TALHA SHAHZAD

Full Stack Developer | AI / ML Engineer

Dera Ghazi Khan, Pakistan • +92 334 657 3003 • mtalhashzdkhn53@gmail.com

linkedin.com/in/talha-shahzad-554614284 • github.com/MtalhaShzd • talhashahzad.lovable.app

PROFESSIONAL SUMMARY

AI Engineer & Full-Stack Developer with hands-on experience building and deploying machine learning, Generative AI, RAG, and full-stack applications. Experienced in Python, FastAPI, LLM integration, vector search, MLOps, REST APIs, Flutter, Laravel, and cloud deployment. Built production-ready systems including multimodal RAG applications and an end-to-end time-series AQI forecasting platform with automated data pipelines, model registry, explainability, and continuous deployment. Focused on developing practical, reliable AI systems that integrate seamlessly with real-world applications.

TECHNICAL SKILLS

Languages	Python, JavaScript (ES6+), PHP, Dart, C#, C++, SQL, Bash, Java
Machine Learning	Scikit-learn, Ridge Regression, Random Forest, XGBoost, LSTM (TensorFlow / Keras), ResNet50 Transfer Learning, Persistence Baselines, Feature Engineering (temporal, cyclical, lag, rolling, momentum), Recursive Multi-Step Time-Series Forecasting, Chronological Train/Val/Test Splits, Model Evaluation (MAE / RMSE / R^2), SHAP Explainability, Pandas, NumPy, Plotly, Matplotlib, Seaborn
MLOps	Hopsworks Feature Store (online RonDB + offline Delta/Parquet, Kafka streaming ingestion), Hopsworks Model Registry, Model Versioning & Promotion Gating, Automated Feature / Training / Backfill Pipelines, GitHub Actions Scheduled CI/CD, Data Validation & Gap Detection, Fallback & Self-Healing Architecture, Google Colab, Jupyter
AI / GenAI	RAG (Retrieval-Augmented Generation), LangChain, LLM Integration (Groq LLaMA 3.3 70B / LLaMA 4), FAISS Vector DB, HuggingFace Embeddings, Semantic Search, Prompt Engineering, OCR (Tesseract), Speech-to-Text, gTTS
Agentic AI / MCP	Model Context Protocol (MCP) — server & client development, JSON-RPC messaging, STDIO & StreamableHTTP transports, sampling, roots and notifications; Claude Code; AI agent & tool-integration workflows
Backend / Web	FastAPI, Flask, Gunicorn, Laravel, React.js, Flutter, Streamlit, ASP.NET / .NET WinForms, REST API Design, Swagger / OpenAPI, JWT & Firebase Auth
Databases	MySQL, PostgreSQL, SQL Server (T-SQL), Firebase Firestore / Realtime DB, SQLite, FAISS
Cloud & Deploy	Render, Streamlit Community Cloud, Hopsworks Serverless, Firebase (Auth, Firestore, Storage), GitHub Actions CI/CD, Linux (Ubuntu), AWS Fundamentals, Microsoft Azure Fundamentals
IoT / Embedded	ESP32-S3 CAM, Arduino IDE, NEO-6M GPS (TinyGPS++), MJPEG Streaming, OpenCV, Leaflet.js
Tools & Practices	Git, GitHub, VS Code, Visual Studio, Postman, MVC Architecture, RBAC, OWASP Security, DB Normalisation (3NF), Data Structures & Algorithms (C++)
CMS / Other	WordPress, WooCommerce, Elementor, Admin Dashboard Design, Provider (Flutter state Management),GoRouter

PROFESSIONAL EXPERIENCE

Data Science Intern — 10Pearls Shine Internship Program (Cohort 9) *Jul 2026 – Sep 2026*

10Pearls Pakistan — 8-Week Program | Project: Pearls AQI Predictor (live in production)

