

03413215170
Karachi, Pakistan
waqargul6@gmail.com

Muhammad Waqar Gul

Azure Certified Data Engineer

github.com/waqarg2001
linkedin.com/in/waqar-gul

EDUCATION

Bachelors in Software Engineering, NED University of Engineering & Technology 2020-2024

SKILLS

Languages and Databases	Python, Bash Scripting, MySQL, SQL Server, MongoDB
Tools and Cloud Services	PySpark, Airflow, Databricks, Data Factory, Data Lake, Synapse, AWS S3, Glue, Redshift, Sagemaker, Athena, Lambda, EMR
Data Science Tools	Tableau, Power BI, Numpy, Pandas, Scikit-learn, Tensorflow, Docker, Kubernetes

EXPERIENCE

Freelance Data Engineer | Level 1 March 2022 — Present
Fiverr Remote

- Successfully completed 20+ data engineering and data science projects, delivering high-quality solutions to clients. Applied advanced data engineering techniques to independently execute a variety of projects.
- Maintained a consistent 5-star rating through exceptional project management, technical expertise, and effective communication.

Machine Learning Intern February 2021 — June 2021
Goddard Discovery Karachi, Pakistan

- Collaborated with an AI team to develop an efficient cloud-based clustering solution that successfully clustered over 600 research articles.
- Presented multiple clustering approaches and utilized Amazon Sagemaker to create clustering models.
- Played a key role in the successful integration of the created models into the Minimum Viable Product (MVP).

PROJECTS

Earthquake ETL Pipeline

- Leveraged Azure services, including Data Factory and Databricks, to establish an end-to-end data engineering pipeline for real-time earthquake data collection and processing.
- Employed PySpark within Azure Databricks for efficient data transformations and analysis, ensuring timely insights into seismic activities.
- Created an interactive Tableau dashboard to visualize and communicate earthquake data in real-time.

Covid-19 Data Pipeline

- Utilized Azure services, including Data Factory, Dataflow, and Databricks for efficient ETL operations and data analysis.
- Ensured data security through Azure Key Vault and Service Principals for credential management.
- Integrated Covid data with population data for comprehensive COVID-19 analysis, utilizing Power BI for interactive visualizations of pandemic trends across Europe.

Youtube Data Pipeline (AWS)

- Utilized AWS services, including S3, Glue, and QuickSight, for efficient ETL operations and data visualization.
- Ensured data security through IAM roles and policies for access management.
- Processed YouTube video statistics data from Kaggle, transforming and storing it in Parquet format for comprehensive analysis and visualization using AWS QuickSight.

ACHIEVEMENTS

- Achieved first rank among 30+ teams in the departmental FYDP poster design competition at NEDUET in December 2023.
- Team secured 1st place out of 20+ teams in the Koderz Kombat competition at Iconics 22, held at NEDUET.
- Finished as runner-up among 10+ teams in a data science competition hosted by the AI Club on Kaggle in Nov 22.
- Ranked 3rd out of 10+ teams in a Python programming competition organized by SAPNED in Feb 2021.

CERTIFICATIONS

- Micorsoft Azure Data Engineer Associate DP-203
- Microsoft Azure Fundamentals AI-900
- Associate Data Analyst Certification by Datacamp
- Data Science Career Track by DataCamp
- Databricks Accredited Lakehouse Platform Fundamentals
- Microsoft Technology Associate (MTA 98- 31) Intro to Python Programming