

Manoj Kumar Thapa

✉ thapam807@gmail.com | ☎ +44-7438-029689 | 📍 Birmingham, UK | 🔗 LinkedIn | 🐙 GitHub | 🧑 Portfolio

SUMMARY

AI Software Engineer with an MSc in Artificial Intelligence, building production-facing AI features across the stack: LLM-powered applications, retrieval-augmented generation pipelines, and agentic tool orchestration. Shipped three distinct AI products end to end, from backend service (FastAPI, Python) through frontend (TypeScript, React), including an agent orchestrating 5 tool integrations via LangGraph and a RAG pipeline grounding LLM output in source documents via ChromaDB. Backed by professional backend engineering experience at Accenture, owning production REST APIs and improving response time by 40%. Strong Python fundamentals combined with a fast-execution, first-principles approach to building and shipping AI capabilities.

TECHNICAL SKILLS

Languages: Python, TypeScript, SQL, Java

LLM & Agentic Frameworks: LangGraph, OpenAI API, Retrieval-Augmented Generation (RAG), Prompt Engineering, Agent Tool Orchestration

Retrieval: ChromaDB (vector search), Embeddings, Semantic Retrieval

Frontend: TypeScript, React

Backend/APIs: FastAPI, Spring Boot, REST API Design, Schema Validation

Databases: PostgreSQL, MySQL, ChromaDB

Cloud/DevOps: AWS (EC2, S3, RDS), Docker, Git, GitHub Actions CI/CD

Machine Learning: PyTorch, Transfer Learning, Model Evaluation & Metrics

Practices: Cross-Functional Collaboration, Performance Optimisation, Automated Testing

PROJECTS

Planck AI, Agentic Application with Tool Orchestration | *Python, LangGraph, FastAPI, TypeScript, React* | Live Demo Sep 2025 – Nov 2025

- Built a full-stack agentic AI application from concept to deployed product, orchestrating 5 tool integrations to plan and resolve complex multi-step queries
- Implemented backend logic in FastAPI and a TypeScript/React frontend, designing tool routing and state management to control context passed to the model at each step

NeuroArc, LLM-Powered Document Tool | *Python, FastAPI, OpenAI API, React* | Live Demo Nov 2025 – Dec 2025

- Built an AI-powered tool processing documents using the OpenAI API, adding schema-level validation and structured error handling to improve production reliability

CogniGraph, RAG Pipeline for Document Analysis | *Python, RAG, ChromaDB, Knowledge Graphs* Jun 2025 – Jul 2025

- Built a retrieval-augmented generation pipeline extracting entities and relationships from unstructured PDFs, using ChromaDB for vector storage and semantic retrieval
- Grounded generated content in retrieved source passages to reduce hallucination and improve factual accuracy of automated output

FruitGuard AI, MSc Dissertation | *Python, PyTorch, YOLOv8, ONNX* | Live Demo Aug 2025 – Apr 2026

- Researched and implemented a two-phase transfer learning pipeline achieving 96.3% classification accuracy from 207 training images, a 95% reduction in required annotation data
- Optimised model performance for production, deploying via ONNX Runtime at 73 FPS for real-time inference

EXPERIENCE

Software Engineer

Accenture

Oct 2021 – May 2022

Bangalore, India

- Built and maintained production backend services in Java for a large-scale enterprise application
- Owned REST API services end-to-end, implementing business logic and PostgreSQL/MySQL data access layers
- Improved system performance, reducing API response time by 40% through SQL query optimisation and elimination of N+1 query issues

- Collaborated cross-functionally with product stakeholders to translate requirements into technical solutions

Mentor

Aug 2025 – Jan 2026

Aston University

Birmingham, UK

- Mentored 10+ MSc AI students on Python and machine learning through weekly one-on-one sessions

EDUCATION

Aston University

Birmingham, UK

MSc Artificial Intelligence (Computer Science) – 70.37%

Jan 2025 – Apr 2026

Dr. Ambedkar Institute of Technology

Bangalore, India

Bachelor of Engineering in Computer Science – 9.45/10

Aug 2017 – Sep 2021

CERTIFICATIONS

5-Day AI Agents Intensive Course with Google (*Kaggle*) | **Machine Learning** (*Stanford*) | **Deep Learning Specialization** (*DeepLearning.AI*)