

Khuzaima Hassan

AI/ML Engineer | Data Science Graduate | Python Developer

Karachi, Pakistan | khuzaimahassan52@gmail.com | +92 321 3937589 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

PROFESSIONAL SUMMARY

Data Science Graduate from NED University with hands-on experience in machine learning and Generative AI, gained through a Data Science Internship at 10Pearls. Proficient in Python, SQL, FastAPI, and Streamlit, with practical experience integrating LLM APIs (Gemini, Groq) and open-source models (Qwen, Llama) for real-time data Q&A, automated reporting, and interactive dashboards. Strong foundation in classification, regression, Deep Learning (LSTM), feature engineering, model evaluation, and end-to-end ML pipeline design.

TECHNICAL SKILLS

Programming: Python, SQL

Machine Learning: Classification, Regression, Deep Learning, LSTM, Feature Engineering, Data Preprocessing, Model Evaluation, Cross-Validation, Model Monitoring, Pipeline Design

Generative AI & LLMs: Gemini API, Groq API, LangChain, LangGraph, Qwen, Llama, LLM Integration, Prompt Engineering, Retrieval-Augmented Q&A, GenAI App Development

Frameworks & Tools: FastAPI, Streamlit, Git, GitHub, GitHub Actions, Hopsworks, Jupyter Notebook, Hugging Face

Databases & Visualization: PostgreSQL, MySQL, Power BI, Pandas, NumPy, Matplotlib, Seaborn

PROFESSIONAL EXPERIENCE

10Pearls | Data Science Intern

Dec 2025 – Feb 2026

- Built an end-to-end ML pipeline for an AQI prediction system, covering data collection, cleaning, and feature engineering using Python, Pandas, scikit-learn, and Hopsworks.
- Automated data pipelines and model deployment workflows utilizing GitHub Actions.
- Trained and evaluated regression and classification models to forecast AQI levels, applying cross-validation to improve generalization.
- Developed and deployed an interactive dashboard to visualize AQI predictions and environmental insights, supporting data-driven monitoring decisions.

Japanese Language & IT Internship (Plus W x OEC) | Internee

Apr 2025 – Sep 2025

- Worked on database-related tasks to support backend development, adding unique constraints to ensure data integrity and prevent duplicate records.
- Wrote and executed database migration scripts to safely update existing tables and maintain data consistency across environments.

PROJECTS

TeleTabib – AI Telehealth Agent (Final Year Project)

- Architected an AI-driven telehealth agent utilizing LangGraph and open-source LLMs to automate patient interactions.
- Integrated Agentic workflows and Retrieval-Augmented Generation (RAG) over medical document datasets to ensure accurate, domain-grounded responses.
- Collaborated within a 4-person team under academic and industrial supervision to deliver a scalable, production-ready healthcare application.

Risk Scoring Platform – End-to-End ML System | [Github](#)

- Built an end-to-end fraud risk scoring platform using Python, PostgreSQL, and FastAPI, covering data ingestion, feature engineering, and model training.
- Implemented model versioning and a model monitoring layer with statistical drift detection to track prediction distributions and flag performance degradation.

Insight Agent – LLM-Powered Data Analysis App | [Github](#)

- Integrated Google's Gemini API to build an LLM-powered app answering natural language queries over user-uploaded datasets in real time.
- Built a Streamlit dashboard with automated dataset summaries, visualizations, and an interactive Q&A interface for non-technical users.

Real-Time Fraud Detection System – Machine Learning | [Github](#)

- Addressed class imbalance using SMOTE; trained and compared multiple ML classifiers with threshold tuning to optimize precision-recall tradeoff.
- Evaluated model performance using precision, recall, F1-score, and ROC-AUC, visualized via an interactive Streamlit dashboard.

EDUCATION

NED University of Engineering & Technology | B.S. Data Science

Sep 2022 – Jun 2026

CGPA: 3.885/4.00 (First Position in Academic Ranking)

Coursework: Machine Learning, Database Systems, Algorithms, Data Visualization

Intermediate (Pre-Engineering) — Adamjee Government Science College

Aug 2020 – Aug 2022

Matriculation — Usman Public School System

Apr 2018 – Apr 2020

CERTIFICATIONS & LANGUAGES

Certifications: Machine Learning Scientist in Python (DataCamp) | Data Analyst in SQL (DataCamp) | Python for Data Science, AI & Development (Coursera) | Machine Learning Engineer (DataCamp) | Huawei HCIA (Datacom)

Languages: English (Professional) | Urdu (Native/Bilingual) | German (Elementary) | Japanese (Elementary)