

Muhammad Huzaifa Shahbaz

AI Software Engineer

mhuzafadev@gmail.com | linkedin.com/in/mhuzafadev | github.com/mhuzafadev | mhuzafa.com | +92 303 2800071

AI Software Engineer with 4+ years of experience designing and building production AI applications, backend systems, and enterprise automation for international clients across the US, UK, and UAE. My work spans applied machine learning, intelligent agents, real-time systems, and cloud infrastructure, with a focus on building reliable and practical solutions. Alongside engineering, I have contributed to research, open-source projects, and technical communities through workshops, mentoring, and student leadership.

SKILLS

- **Programming:** Python, TypeScript, JavaScript, Go, SQL
- **AI & Applied ML:** PyTorch, TensorFlow, LangGraph, LangChain, RAG, agentic AI systems, multimodal AI, computer vision, LLM workflows, Sentence-BERT, model fine-tuning and inference optimisation
- **Backend & Product Engineering:** FastAPI, Django, Node.js, REST APIs, WebSockets, PostgreSQL, Redis, MongoDB, event-driven systems, real-time and distributed backends
- **Cloud & Infrastructure:** AWS, Azure, GCP, Docker, Kubernetes, Linux, GitHub Actions, CI/CD, Nginx, GPU inference, CUDA
- **Automation & AI Platforms:** n8n, Retell AI, RabbitMQ, Celery, vector search, self-hosted LLMs, voice AI

PROFESSIONAL EXPERIENCE

AI Systems Engineer

Sep 2024 — Present

[Digital Dividend Global](#) | Digital Transformation Company based in Sweden

- Working on a system that communicates multiple agent systems with HITL as the standard procedure, built on both on-prem LLM and OpenAI, using Nintex K2 and Microsoft Power Automate for MOCCA, UAE.
- Actively working on an internal enterprise multi-agent platform for a Lebara - UK telecommunication client, including self-healing Databricks/Azure Data Factory workflows and automated code-review agents, operating within GDPR, security and AI governance requirements.
- Working on an AI CFO platform integrating QuickBooks and Dropbox for Apparatus Contractor Services in the US, serving high-value business customers, addressing access control, financial-data privacy and AI governance requirements at the application layer.
- Lead backend and AI/IoT workstreams for a 3-engineer team, owning platform services and production delivery
- Built PseudoCompiler, a real-time WebSocket translation platform (FastAPI, AWS), scaling to 50,000+ concurrent users at under 50ms latency while cutting AWS infrastructure cost by ~35%
- Designed the OptyAI recommendation engine for OptySleep using Redis caching, A/B testing, and Sentence-BERT semantic similarity, contributing to a ~70% increase in iOS app engagement
- Delivered the Staging Pro IoT backend (Django, PostgreSQL, Redis, WebSockets) processing live race telemetry at under 100ms latency

Backend Engineer

Jul 2022 — Aug 2024

[Lenaar Digital LLC](#) | AI Software Services Company, (Remote, USA)

- Built and scaled backend systems on Azure and Kubernetes for products used by 100,000 to 1,000,000+ users.
- Built multi-agent computer vision pipelines using YOLO, DeepSORT and SAM2 for automated detection, tracking and tagging.
- Deployed self-hosted LLMs and automated CI/CD using Docker and GitHub Actions, reducing deployment time and infra overhead.

Machine Learning Engineer, Research Team

Jan 2022 — Jun 2022

[AKDN Digital Health Resource Centre](#), Aga Khan University Hospital, (Karachi, Pakistan)

- Built machine learning pipelines for AI-assisted breast cancer screening on mammography and ultrasound imaging, achieving 92% accuracy
- Contributed to an 87% reduction in radiology review time through automated triage, sorting, and tagging tools developed in collaboration with clinicians. Work was conducted within a hospital lab incubation setting linked to final-year thesis research (Mammory)

Machine Learning Research Intern

May 2020 — Oct 2020

[Virufy Organization](#) | COVID-19 Audio ML Research Organisation (Remote, USA)

- Developed signal-processing and ensemble-ML pipelines for COVID-19 detection from cough & breath audio; prototype achieved 89% accuracy.

