

YOGESH KUMAR *Applied AI & Machine Learning Engineer*

✉ 0505yogesh.p@gmail.com ☎ +91 8447386601 📍 Delhi-NCR, India 🔗 theyogeshpoonia.in

🌐 linkedin.com/in/yogeshkumar0001 🐙 github.com/yogeshpoonia0001

PROFESSIONAL SUMMARY

Applied AI & Machine Learning Engineer experienced in developing intelligent systems, multi-agent workflows, and semantic retrieval pipelines across fast-paced startup environments. Demonstrated experience in taking AI solutions autonomously from data evaluation to production deployment, engineering scalable backend APIs, and integrating vector search with strict schema validation. Google Cloud certified in foundation models, focused on delivering reliable, end-to-end AI softwares.

WORK EXPERIENCE

Artificial Intelligence Intern, *Padmas Technologies*

Jun 2026 – Aug 2026

- Architected multi-step reasoning workflows using LangGraph and LangChain to automate tool execution, improving multi-step task completion rate to 85%+ across benchmark suites.
- Integrated semantic search pipelines using ChromaDB vector embeddings, evaluating retrieval context to improve response consistency.
- Engineered asynchronous FastAPI endpoints and optimized PostgreSQL queries using Pydantic schemas, delivering 10+ core backend features.

Frontend Developer Trainee, *Fonix Technologies*

Apr 2025 – Mar 2026

- Promoted from Intern → Trainee in 5 months for high feature output and technical reliability.
- Shipped 11+ production web applications and multi-user administrative dashboards using Next.js and Tailwind CSS, increasing user retention by 65% and CSAT by 23% via rendering optimizations.
- Orchestrated modular UI component architectures with 3+ role-based access levels (RBAC), Stripe payment workflows, and 4+ locale translations (i18n).

Frontend Development Intern, *Cosmo Logic IT Solutions Private Limited*

Nov 2024 – Mar 2025

- Scaled React.js interfaces on AWS, reducing client-side load times and frontend defects by 36% through modular component refactoring.
- Developed administrative workflows supporting 50+ internal users with role-based access control (RBAC) for catalog and order management.

TECHNICAL SKILLS

Languages — Python, C++, SQL, JavaScript, C

Agentic AI & Orchestration — Multi-Agent Systems, LangGraph, LangChain (LCEL), Model Context Protocol (MCP)

Generative AI & LLMs — RAG Pipelines, Hugging Face Transformers, ChromaDB, Embeddings, GCP (Vertex AI), LLM Integration

Machine Learning & DL — PyTorch, Scikit-Learn, CNNs, XGBoost, Matplotlib, Fine-Tuning, Model Evaluation

Full-Stack & Systems — FastAPI, Next.js, React.js, PostgreSQL, Docker, RESTful APIs, Pydantic, Streamlit, Git, CI/CD

PROJECTS

Research Mind · Multi-Agent AI Research Engine 🔗

Aug 2026 – Sep 2026

Tech Stack: Python, LangGraph, LangChain (LCEL), ChromaDB, FastAPI, Pydantic, Tavily Search API, Streamlit

- Architected a multi-agent NLP pipeline coordinating 4 specialized agents for query decomposition, real-time web scraping, and multi-source context synthesis.
- Implemented dense-vector embeddings and semantic reranking in ChromaDB to ground tool outputs in sources, reducing unsupported claims by 12%.
- Built asynchronous FastAPI microservices and Pydantic models to ensure strict API data contracts, streaming real-time agent execution states and structured JSON reports to the client UI.

AI Barista · Grounded Conversational Support & Recommendation System

Jul 2026 – Aug 2026

Tech Stack: GCP Vertex AI, Gemini Models, Python, Vector Search, ChromaDB, Docker, Streamlit

- Engineered an end-to-end multi-turn conversational agent handling real-time intent classification across 15+ dialogue states to execute deterministic tools for inventory search and ordering.
- Designed and evaluated a semantic RAG pipeline indexing 200+ product catalog entities, applying nearest-neighbor vector search to maximize retrieval precision and response relevance.
- Containerized services using Docker and deployed production inference endpoints on Google Cloud Platform with strict schema enforcement.

EDUCATION

B.Tech, Computer Science Engineering | CGPA: 8.5, *Jaipur Engineering College, Kukas*

2023 – 2027

Relevant Coursework: Machine Learning, Deep Learning, Natural Language Processing, Data Structures & Algorithms, DBMS

HONORS & CERTIFICATIONS

Google Cloud Certified: Gen AI Academy (APAC) – Agent Orchestration & Vertex AI 🔗 ,

Certified in building, orchestrating, and deploying intelligent AI agents on Google Cloud.

Winner: Inter-University Hackathon (JECRC University), *Built an AI-driven Next.js platform*

Vice President E-Cell, *Led campus technical workshops, startup hackathons, and entrepreneurship initiatives*