- Built and deployed an **end-to-end MLOps system** forecasting the Air Quality Index for Lahore across 24/48/72-hour horizons — covering automated ingestion, feature store, comparative model training, model registry, explainability, REST API and dashboard.
- Engineered **29 features** (temporal, cyclical sine/cosine encodings, 1h/24h lags, 6h rolling mean & std, AQI change-rate) over a validated hourly dataset spanning January 2024 onward, sourced from the Open-Meteo Weather and Air Quality APIs.
- Trained and objectively compared **five modelling approaches** — Ridge Regression, Random Forest, XGBoost, LSTM and a persistence baseline — on a chronological 70/15/15 split; selected **Ridge Regression** (validation MAE 1.00, RMSE 1.52, R^2 0.9988; test R^2 0.9868, 90% of predictions within ±7 AQI) for accuracy, retraining speed and exact SHAP interpretability.
- Implemented a **Hopsworks Feature Store** feature group with streaming (Kafka) ingestion and both online (RonDB) and offline stores, plus a **Model Registry** gated by an --only-if-better RMSE check to prevent version sprawl from daily retraining.
- Built a **recursive 72-step forecast engine** that feeds each prediction back into lag/rolling features, and per-horizon **SHAP explanations** via LinearExplainer against the scaled Ridge pipeline.
- Shipped a **Flask REST API** (9 endpoints, Gunicorn on Render) and an interactive **Streamlit dashboard** (Plotly charts, health classifications, downloadable forecasts, SHAP visualisations) on Streamlit Community Cloud.
- Automated the full lifecycle with **GitHub Actions**: hourly feature ingestion, daily retraining and registration, and a scheduled gap-detecting backfill pipeline — zero manual intervention.
- Independently **validated against Punjab EPA** official monitoring data — live prediction of 166 AQI vs. the government reading of 160 (~4% deviation)

Full Stack Web, Mobile & AI Developer *Jan 2024 – Present*

Freelance / Self-Employed — Remote (Direct & Local clients)

- Architected and deployed a **multimodal RAG chatbot** (Flutter + FastAPI + FAISS + Groq LLaMA 3.3 70B + Firebase) ingesting PDF, DOCX, PPTX, XLSX, CSV, images (OCR) and MP4 (speech-to-text), with bilingual English/Urdu voice I/O and real-time Firestore chat history.
- Designed and delivered **10+ production web applications** in PHP, Laravel, React.js and WordPress following MVC architecture, role-based access control and OWASP security standards.
- Built an **IoT AI drone waste-detection system**: ESP32-S3 CAM firmware (C++) + Flask server + ResNet50 transfer-learning model (~80% accuracy) with live GPS geospatial mapping in Leaflet.js.
- Developed cross-platform **Flutter mobile apps** on a Firebase backend (Auth, Firestore, Storage), shipping Android and iOS from a single codebase.
- Built a retail **POS desktop application** (C#, WinForms, SQL Server, T-SQL) with inventory, billing and sales reporting, plus a **Real Estate Rental Management System** on a fully 3NF-normalised schema.
- Implemented and optimised relational schemas (MySQL, PostgreSQL), REST APIs and admin dashboards; documented all FastAPI endpoints with **Swagger / OpenAPI**.

KEY PROJECTS

Pearls AQI Predictor — Production MLOps Forecasting System — *Python · Scikit-learn · XGBoost · TensorFlow · Hopsworks · SHAP · Flask · Streamlit · GitHub Actions*

GitHub | Live Dashboard

- Nine-stage pipeline documented across Jupyter notebooks (collection → preprocessing → EDA → feature engineering → feature store → model comparison → registry → SHAP → recursive forecasting) and productionized into automated scripts.
- Model comparison: Ridge Regression (selected, R^2 0.9988 validation) vs. persistence baseline, Random Forest, XGBoost and LSTM, evaluated on MAE, RMSE and R^2 with a non-shuffled chronological split reflecting real forecasting conditions.
- Engineered a resilient production architecture with Hopsworks/local-CSV freshness checks, model-artifact fallback with 30-minute auto-retry, three-tier weather API fallback, computation caching, and graceful dashboard degradation
- Deployed the complete ML system on cloud free-tier infrastructure using Render, Streamlit Community Cloud, Hopsworks Serverless, and GitHub Actions, implementing safeguards for cold starts, API limits, and platform quota constraints.

