Yash Nilesh Gori

Mumbai, MH | yashnileshgori@gmail.com | 7718081766 | Portfolio | linkedin.com/in/yashgori20 | github.com/yashgori20 | Huggingface

Summary

AI Engineer experienced in building and deploying ML and GenAI systems, including Retrieval-Augmented Generation (RAG), predictive modeling, and scalable APIs. Experienced in building scalable AI platforms (GPT, Mixtral, Gemini), optimizing compliance systems, and deploying systems with sub-second latency.

Experience

AI Engineer (IC) | Webotix IT Consultancy   *Dec 2024 – June 2025*
- Architected and delivered an AI-powered QC automation platform using Azure OpenAI GPT-4o + RAG, slashing template creation time by 90% while sustaining 80% compliance accuracy.
- Designed multi-tenant architecture (Azure Cosmos DB + Redis) optimized for sub-second responses across enterprise clients.
- Integrated Azure Document Intelligence OCR to digitize and process legacy QC documents, enabling large-scale automation.
- Translated regulatory guidelines into 25+ parameterized compliance templates, creating a reusable pipeline for regulated industries.
- Deployed the platform into production as a B2B SaaS solution, securing $5K Microsoft AI Hub funding as Webotix's first enterprise AI product to receive external backing.
- Owned end-to-end execution: from technical architecture to deployment.

Business Analyst Intern | N. K. Engineering   *June 2024 – Nov 2024*
- Built automated Power BI dashboards tracking 15+ KPIs, uncovering ₹5Cr+ revenue opportunities in underserved engineering niches.
- Used predictive analytics to flag growth pockets; recommendations contributed to 12% market share expansion.
- Designed a responsive e-commerce platform and interactive product catalog, improving customer engagement.
- Applied ML-driven competitive analysis to identify five high-value investor prospects all closed.

Projects

Swift Check AI: QC Platform | *Azure OpenAI, Flask, Document Intelligence, Cosmos DB, Redis, Docker*
- Architected multi-tenant RAG microservice (Azure OpenAI + FAISS-style retrieval) for template generation and compliance validation; integrated Azure Document Intelligence OCR.
- Cut document generation time 90%; supported 25+ configurable parameters and maintained ~80% compliance accuracy with sub-second cached responses.

Financial Compliance Automation Tool | *Python, Mixtral LLM, GROQ Cloud, FAISS, Model Training*   Link
- Built custom RAG pipeline trained on RBI docs; iterative refinement achieved ~80% compliance prediction accuracy.
- Demonstrated 10% potential annual cost-savings opportunity for banks via automated checks.

Inhance: LinkedIn and Resume Profile Optimization Platform |*Streamlit, GROQ Cloud, Mixtral LLM, Multi-Agent System*   Link
- Implemented multi-agent evaluation and ATS-scoring pipeline; produced role-specific optimization suggestions and LaTeX resume generation.
- Delivered automated, actionable recommendations and an interactive enhancement agent for profile tuning.

DocuTalk: AI-Powered Document Intelligence Platform |*Python, FAISS, LangChain, Gemini Embeddings, Flask*   Link
- Engineered semantic search + conversational layer using Gemini embeddings + FAISS; handled context-aware Q&A for docs 100+ pages.
- Cross-platform deploy (Flutter) and reduced manual review from hours to minutes; adoption improved 40%.

Skills

AI/ML Core: Retrieval-Augmented Generation (RAG) systems, vector databases (FAISS, Pinecone), LLM integration (GPT, Mixtral, LLaMA, Gemini), OCR processing, NLP, Context Engineering, Huggingface Transformers.

Programming & Frameworks : Python, LangChain, Streamlit, FastAPI/Flask, Azure (OpenAI, Cosmos DB, Container Apps), Docker, GitHub Actions.

Data & Analytics: Pandas, NumPy, Seaborn, model evaluation (AUC, F1), Power BI.

Education

K. J. Somaiya College of Engineering   Mumbai, Maharashtra
B.Tech. in Information Technology | CGPA: 8.12/10   *2021–2025*

Impact and Achievements

Semi-finalist | Devopia 7.0 Hackathon - Top 3 teams for AI-powered solution.
Volunteer Instructor | Vacha NGO - Designed educational programs and taught for 30+ underprivileged students.
Microsoft AI Hub - Secured funding through competitive pitch for Swift Check AI platform