines and predictive modeling to LLM-powered RAG systems, using Python, SQL, PyTorch, TensorFlow, Hugging Face, LangChain, and LangGraph. Strong background in experimental design, model evaluation, and optimization, with a track record of translating data science research into production-ready systems and clear, actionable recommendations for cross-functional stakeholders.

TECHNICAL SKILLS

ML / GenAI: LLMs, RAG, Agentic AI, Knowledge Graphs, Prompt Engineering, LoRA, QLoRA, PEFT, Transformers, Vision-Language Models, CLIP
Frameworks: LangChain, LangGraph, PyTorch, TensorFlow, Scikit-learn, Hugging Face, NLTK, FAISS, FastAPI
Programming: Python, JavaScript, SQL, R, C
Databases: PostgreSQL, MySQL, MongoDB, Redis, NoSQL
Cloud & HPC: Azure (AI Search, AI Foundry, Blob Storage, App Service), AWS (S3, EC2, SageMaker), NVIDIA GPU Clusters
DevOps: Docker, Kubernetes, CI/CD, Git, GitHub
Certifications: NVIDIA DLI Certified, Generative AI Fundamentals, Python Developer

EXPERIENCE

Texas A&M AISLS Lab
College Station, TX

Research Assistant
June 2025 – Present

- Engineered a triple-retriever RAG pipeline (KG semantic, Qdrant dense, BM25) with Reciprocal Rank Fusion and cross-encoder reranking, achieving 0.95 faithfulness and 89% answer accuracy on a 10,000-question domain eval set.
- Boosted context recall from 0.38 → 0.94 by constructing a domain knowledge graph of 7,405 entities and 11,249 relationships from 235 research papers using a multi-pass Qwen3 14B extraction pipeline with 93% verbatim quote verification.
- Developed an agentic retriever with 8 targeted graph and passage search actions over a Kuzu graph database and embedded Qdrant index, achieving 91% gold-page retrieval and 86% gold-quote recall, measured deterministically against source PDFs.
- Designed a React frontend with clickable [F#]/[P#] citations that open the exact source PDF page with the retrieved passage highlighted and per-citation retrieval scores surfaced, delivering full end-to-end citation transparency.

Texas A&M Department of Animal Science
College Station, TX

Graduate Research Assistant
January 2024 – May 2025

- Enabled minute-level per-animal energy tracking and monthly methane forecasting, validated against live MQTT sensor ground truth, by engineering a sensor-driven digital twin using a Python Mesa agent-based simulation model on university HPC infrastructure.
- Eliminated dependency on general-purpose cloud APIs for livestock domain queries, serving 3 research semesters of faculty-validated production use, by fine-tuning LLaMA 3 with LoRA/PEFT on a proprietary animal science corpus and integrating Azure AI Search with Azure OpenAI for grounded, citation-backed generation.

Tata Consultancy Services
Hyderabad, India

Machine Learning Engineer
August 2021 – July 2023

- Processed **100M+** sensor records per day for downstream ML fault prediction, measured by pipeline throughput and feature freshness SLAs, by designing a PySpark ETL framework on Hadoop to extract, filter, and aggregate telecom gateway logs.
- Cut loan processing time by **30%**, measured by end-to-end decision turnaround across client portfolios, by developing classification models on Azure ML with AutoML-driven selection and statistical feature engineering on approval probability scores.
- Eliminated **30%** of manual hardware log triage and improved issue resolution time by **18%**, measured against baseline manual review throughput, by engineering an NLP classification pipeline using TF-IDF and Random Forest across **500,000+** daily telecom log entries.

- Reduced post-deployment model decay by **22%** and cut deployment turnaround by **30%**, measured by metric drift and release cycle time across production, by establishing drift monitoring dashboards and containerizing ML applications with Docker orchestrated on Kubernetes.

Machine Learning Engineer Intern *May 2021 – July 2021*

- Achieved **98.5%** binary fraud classification accuracy on KYC anomaly detection, measured against a held-out production test set, by implementing a multimodal PyTorch model fusing 1D CNN audio embeddings and 2D CNN document image encodings via fully connected classification layers.
- Reduced team onboarding and knowledge handoff time by **25%**, measured across 2 project handoffs, by authoring end-to-end documentation covering model architectures, training assumptions, and evaluation results for senior engineers and stakeholders.

PROJECTS

AI-Powered Retinal Disease Diagnosis with Vision-Language Models

- Achieved **91%** diagnostic accuracy on a 12-class retinal disease classification task, validated against **1,200+** expert-annotated fundus images, by fine-tuning MedGemma 4B on paired image and clinical note inputs using vision-language modeling.

Image-Text Retrieval with Contrastive Learning (CLIP-style)

- Improved image-text retrieval by **17%** on MSCOCO and zero-shot classification by **14%** on ImageNet, benchmarked against baseline CLIP, by developing a dual-encoder ResNet-50 + DistilBERT model trained with NT-Xent and iSogCLR loss functions and AdamW/RAdam optimizers.

Exoplanet Discovery | 2025 NASA International Space Apps Challenge

- Achieved **96.64%** classification accuracy on confirmed exoplanets, candidates, and false positives, benchmarked across Kepler and TESS mission datasets, by training and ensembling Random Forest, XGBoost, LightGBM, Gradient Boosting, and MLP models with interactive Streamlit deployment.

LEADERSHIP

Texas A&M University | Aggie Research Program *College Station, TX*

Project Lead *Aug 2024 – May 2025*

- Delivered research outputs integrated into 2 active faculty publications, measured by professor-approved contributions across 2 consecutive semesters, by leading 2 concurrent teams of 3 undergraduate researchers, setting weekly task milestones, and reporting progress directly to the supervising professor.
- Elevated 6 undergraduates to publication-ready contribution standard, measured by deliverables accepted into active research outputs, by providing hands-on mentorship in applied ML, reviewing all technical work, and aligning team output with faculty research goals.

EDUCATION

Texas A&M University *Aug 2023 – May 2025*

M.S. in Data Science · GPA: 3.9 / 4.0

Gayatri Vidya Parishad College of Engineering *Aug 2017 – Jul 2021*

B.Tech. in Electronics & Communication Engineering · First Class with Distinction