Multimodal RAG Chatbot — *Flutter · FastAPI · Groq LLaMA 3.3 70B · FAISS · Firebase · Python*

github.com/MtalhaShzd/rag-chatbot-app

- Full-stack AI assistant with General Knowledge and Document-Based Q&A modes, powered by FAISS semantic search over HuggingFace sentence-transformer embeddings.
- Multimodal ingestion (PDF, DOCX, PPTX, XLSX, CSV, images via OCR, MP4 via speech-to-text), Firebase Auth (email + Google Sign-In), bilingual voice I/O via gTTS, dark/light theme and searchable grouped chat history.
- Architecture: Flutter frontend (Provider + GoRouter) → FastAPI backend → Groq LLM → FAISS retrieval → Firestore storage, with full Swagger API documentation.

AI Voice Chatbot with RAG — *Python · Streamlit · Groq LLaMA 4 · FAISS · Tesseract OCR · FFmpeg*

github.com/MtalhaShzd/ai-rag-voice-chatbot

- Enterprise-grade RAG assistant with multi-source ingestion (PDF, DOCX, images via Tesseract OCR, MP4 via FFmpeg transcription) and a persistent FAISS vector index across sessions.
- Groq LLaMA 4 inference for sub-second responses, English/Urdu voice input and output, index reset controls and secrets-based API key handling.

IoT AI-Driven Drone Waste Detection & Geospatial Mapping — *C++ · Python · ESP32-S3 · TensorFlow · ResNet50 · Flask · OpenCV · Leaflet.js*

github.com/MtalhaShzd/IoT-based-AI-Driven-Drone-Waste-Detection-Geospatial-Mapping-System

- 6th-semester IoT capstone: ESP32-S3 CAM streams MJPEG video to a Flask server running a ResNet50 transfer-learning classifier (~80% validation accuracy) labelling scenes clean or garbage every 10 frames.
- GPS coordinates polled every 5 seconds (NEO-6M, TinyGPS++) and plotted live on an interactive Leaflet.js map with colour-coded markers.

POSApp & Real Estate Rental Management System — *C# · SQL Server (T-SQL) · WinForms · .NET*

github.com/MtalhaShzd/POSApp | github.com/MtalhaShzd/RealEstateRentalManagementSystem

- Retail POS for a store: product catalog, inventory management, billing engine and sales reporting dashboard.
- Rental platform covering tenant records, lease lifecycle and financial reporting on a 3NF-normalised schema with full PK, FK, UNIQUE, CHECK and NOT NULL constraints.

EDUCATION

B.S. Information Technology (In Progress) *Sep 2023 – 2027 (Expected)*

Ghazi University, Dera Ghazi Khan, Punjab, Pakistan

Relevant coursework: Data Structures & Algorithms, Object-Oriented Programming, Database Management Systems, Software Engineering, Web & Mobile Application Development, Artificial Intelligence, Computer Networks, Information Security, Internet of Things.

CERTIFICATIONS

- AI Fluency: Framework & Foundations — Anthropic *(Jul 2026)*
- Introduction to Model Context Protocol — Anthropic *(Jul 2026)*
- AI For Everyone — DeepLearning.AI *(Dec 2024)*
- Networking Basics — Cisco Networking Academy *(Dec 2024)*
- AWS Cloud Fundamentals
- Model Context Protocol: Advanced Topics — Anthropic *(Jul 2026)*
- Claude Code 101 — Anthropic *(Jul 2026)*
- Microsoft Azure SQL — Microsoft (Coursera) *(Dec 2024)*
- Introduction to Cybersecurity — Cisco Networking Academy *(Sep 2024)*
- Microsoft Azure Fundamentals

ADDITIONAL INFORMATION

Languages	English — Professional Fluency | Urdu — Native
Availability	Open to remote, hybrid and on-site opportunities
Interests	Applied AI & LLM Engineering, RAG Systems, Agentic AI & MCP, MLOps, Data Science, Full Stack Development