AALOK PANDIT

+91 9035442978 ◇ Bengaluru, India
paalok333@gmail.com ◇ linkedin.com/in/alowkii ◇ github.com/alowkii ◇ aalokpandit.netlify.app

SUMMARY

Software engineer building ML platform infrastructure: inference services, APIs, and training-to-serving pipelines for small language models, plus the data pipelines and internal tools around them.

EXPERIENCE

Adapt Ready Inc. – **AI/ML Platform Engineer** Nov 2025 – Present

- Owned the end-to-end lifecycle of **small language models (SLMs, 0.6B–1B)**: **LoRA** fine-tuning (Unsloth, TRL) → **GGUF** quantization → **Ollama** serving in production at **p95 latency** $\leq$ **1 s**.
- Automated data extraction and processing procedure from **20 days to under 60 seconds** (>**99.99% reduction**)
- Built **evaluation, analytics and internal developer tooling** for the ML and engineering teams, including an offline harness for base-vs-fine-tuned comparison and error analysis.
- Built a **FastAPI** inference service with an OpenAI-compatible API, **JSON-schema-constrained decoding** for guaranteed valid output, Pydantic request validation, concurrent request handling, and automated tests in **CI**.
- Drove applied **R&D**, including developing a **federated learning** system for privacy-preserving model training, and collaborated with data scientists and data engineers on production ML systems.

Adapt Ready Inc. – **AI/ML Intern** May 2025 – Nov 2025

- Built agentic workflows with **LangGraph** and **MCP** tools, grounded in **Neo4j** knowledge graphs and **FAISS** retrieval.
- Built scheduled **Airflow** data pipelines and containerized backend services with **Docker**.

PROJECTS

nia – Local Voice Assistant Python, Whisper, Ollama

- Built an on-device voice assistant using wake-word detection, **Whisper** ASR, **Ollama** and structured **tool calling**, with contract tests validating tool routing.

Image Forgery Detection (RDLNN) 1st Place, INVATI Expo 2025

- Built image forgery detection and localization using wavelet transforms, **CNN**-derived features and a regression DNN in **PyTorch**; served via REST API with batch inference.

CookLyt – Restaurant POS & CRM Node.js, PostgreSQL, React

- Led a 5-person team building a POS & CRM platform: **Express** + **PostgreSQL** backend, real-time **WebSockets**, and **React** POS, admin and QR-ordering apps.

SKILLS

Languages	Python, SQL, TypeScript/JavaScript, C/C++, Cypher
ML / LLMs	PyTorch, HF, TRL, LoRA/QLoRA, Federated Learning, LLM Structured Outputs
LLM Systems	Ollama, GGUF quantization, LangGraph, LangChain, RAG, tool calling, MCP, FAISS
Backend & Data	FastAPI, Node.js/Express, REST APIs, WebSockets, Airflow, Pandas
Data / Infra	PostgreSQL, Neo4j, MongoDB, Docker, Linux, Git, CI, CUDA

EDUCATION

B.E. Computer Science and Engineering, Nitte Meenakshi Institute of Technology 2021–2025