EDUCATION

B.S. Software Engineering,

2019 — 2023

[Sir Syed University of Engineering & Technology \(SSUET\)](#), Karachi, Pakistan

AWARDS & HONOURS

- ✓ Sponsored Trip to KSA for Leap 2026 by Digital Dividend Global for best Overall Performance & Client Communication
- ✓ Highest GPA in Final Year Project in Sir Syed University of Engineering & Technology, 2023 (Mammory)
- ✓ Microsoft for Startups Founders Hub Award, USD 25,000 in Azure credits, 2024

PUBLICATIONS

- M. H. Shahbaz, N. Ullah, and M. Ahmed, “**AI-Assisted Seismic Performance Assessment of Sustainable High-Rise Structures Using Smart Material Technologies in Pakistan**,” *Spectrum of Engineering Sciences*, vol. 4, no. 5, pp. 1724–1734, May 2026. DOI: [10.5281/zenodo.20305243](#)
- U. Ali, I. A. Kandhro, R. S. Ahmed, A. A. Khan, M. H. Shahbaz, and M. Osama, “**The Future of Third Web: A Role of Blockchain and Web 3.0**,” *International Journal of Emerging Sciences and Digital Economy (IJESDF)*, vol. 17, no. 3, pp. 391–403, Apr. 2025. DOI: [10.1504/IJESDF.2025.145880](#)
- K. Mahboob, M. H. Shahbaz, F. Ali, and R. Qamar, “**Predicting the Karachi Stock Price Index with an Enhanced Multi-Layered Sequential Stacked Long-Short-Term Memory Model**,” *VFAST Transactions on Software Engineering*, vol. 11, no. 2, pp. 249–255, Jun. 2023. DOI: [10.21015/vtse.v11i2.1571](#)
- M. H. Shahbaz, Zain-Ul-Abidin, K. Mahboob, and F. Ali, “**Enhancing Contextualized GNNs for Multimodal Emotion Recognition: Improving Accuracy and Robustness**,” in *Proc. 7th Int. Multi-Topic ICT Conf. (IMTIC)*, Jamshoro, Pakistan, 2023, pp. 1–7. DOI: [10.1109/IMTIC58887.2023.10178481](#)
- T. Mubeen, Zain-Ul-Abidin, M. H. Shahbaz, P. O. Roth, and M. A. L. Nieto, “**Analyzing the Classification Performance of DenseNet121 on Pre-processed MIAS Dataset**,” in *Proc. Global Conf. Wireless Opt. Technol. (GCWOT)*, Málaga, Spain, 2023, pp. 1–7. DOI: [10.1109/GCWOT57803.2023.10064663](#)
- M. H. Shahbaz, Z. Ul Abidin, U. Marfani, M. Abbasi, and T. Mubeen, “**Mammory – Breast Cancer Detection using AI on Mammography & Ultrasonography**,” B.S. Thesis, Sir Syed University of Engineering & Technology, Karachi, Pakistan, 2023. DOI: [10.6084/m9.figshare.22340170.v1](#)
- M. H. Shahbaz, R. W. Baig, and Zain-ul-Abidin, **PyKids 2023 (book)**, Apr. 2023. ISBN: 9798390752586 – Amazon Kindle. <https://www.amazon.com/dp/B0C1YBFWZY>

SELECTED PROJECTS

Agentic AI Internal Ops – Ministry of Environment and Climate Change, UAE

2026–Present

Contributing through my employer to the technical design of an AI modernization initiative proposed for the UAE Ministry of Climate Change and Environment (MOCCA), focusing on environmental data workflows, responsible AI and human oversight.

Agentic AI Code Review & Data ETL pipeline – Lebara

2026

Developed an agentic orchestrated pipeline on the top of Databricks and Azure Data factory, for their internal ops, Code reviews, runtime QA, and quick runtime analysis on Data ETL pipelines.

MyCFOSimon AI – Apparatus Contractor Services

2025–2026

A virtual avatar-based CFO that connects to clients' individual accounts data and QuickBooks, reporting real-time suggestions and providing advice just like a human CFO, designed exclusively for their clients rather than as a subscription service.

VikingRAG – (Open-Source AI Research)

Independently implemented a structure-aware RAG system with hierarchical retrieval, adaptive search, experience-based learning, and configurable retrieval budgets using Python, FastAPI, and PostgreSQL/pgvector. Release: <https://github.com/mhuzaifadev/VikingRAG>

AI Legal Assistant

2023–2024

Retrieval-Augmented Generation system (Next.js, LangChain, Gemini) for citation-backed legal document review, reducing review time by approximately 70%

CERTIFICATIONS

Amazon Web Services Solutions Architect Associate (AWS SAA C03), *Oct 2026*

Google AI Essentials, Google (via Coursera), *Sep 2025*

Building AI Voice Agents for Production, DeepLearning.AI, *Jun 2025*

Microsoft Certified: Azure AI Fundamentals, Microsoft, *Jan 2021*