



Hafiz Muhammad Tariq Raza

☎ Phone number: (+92) 3430066353

✉ Email address: muhammadtariq2001@outlook.com

🌐 LinkedIn: <https://www.linkedin.com/in/tariqinai/>

🐙 GitHub: <https://github.com/tariqinai>

WORK EXPERIENCE

🏠 Upwork & Fiverr - Remote (Part-Time)

Freelance Data Scientist -

[09/2021 – Current]

- **For Humans (Germany):** Architected a fully **automated Shopify analytics pipeline** (Python, Pandas, Streamlit) that ingests daily sales data, calculates 15+ KPIs (**CLV, CAC, repeat rate**), and delivers real-time interactive dashboards—slashing report generation time by 90% and empowering data-driven product decisions.
- **Auto Finance Canada (Canada):** Designed and deployed an **automated data integrity** system (Python, Google Sheets API) for over **10M customer records**. The system utilized machine learning-based **anomaly detection** alongside rule-based validation, **reducing manual data preparation** by **80%** while ensuring regulatory compliance.
- **MaineWorks & YourModernAccountant:** Developed advanced **Power-BI & Looker Studio dashboards** (SQL, Python) to monitor workforce reintegration success metrics and **financial KPIs**—boosting reporting **efficiency by 75%**, uncovering actionable insights, and driving repeat engagements.
- **Trusted by 150+** international clients for end-to-end data science solutions (**ETL, predictive analytics, forecasting, visualization**), earning top-rated status and consistent 5-star feedback for delivering measurable business impact.

🏠 MECHESOL CO - Hybrid (Full-Time) – Multan, Pakistan

City: Multan | Country: Pakistan

Machine Learning Engineer -

[07/2022 – 01/09/2025]

- Developed and launched an internal **NLP-powered chatbot** to serve as a technical knowledge base for 50+ engineers. The system **parsed** thousands of pages of technical documents (BOMs, spec sheets), **enabling efficient information retrieval** via natural language queries and reducing information search time by over **50%**.
- Engineered a **predictive quality** control system by analyzing manufacturing process data. Trained a **Random Forest classification model** on inspection logs (features included machine settings, material batches, operator IDs) to predict defect likelihood with **85% accuracy**, cutting material scrap by **12%** and reducing manual inspection overhead by **30%**.
- Built and deployed a **demand-forecasting model** (Prophet, SQL) using **historical Bill of Materials (BOM)** usage to optimize inventory levels. This improved **forecast** accuracy by **18%** and reduced component stockouts by **25%**, directly **accelerating prototyping cycles**.

🏠 UrbanMop - Remote – Toronto, Canada

City: Toronto | Country: Canada

Data Scientist/ Analyst -

[10/2023 – 12/2024]

- **Engineered and deployed a predictive pipeline** to model customer behavior, **forecasting churn** and **lifetime value (LTV)**. This system enabled proactive marketing interventions, improving **customer retention** by **10%**.
- **Developed API-driven ETL workflows** (Python, Launch27 API, Cron on VPS) to ingest booking and cleaner performance data, automatically **flagging late arrivals** and no-shows—cut **manual tracking** by **90%** and **reduced no-shows** by **30%**.

- Created an **automated revenue recognition system** (Python, Google Sheets API, Power BI) that processed **time series booking data** into **financial reports**—halved month-end close time and improved **revenue accuracy by 15%**.
- Designed **comprehensive PowerBI dashboards** consolidating **bookings, revenue, and operational KPIs**—boosted scheduling **efficiency by 25%** and provided **real-time visibility** that empowered **data-driven decision-making**.

PROJECTS

Human Activity Recognition (HAR) from Sensor Data

- Engineered a complete **ML pipeline** to **classify human activities** (e.g., walking, sitting, standing) from the **time-series** sensor data in the **Pamap2 dataset**.
- Evaluated six different **classification algorithms**, with **Random Forest** selected as the production model for achieving a **98% accuracy rate** on the held-out test set.
- **Technologies:** Python, Scikit-learn, Pandas, NumPy, Matplotlib

AI-Powered Q&A System for Scientific Research

- Developed a **Retrieval-Augmented Generation** (RAG) system to answer **natural language** questions based on a corpus of **100+** ArXiv research papers.
- Implemented a **pipeline** that ingests PDFs, chunks text, generates **embeddings** (using a Hugging Face model), stores them in a **vector database** (FAISS), and uses a **Large Language Model** to synthesize answers from retrieved context.
- **Technologies:** Python, LangChain, Hugging Face Transformers, FAISS, Streamlit

PROFESSIONAL DEVELOPMENT & CERTIFICATIONS

1. Deep Learning Specialization

DeepLearning.AI / Coursera

2. TensorFlow Developer Professional Certificate

TensorFlow / Google

3. Natural Language Processing Specialization

DeepLearning.AI / Coursera

4. Google Data Analytics Professional Certificate

Google / Coursera

SKILLS

Langchain / Hugging Face Transformers / Vector Database

Programming Language & Tools

Python / R / SQL / Power BI / Tableau

Framework & Libraries

Python (TensorFlow, Keras, PyTorch, Pandas, NumPy, Scikit-learn, GDAL, OpenCV, Seaborn, NLTK) / Python DSci/ML Libraries (Numpy, Pandas, Sklearn, Scipy) / Python visualization libraries (Matplotlib & Seaborn)

Data Science & AI Techniques

Supervised Machine Learning: Regression and Classification / Unsupervised Machine Learning / Deep Learning, Reinforcement Learning, Regression, SVM, RF, Naive Bayes, Clustering, LASSO / Time series Forecasting

Databases

MySQL, SQL, SQLite, PostgreSQL / MongoDBCompass

EDUCATION AND TRAINING

Bachelors in Computer Sciences (BSCS) -

Bahauddin Zakariya University Multan [09/2018 – 06/2022]

City: Multan | Country: Pakistan | Final grade: 3.52 / 4.00

HOBBIES AND INTERESTS

Competitive team-based strategy and the psychology of performance and systems thinking.
