

ALOK SINGH

AI Engineer

SUMMARY

AI Engineer with strong hands-on experience in building production-grade AI solutions including OCR pipelines, RAG-based agents, call summarization systems, and Model Context Protocol (MCP) integrations for enterprise databases. Skilled in Python, LLMs, automation, embeddings, and cloud deployment. Experienced in designing intelligent AI systems used in real business workflows. Passionate about building efficient, scalable AI tools and also exploring innovative personal projects like AI job-matching and smart scraping agents.

EDUCATION

IIT BHU, Varanasi

Bachelor's of Technology – Electrical Engineering
2020 – 2024

SKILLS

AI & LLMs

- RAG systems, OCR, MCP tools
- LLM agents, model fine-tuning
- Prompt engineering, embeddings, vector search

Programming & CS Fundamentals

- Python, JavaScript, Node.js
- Playwright automation
- REST APIs, microservices
- Data Structures & Algorithms (DSA)
- Strong problem-solving ability

ML & Data

- PyTorch / TensorFlow (if applicable)
- Data preprocessing, feature extraction
- Text summarization, classification

Frontend / Tools

- Vue.js, Tailwind CSS
- Docker, Render, Railway, Vercel

Databases

- PostgreSQL, MongoDB
- Vector DBs (FAISS / Pinecone optional)

DevOps / Deployment

- CI/CD Pipelines
- Docker
- Render, Railway, Vercel
- Automated deployments, build workflows

PROFESSIONAL EXPERIENCE

AI Engineer

ITAakash Strategic Software Private Limited | Aug 2024 – Present

- Built **production OCR pipeline** (LLMs + vision models) **processing 1000+ docs/day at 95%+ accuracy**.
- Developed **RAG agents** with optimized embeddings, **reducing query time by 70%** via **natural language interface**.
- Deployed **call summarization system** with sentiment analysis, **saving 15+ hrs/week** and generating **CRM-ready insights**.
- Integrated **MCP protocol** for secure **AI-database operations, boosting productivity by 40%**.
- Optimized workflows through **prompt engineering** and **context tuning** for **maximum accuracy and reliability**.

PROJECT

AI Job Matching Platform | [Link](#)

April 2025 – July 2025

- Built **LLM-powered matching system** with **embeddings and similarity** scoring to auto-match resumes to jobs via **skill extraction and ranking engine**.
- Automated **web scraping** using **Playwright + Gemini** for **intelligent job extraction** with **pagination detection and HTML parsing**.
- Developed **interactive dashboard** with **job cards, profile insights, and personalized recommendations**.