

MANOKARAN K

Python Developer · Generative AI Engineer · RAG & LLM Systems

+91 7373232074 • manokarank1916@gmail.com • Chennai, Tamil Nadu, India • github.com/manokaran7373

PROFESSIONAL SUMMARY

Results-driven Python & Generative AI Engineer with 2.6 years of experience designing and deploying production-grade AI systems, RAG pipelines, and scalable FastAPI microservices. Deep expertise in LLM orchestration (LangChain, LlamaIndex), vector databases (ChromaDB), and agentic AI frameworks (CrewAI, MCP). Proven track record of modernizing legacy architectures into high-performance, cloud-native solutions—achieving up to 30% improvement in system latency. Passionate about building intelligent applications that bridge cutting-edge AI research with real-world enterprise value.

CORE TECHNICAL SKILLS

AI / LLM	LangChain, LlamaIndex, CrewAI, OpenAI API, Gemini API, Prompt Engineering, Retrieval-Augmented Generation (RAG), Agentic AI, FastMCP, MCP
Vector Stores	ChromaDB, Semantic Search, Embedding Models, Cohere Embeddings (embed-english-v3.0)
LLM Models	GPT-4, Gemini Pro, Mistral-7B-Instruct (HuggingFace), Open-Source LLM Deployment
Backend	Python 3.x, FastAPI, Django, Flask, RESTful API Design, Async Programming, SSE Streaming
Databases	PostgreSQL, MySQL, Vector Database Design, SQL Query Optimization
Security	JWT Authentication, AES-256, Secure Document Handling, OAuth 2.0
Cloud & DevOps	AWS (EC2, S3, Amplify), Docker, CI/CD, Background Task Processing
Frontend	React, Next.js, HTML5, CSS3, Bootstrap, JavaScript
Data / ML	Pandas, NumPy, Scikit-learn, Data Pipeline Design
Tools	Git, GitHub, Postman, uv (Python package manager), VS Code

WORK EXPERIENCE

Generative AI Developer | VPearl Solutions

January 2025 – Present

- Architecting and developing enterprise RAG systems using LangChain, ChromaDB, and Cohere embeddings (embed-english-v3.0) with semantic search for intelligent document retrieval.
- Built and deployed the AI Data Warehouse Chatbot—an NLP-driven platform enabling users to query and analyze databases using natural language via GPT-4 and MCP protocol integration.
- Implemented MCP (Model Context Protocol) to standardize AI access to enterprise resources, enabling full data governance and auditability.

Python Developer | Adshi5.com

July 2024 – December 2024

- Developed and maintained scalable RESTful APIs using FastAPI and PostgreSQL, serving high-concurrency enterprise workloads.
- Built the Nexus ERP System—a centralized platform for managing employees, vendors, customers, materials, and financials with SMTP email notifications and AWS S3 file storage.
- Modernized legacy monolithic services into FastAPI microservices, resulting in a measurable 30% improvement in system response latency.

- Contributed to backend Python services for data-intensive web applications, focusing on API performance and code quality.
- Collaborated in Agile sprints to deliver features on schedule, participated in code reviews and maintained high test coverage.
- Supported cloud infrastructure on AWS EC2 and S3 for deployment and static asset management.

KEY PROJECTS

AI Data Warehouse Chatbot

Tech Stack: Python · FastAPI · GPT-4 · MCP (Model Context Protocol) · PostgreSQL · Prompt Engineering

- Designed an intelligent NLP interface allowing non-technical users to query enterprise databases using plain English—eliminating manual SQL writing.
- Leveraged MCP protocol to standardize AI access to enterprise data sources, providing full auditability of how data is queried and delivered.
- Integrated GPT-4 with carefully engineered system prompts to generate accurate, parameterized SQL and return structured results in real time.
- Improved data governance and compliance: organizations gain complete visibility into AI-driven data access patterns.

Nexus ERP System

Tech Stack: Python · FastAPI · PostgreSQL · AWS S3 · SMTP · Docker · JWT

- Built a comprehensive ERP platform managing companies, employees, vendors, customers, materials, banking, and expenses from a single admin interface.
- Automated SMTP email workflows for expense approvals, vendor onboarding, and employee notifications, reducing manual admin overhead.
- Achieved 30% latency improvement by migrating from a monolithic Django app to asynchronous FastAPI microservices.
- Integrated AWS S3 for secure document storage with pre-signed URL access and lifecycle policies.

CERTIFICATIONS

- Generative AI, Agentic AI & RAG Development — Udemy
- MCP Server Development — LinkedIn Learning
- Web Application Development — Certified

EDUCATION

Bachelor of Science (B.Sc.) — CGPA: 6.9 | 2020

HSC: 86% | SSLC: 89%

AREAS OF INTEREST

Generative AI & LLM Systems · RAG Architecture · Agentic AI & Multi-Agent Systems · AI-Powered Backend Engineering · Conversational AI · Full-Stack AI Applications
