

Tahir Malik

tahir.malik296@gmail.com

+91-997-158-7221

[LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

Amity University

Bachelor of Technology – Information Technology; GPA: 8.45

Gurugram, India

June 2017 – August 2021

Relevant Coursework: Python, SQL, Big Data & Fundamentals, Economics (Minor)

SKILLS SUMMARY

- Languages:** Python, SQL, C, C++
- Data Engineering & Processing:** Data Cleaning, Data Transformation, ETL
- Databases:** MySQL, MongoDB, PostgreSQL
- Data Tools & Visualization:** Power BI, Tableau, Microsoft Excel
- Big Data Tools:** Hadoop, Hive, Kafka, Spark, Databricks, Airflow
- Cloud:** AWS, Microsoft Azure, and Google Cloud Platform
- Development Tools:** Jupyter Notebook, Visual Studio Code, Git
- Soft Skills:** Problem Solving, Strong Communication, Team Collab.

WORK EXPERIENCE

DATA ENGINEER INTERN | MACTORES | MUMBAI

April 2025 – Nov 2025

- Gained hands-on experience in data engineering workflows on the Databricks Lakehouse Platform.
- Monitored, debugged, and optimized ETL pipelines built with Apache Spark and Delta Lake to ensure reliability & data quality.
- Worked on implementing real-time streaming pipelines using Kafka, Delta Lake, and Databricks Auto Loader.
- Achieved AWS Cloud Practitioner and Databricks Certified Data Engineer (Associate & Professional) credentials during this role.

PROJECTS

Fintech Data Migration | [LINK](#)

- Developed scalable pipelines for migrating and transforming data of a fintech company from their traditional SQL database to Azure Data Lake, ensuring data consistency and minimal downtime.
- Implemented a medallion architecture within ADLS Gen2 using PySpark & Delta Lake to effectively manage and process data, improving data accessibility and reliability.

E-Commerce Data Streaming Analytics | [LINK](#)

- Developed a robust data pipeline leveraging Confluent Kafka to collect e-commerce clickstream and user activity data, enabling real-time analytics capabilities.
- Combined Azure Data Lake Storage Gen2 (ADLS Gen2) with Databricks to efficiently store, cleanse, and convert raw data into formats suitable for analysis, facilitating actionable insights.

CERTIFICATES

Databricks Certified Data Engineer Professional

[LINK](#)

Databricks Certified Data Engineer Associate

[LINK](#)

AWS Certified Cloud Practitioner

[LINK](#)

Google Data Analytics Specialization

[LINK](#)