M Maaz

Product-focused engineer with 4 years of experience specializing in distributed systems, serverless workflows, and applied AI. I design and ship architectures that scale reliably, cut latency, and reduce operational cost, whether rebuilding legacy systems into headless platforms, engineering edge-first applications, or deploying custom RAG pipelines. I combine pragmatic execution with data-driven decisions to deliver systems that perform in production

✉ mmaaaazu@gmail.com   📞 03353393916   📍 Karachi, Pakistan
🐙 https://github.com/mmaaaaz

PROJECTS

AskShu - University AI Agent   2024
- Engineered a serverless RAG pipeline entirely on Cloudflare Edge, eliminating cold starts on typical Lambda Functions
- Implemented a Python ingestion engine to scrape, chunk, and vectorize thousands of documents directly into Cloudflare Vectorize, removing reliance on managed vector DBs
- Integrated LangChain, Cloudflare Workers/DB, LLM and custom prompt tuning
- Deployed Llama 3.3 70B via Workers AI with <500ms TTFT, cutting inference cost by ~60% vs GPT-4 API usage.

ZenMall - Headless Commerce Platform   Jun 2024 - present
- Achieved 99/100 Core Web Vitals using RSC, edge caching, and optimized data loading
- Engineered a consistency-safe cart system using 3-way sync (URL params, LocalStorage, Atoms)

AhsanAutos.pk   Oct 2021 - March 2023
- Migrated BigCommerce monolith to a Headless Next.js, reducing bounce rate by 40%
- Integrated BigCommerce GraphQL APIs to unlock custom checkout flows and dynamic product logic
- Integrated BigCommerce GraphQL APIs to unlock custom checkout flows and dynamic product logic

SmartInsure.pk   Dec 2021 - Jun 2022
- Developed a real-time insurance quoting engine (React + Firebase) delivering complex quotes <2s
- Implemented Cloud Functions + SendGrid workflows to automate transactional communications

EDUCATION

Salim Habib University

Bachelor of Science in Computer Science   2022 - ongoing
- Current Status: 7th Semester (Evening/Flexible schedule available for part-time roles).

SKILLS

- AI Engineering: RAG, Vector Search, LangChain, OpenAI / Workers AI
- Frontend: React, Next.js, TypeScript, Astro, Svelte
- Backend & Infra: Cloudflare Workers, Node.js, Hono.js, GraphQL, REST
- Data & Storage: PostgreSQL, Supabase, Firebase, Redis
- DevOps & Platforms: Docker, GitHub, Cloudflare, Vercel, Turborepo, CI/CD