Preetham Anand

AI Engineer | Applied ML | LLM Applications | Python & FastAPI

Hyderabad, India | +91 81210 78989 | preetham0802@gmail.com
linkedin.com/in/preethamanand | github.com/PreethamAnand | preetham.live

Summary

Final-year B.Tech (AI & ML) student who builds AI applications end to end: models and LLM pipelines served through FastAPI, with validated outputs, fallbacks and automated tests. Work spans grounded LLM root-cause reasoning for industrial QA, a defense against poisoned knowledge in RAG, and a deployed LangGraph-powered platform, plus three internships in ML and full-stack development.

Education

B.Tech, Computer Science & Engineering (AI & ML) – Vignan Institute of Technology and Science, HyderabadCGPA 8.6/10 | 2023 – 2027

Technical Skills

Languages: Python, TypeScript, JavaScript, SQL
AI / ML: PyTorch, scikit-learn, XGBoost, Hugging Face Transformers, Sentence-Transformers, OpenCV, TensorFlow/Keras
LLM & GenAI: RAG, LangGraph, LangChain, embeddings, ChromaDB, Pydantic structured outputs, Groq, OpenRouter, Llama 3, Qwen2.5
Backend & Data: FastAPI, Flask, Node.js/Express, REST APIs, PostgreSQL (Supabase, PostGIS), SQLite, MongoDB
Engineering: pytest, Git/GitHub, GitHub Actions, Vercel, Render, Streamlit

Experience

AI/ML Intern – Cothon Solutions Mar 2025 – May 2025

- Built an airfare price model on 300K flight records, comparing Linear Regression, Random Forest and XGBoost (best R^2 0.973, 5-fold CV), and served it through a Streamlit app.
- Trained a Keras CNN to classify real vs. deepfake faces from FaceForensics++ video frames and exposed it through a Flask demo.

Full Stack Developer Intern – Nirvaha Wellness LLP Oct 2025 – Feb 2026

- Built a Node.js/Express + MongoDB REST API (35 endpoints) with JWT auth and Socket.IO real-time sync.

AI Intern – Edunet Foundation (AICTE TechSaksham, SAP & Microsoft) Aug 2025 – Oct 2025

- Trained and evaluated five scikit-learn classifiers for disease risk prediction, shipped in a multi-page Streamlit app.

Projects

VisionForgeX – Closed-Loop Industrial QA & Root-Cause Reasoning | *Python, PyTorch, Pydantic, NetworkX, Streamlit*Repo | Oct 2026

- Built the LLM reasoning layer for a 3-member Infosys Hack-AI-Thon team: Llama-3.1-8B on Qualcomm Cloud AI 100 with timeout, retry and rule-engine fallback, plus a Pydantic validator that rejects root causes citing unknown evidence.
- Engineered telemetry anomaly detection (past-only rolling z-scores) fused with defect detections into a NetworkX evidence graph; the team's DINOv2 detector reached 0.98 AUROC at 141 ms/image on CPU.
- Delivered a 9-tab Streamlit dashboard, PDF audit reports and 70+ pytest tests across 5 end-to-end scenarios.

RepairRAG – Repairing Poisoned Knowledge in RAG | *PyTorch, Transformers, Qwen2.5, NLI, FastAPI*Team of 4 | 2026 – Present

- Co-developing (team of 4) a defense that traces a poisoned RAG answer to its source documents, rewrites only the poisoned sentences from a trusted corpus, and commits a fix only if an NLI-based confidence gate passes.
- Pilot run (Natural Questions): attack success cut from 100% to 6.3% (vs. 12.5% for document removal), 0.96 attribution F1, zero benign documents modified; 61 tests at 83% coverage.

PulseConnect 2.0 – AI-Assisted Blood Donation Platform | *FastAPI, LangGraph, Supabase/PostGIS, React* Live | Oct 2025 – Aug 2026

- Built a FastAPI AI service whose scheduled LangGraph workflows escalate stale emergency requests and broadcast shortage alerts, with every run logged to an admin monitoring view.
- Implemented donor matching with a PostGIS radius search, blood-group compatibility rules and weighted proximity, recency and availability scoring over a 17-table Postgres schema.

ResearchPilot – RAG Research Assistant | *FastAPI, ChromaDB, Llama-3.3-70B, React* Live | Feb 2026

- Co-built and deployed (Vercel + Render) a PDF question-answering app with chunked embeddings, similarity-filtered retrieval and source-cited answers; fixed Render out-of-memory crashes by lazy-loading the model.

Achievements & Certifications

Infosys ENG NEXT 2026 Hack-AI-Thon; NASSCOM KG Hackathon; Amazon ML Challenge 2025 | Deloitte Technology & Data Analytics Job Simulations (2026); NPTEL: Python, DSA using Python, Cloud Computing; HackerRank Python (Gold)