

Syed Azfar Ali

Karachi, Pakistan | +92-340-2007760 | azfari021.aa@gmail.com
linkedin.com/in/azfarali-saa0021 | github.com/azfarali16

SUMMARY

Data Scientist with experience in building ML models, crafting data-driven insights, and deploying data solutions. Skilled in Python, SQL, ETL, and machine learning with a strong foundation in data engineering.

EDUCATION

NED University of Engineering and Technology
Bachelor of Science in Computer Science

Karachi, Pakistan
Nov. 2021 – June 2025

EXPERIENCE

Firstplant

Remote

Data Engineer Intern & Data Engineer

Nov 2024 – Present

- Worked with cross-functional teams to design methods for acquiring and validating eco-friendly product data from both structured and unstructured sources.
- Designed and implemented scalable data management systems, including hands-on ETL processes to extract, transform, and load data.
- Contributed to data quality assurance efforts and explored modern tools to enhance data systems, while supporting machine learning applications to improve data procurement.
- Continuing development of scalable data pipelines and management systems to support evolving business needs.
- Driving improvements in data quality and supporting advanced analytics and machine learning workflows.

StaunchTie Innovations

Remote

Software Developer

Oct 2024 – Dec 2024

- Developed analytical dashboards using Dash and Plotly to visualize key data insights.
- Collaborated on frontend development and interface improvements to enhance user experience.
- Implemented data visualization features supporting business decision-making and performance tracking.
- Worked with cross-functional teams to deliver scalable, user-friendly solutions.

PROJECTS

PhishTracer: Phishing Detection System

- Developed an email phishing detection system using ML and RESTful API with Python.
- Cleaned large-scale email datasets and improved model accuracy through feature engineering.
- Deployed as a Python microservice integrated with a Node.js app.

E-Store ETL Data Pipeline

- Automated ETL pipeline to extract data from Amazon RDS, clean and transform it into a star schema.
- Loaded transformed data into MySQL for efficient querying and analytics.
- Designed a data warehouse supporting business intelligence tasks.
- Integrated MySQL with Power BI for dynamic dashboards on inventory and sales.

Leads Sales Agent

- Built a lead generation bot using Google ADK and Gemini 2.0 (LLM) for interaction.
- Integrated with an HTTP form that triggers the bot to ask three questions: age, country, and interest.
- Stored user responses in a CSV file for further analysis or contact.
- Implemented timeout logic to send up to 3 reminders before automatically ending the session.

CERTIFICATIONS

Data Analyst with Python – DataCamp
Machine Learning Specialization – DeepLearning.AI (ongoing)

TECHNICAL SKILLS

Languages: Python, SQL

Frameworks: Flask, TensorFlow, PyTorch, Dash

Tools: Git, REST APIs, Power BI

Data Science: Pandas, Scikit-learn, Machine Learning, Deep Learning, EDA, Data Warehousing Concepts