

Khuzaima Hassan

Data Scientist

khuzaimahassan52@gmail.com | 923213937589 | khuzaima-hassan-2004mkh | github.com/KhuzaimaHassan | Karachi, Pakistan

PROFILE

Data Science undergraduate with strong problem-solving skills and hands-on experience in machine learning, LLM-powered applications, and data-driven systems. Skilled in Python, SQL, Streamlit, and model evaluation pipelines. Passionate about building intelligent tools that simplify data exploration and decision-making.

PROFESSIONAL EXPERIENCE

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

01/2025 – Present | Remote

- Worked on real-world tech assignments.
- Gained practical experience with IT workflows, documentation, and coding tasks.
- Collaborated on internal prototypes using Python and data visualization.

PROJECTS

Personalized News Summarization Agent (Capstone), Kaggle 5-Day Generative AI Course

- Built a pipeline that ingests news, produces personalized summaries using fine-tuned/lightweight models and ranks results; focused on evaluation, prompt design, and UX for non-technical users.
- Implemented end-to-end demo with reproducible notebooks and evaluation metrics.

AI Academic Mentor, Data Science Hackathon

- Integrated Groq API to generate personalized student academic reports based on performance.
- Built a streamlit UI to showcase student data and report generation.

Insight Agent, LLM Integration & Streamlit App

- Integrated Gemini API (LLM) to answer queries over uploaded datasets in real time.
- Built a Streamlit dashboard to showcase dataset summary, visualizations, and interactive Q&A.

Real-Time Fraud Detection System, Machine Learning

- Designed balanced datasets using SMOTE, threshold tuning, and multiple ML models.
- Prepared evaluation metrics and Streamlit dashboard integration.

Real Estate Management System, MySQL Database

- Designed normalized relational schema and CRUD operations for property listings.
- Implemented stored procedures and query optimizations to support common use cases (search, filter, transactions).

EDUCATION

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

09/2022 – Present

- CGPA: 3.87
- Relevant coursework: Machine Learning, Database Systems, Algorithms, Data Visualization

Intermediate, Adamjee Govt. Science College

08/2020 – 08/2022

- Secured A-1 Grade with 93.51% marks.

Matriculation, Usman Public School System

04/2018 – 04/2020

- Awarded consecutive shields for achieving A+ grades in all subjects.

TECHNICAL SKILLS

Python

pandas, numpy, scikit-learn, matplotlib, seaborn

ML & AI

Classification models, regression, model evaluation, cross-validation, pipeline design, LLM integrations

Visualization

Power BI dashboards, modeling

SQL

Database design, Query optimization, Joins, Indexing, CRUD operations

Tools

Streamlit, Hugging Face, Jupyter Notebook, Git/GitHub

Networking

Huawei eNSP (RIP, VLAN, DHCP, Eth-Trunk/LACP)

CERTIFICATION

Machine Learning Scientist in Python

Kaggle

Python for Data Science, AI & Development

Coursera

Huawei Certified ICT Associate (HCIA)

Datacom

Data Analyst in SQL

DataCamp

Power BI Fundamentals

DataCamp

LANGUAGES

English: Professional working proficiency

Urdu: Native or bilingual

German: Elementary proficiency

Japanese: Elementary proficiency