

Aijaz Ali Shah Syed

Data & ML Engineer – Databricks & Fabric Associate

✉ lakyari.aijaz@gmail.com · 📞 +49 157 3302 1370 · 📍 Düsseldorf, Germany

DOB: 19.10.1999 · Nationality: Pakistani · Visa: Job-Seeking

[LinkedIn](#) · [GitHub](#) · [Certifications](#)



PROFILE SUMMARY

3+ years of experience in Data Engineering and Machine Learning. Production experience building forecasting system on Azure Databricks, owning the full ELT and ML lifecycle from data ingestion to training models. Experienced in building scalable data pipelines and end-to-end deployable workflows.

TECHNICAL SKILLS

Languages	Python, SQL, R, C++
Data Engineering	ETL/ELT Pipelines, Delta Lake, Spark, Airflow, Databricks, Snowflake, Fabric
Cloud & Ops	Azure (ADLS Gen2, ADF, Synapse), Terraform, DABs, CI/CD, Docker, Git
Databricks	Auto Loader, Unity Catalog, DLT/SDP, Databricks SQL, Jobs, Genie
ML & MLOps	PyTorch, TensorFlow, Optuna, scikit-learn, MLflow, Spark Streaming, Model Serving
AI Stack	ChatGPT, Claude (code, cowork), Copilot, n8n, AI Studio
Currently exploring	GCP/ BigQuery, SAP BDC, LLMops, RAG, Agentic AI

WORK EXPERIENCE

Working Student – Data Engineering & MLOps

10/2023 – 06/2025

Stadtwerke Krefeld AG

- Developed **Azure Databricks** forecasting platform serving district heating and energy systems.
- Engineered **PySpark & Delta Lake** ELT pipelines, cutting data processing time by **30%** to benchmark.
- Orchestrated **ADF** and **Databricks Asset Bundles** for repeatable, production-grade ELT deployments.
- Implemented **MLflow** tracking and model versioning across multi-environment **CI/CD** pipelines.
- Deployed **Spark Structured Streaming** endpoints, reducing time-to-production from weeks to days.
- Governed **Unity Catalog** metadata standards, ensuring data quality across the full platform.
- Automated **Power BI** dashboards for real-time energy forecast visibility across business units.
- Presented Data and ML architecture decisions to technical and business stakeholders.

Master Thesis – Hybrid ML-MILP Unit Commitment Optimization

01/2025 – 06/2025

Stadtwerke Krefeld AG & TU Dortmund

- Designed hybrid **ML-MILP** framework for unit commitment, achieving **98%** solution optimality.
- Reduced **Gurobi** runtime by **57.1%** by replacing constraints with ML-predicted warm-start bounds.
- Automated **MLflow** pipeline with Optuna hyperparameter tuning and k-fold cross-validation.
- Enforced production-readiness via data validation and feasibility assertions on model outputs.

Graduate Student Assistant – Research

11/2022 – 06/2023

TU Dortmund (ie3 Institute)

- Simulated 5G network communication requirements for solar power plant using **NS-3** and **MATLAB**.
- Analysed simulation outputs to evaluate latency, throughput, and network performance metrics.
- Documented findings and supported research publications at TU Dortmund's ie3 Institute.

Data Engineer Trainee

01/2021 – 06/2022

Torino Technologies, Pakistan

- Managed **SQL** databases including schema design, query optimisation, and ensured data quality.
- Migrated on-premises database to **Azure** cloud, supporting scalable and cost-efficient data storage.
- Built and maintained **ETL/ELT** pipelines automating extraction, transformation, and loading.
- Developed automated CI/CD pipelines and data integration scripts using **Python** and **SQL**.
- Delivered data engineering solutions for international and local clients across multiple projects.

PROJECTS

Databricks End-to-End Forecasting Pipeline – [GitHub](#)

- Databricks-based energy forecasting pipeline with **DLT**, **MLflow**, and **Unity Catalog**; deployed via Databricks Asset Bundles (DABs) with fully automated dev-to-prod CI/CD.

Hybrid ML-MILP Optimization Framework – [GitHub](#)

- End-to-end MLOps framework combining LSTM & LightGBM with MILP solvers — Optuna tuning & **MLflow versioning**, k-fold time-series validation.

Virtual Power Plant Simulation – MSc Project (12 ECTS) – [GitHub](#)

- Simulated renewable generation & flexible load dispatch using **VPPLib**; implemented operational strategies for cost-effective consumer energy management.

Energy Trading Platform – MSc Project (12 ECTS)

- Agent-based smart factory energy bidding using **Pyomo**; machines bid for time slots by operational priority & real-time energy prices.

CERTIFICATIONS

- Databricks Data Engineer Associate
- Databricks ML Associate
- Azure Databricks Platform Architect
- Microsoft Fabric DP-700
- Snowflake Essentials (Badge 1 & 2)
- dbt Fundamentals

EDUCATION

M.Sc. Automation & Robotics 10/2021 – 09/2025

Technical University Dortmund

Diploma in Technology Management 04/2023 – 10/2024

Technical University Dortmund

B.E. Mechatronics 10/2016 – 09/2020

Mehran University of Eng. & Tech.

LANGUAGES

English – C1

German – B1

Urdu – Native

Sindhi – Mother tongue

ADDITIONAL INFORMATION

Hobbies: Sports, Traveling, Cooking

Driving License: German Class B