

Maarij Aqeel

📍 Faisalabad, Pakistan ✉️ maarijaqeel3200@gmail.com ☎️ +92 3328885574
🌐 maarijaqeel.vercel.app 🐙 github.com/Maarij-Aqeel 🔗 linkedin.com/in/maarijaqeel

Summary

AI Engineer with **1 year** of experience architecting production LLM applications and scalable backend systems, specializing in RAG pipelines, fine-tuned language models, and API-driven architectures serving **10K+ daily users**. Proven track record of deploying AI solutions that reduce operational costs by **75%+** and accelerate workflows from **days to hours**.

Technical Skills

- **Languages:** Python, JavaScript, TypeScript, SQL
- **AI & MLOps:** RAG Pipelines, LangChain, Ollama, Hugging Face, Fine-Tuning (LORA/QLORA), Vapi, PyTorch, TensorFlow
- **Backend & Cloud:** FastAPI, Django, PostgreSQL, MongoDB, Redis, Azure, AWS, Docker, CI/CD, Git
- **Frontend:** Next.js 16, React 19, Tailwind CSS, RESTful API Design
- **Automation:** n8n, Make, Zapier, Workflow Automation, Agentic Design

Experience

Contract AI Engineer

2025 – Present

Melonlabs Limited, 16x9 Studio / Remote

- Led backend and Voice AI development as Team Lead at **Melonlabs Limited** for Wordbuddy, managing database infrastructure and architecting the "Melon Tutor" real-time conversational agent (LiveKit, Deepgram, Gemini, ElevenLabs).
- Engineered AI content automation pipeline for **16x9 Studio** using n8n and LLM integrations, reducing manual creation time by **100%**.
- Deployed multi-tenant private cloud using OpenStack and Ansible, cutting developer setup time from **4 hours to 10 minutes (96% reduction)**.

AI Engineering Intern

Sep 2024 – Dec 2024

Software Productivity Strategists (SPS) / Remote

- Productionized FastAPI-based REST endpoints serving **10K+ daily requests**, optimizing PostgreSQL queries and implementing Redis caching to cut API response time from **450ms to 315ms (30% improvement)**.
- Integrated IBM WatsonX LLM into Greenkeeper app, engineering RAG pipelines that automated **12-step approval workflows**, reducing project approval time from **3 days to 4 hours (90% reduction)** for **50+ concurrent projects**.

AI & Backend Intern

Jun 2024 – Aug 2024

Nexium / Remote

- Architected Intervue AI platform using Next.js, Vapi Voice API, and n8n with dual-role authentication and real-time analytics, reducing interview setup time by **30%**.
- Productionized AI blog summarization pipeline using Gemini and LangChain, processing **200+ articles weekly** and reducing content review cycle by **75% (2 hours to 30 minutes)**.

Education

Bachelor of Science in Artificial Intelligence

Expected Graduation: 2027

FAST NUCES

Chiniot, Pakistan

Projects

VisionSnip – AI Video Clipping & Subtitling Engine

Python, Whisper, FFmpeg, LLMs

- Developed an AI-powered video processing pipeline utilizing OpenAI Whisper to analyze long-form content and automatically extract short clips with **high viral potential**.
- Implemented automated, highly accurate subtitle generation and temporal synchronization, **significantly reducing** manual video editing time for social media shorts.

RoleSync AI – ATS Resume Optimizer

Next.js 16, React 19, TipTap, Zustand

- Architected AI-powered resume optimization platform with real-time keyword injection via DOM-scanning AI, scoring resumes against JDs and reducing manual editing time by **80%**.
- Implemented dynamic LaTeX PDF export pipeline and a zero-upload cover letter generator featuring strict prompt engineering for technical recruiter personas.

Semantic Image Search – Multimodal Retrieval Engine

OpenAI CLIP, Pinecone, COCO Dataset

- Engineered multimodal search pipeline using OpenAI CLIP for embedding generation and Pinecone for vector indexing, achieving **<100ms retrieval latency** across **50,000+ image dataset**.
- Implemented text-to-image, image-to-image, and hybrid search workflows, enabling nuanced discovery of semantically relevant visual content.