

MUHAMMAD ASIF JAVED

axif.javed@gmail.com | +92 332 464 7331 | Lahore, Pakistan | [linkedin.com/in/maxifjaved](https://www.linkedin.com/in/maxifjaved) | github.com/maxifjaved

PROFESSIONAL SUMMARY

AI/ML Engineer specializing in LangChain RAG pipelines and Hugging Face fine-tuning—delivering 40% inference-latency reduction, 35% cost savings, and 99.9% uptime for 25K+ users. Built & scaled six production platforms across edtech, enterprise, BI, and cloud-cost optimization. Seeking to drive AI-first solutions at innovative technology companies.

TECHNICAL SKILLS

AI/ML & Data: Python · PyTorch · Transformers · LangChain · RAG Systems · Vector DBs · Prompt Engineering · OpenAI API
Web & Infra: React · Next.js · Node.js · FastAPI · REST/GraphQL · Docker · AWS (EC2/S3/Lambda) · CI/CD · PostgreSQL · MongoDB
Also: Vue.js · Laravel · Redis · WebRTC · Ember.js · Angular · Microservices

PROFESSIONAL EXPERIENCE

Founder & AI/ML Engineer | QuranGallery.App

April 2023 – Present

- Lead architect for LangChain RAG pipelines (50+ languages), cutting inference latency 40% and AWS spend 35% while serving 25K users at 99.9% uptime
- Fine-tune Hugging Face models on 100K-sample corpus, boosting answer relevance 25% and reducing manual curation 30%
- Manage \$2K/mo AWS infra with auto-scaling (EC2, Lambda) and observability (CloudWatch), handling 1M+ audio requests monthly
- Build real-time, multi-modal learning features—interactive tools and specialized memorization systems—using React, FastAPI, and WebSockets

Technologies: Python, PyTorch, Hugging Face, LangChain, React, Next.js, FastAPI, AWS

Senior Software Engineer | DigitalDot.io

January 2021 – February 2023

- Led 8-member team to re-architect monolith into microservices, cutting downtime by 30% and accelerating delivery by 20%
- Built CI/CD pipelines with GitHub Actions and Docker, reducing manual deploy errors by 90% across 5 major releases
- Migrated legacy systems supporting 10K+ concurrent users with zero data loss, managing \$15K/mo infrastructure budget
- Collaborated with 3-person product team to deliver scalable technical solutions meeting diverse client requirements
- Implemented performance optimizations improving application response times by 50% under peak load

Technologies: React, Node.js, TypeScript, Docker, PostgreSQL, MongoDB, Microservices

Web Application Developer | Renesis Tech

May 2017 – January 2021

- Developed full-stack applications for 15+ clients across healthcare and finance, managing \$50K+ project budgets annually
- Implemented performance optimizations reducing page load times by 60% for high-traffic applications serving 5K+ daily users
- Built custom APIs supporting complex business workflows with 99.5% reliability and sub-200ms response times
- Led technical support team maintaining sub-4-hour issue resolution SLA across production systems
- Architected database solutions and integrations for enterprise clients requiring HIPAA compliance

Technologies: JavaScript, PHP Laravel, MySQL, Performance Optimization

Software Engineer | AA'LA Solutions

April 2015 – May 2017

- Developed applications using React, Ember.js, and Node.js across 20+ project deliveries for diverse client base
- Collaborated with 4-person development teams from planning through deployment, maintaining 95% client satisfaction
- Provided technical support resolving complex issues within 24-hour SLA, improving customer retention by 20%
- Contributed to architecture decisions ensuring scalable solutions meeting long-term business objectives

Technologies: React, Ember.js, Node.js, JavaScript, Laravel

FEATURED PROJECTS

[QuranGallery.App](#) — Founder & Lead Engineer; RAG-powered edtech (50+ langs), 25K users, 40% latency cut
[altconsole.com](#) — Architect; real-time cloud-cost optimization, 30% spend reduction (AWS/GCP)
[crystalball.pro](#) — Full-Stack Dev; BI platform with auto reporting, generating 10K daily insights

EDUCATION & CERTIFICATIONS

MSc, Computer Science, PUCIT — May 2015

BSc, Government College University, Lahore — May 2013

Certifications: Meta Front-End (Aug 2023); IBM Full-Stack (Jul 2023); freeCodeCamp Full-Stack (Oct 2023); React Dev (Jun 2023)

Additional: Git & GitHub Essentials · Cloud Computing · Node.js & Express

Languages: English (Professional), Urdu (Native) • **Open to:** Remote roles globally