

Anmol Garg

+91 8650660652 | anmolgarg848@gmail.com

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

PROFILE

Data Engineer with **3+** years of experience building large-scale data pipelines, ETL processes, and data warehouse solutions. Utilized technologies like **Python, SQL, Spark, Docker, Databricks, AWS and Kafka** to develop multiterabyte scalable big data solutions for Fortune 100 Finance and Telecom companies.

TECHNICAL SKILLS

Programming Languages: Python, SQL, Scala, Java

Big Data & Distributed Systems: Apache Spark (PySpark), Hadoop, Hive, YARN, Delta Lake, Apache Hudi,

Streaming Technologies: Kafka, Spark Structured Streaming, Apache Flink

Cloud Platforms: Azure (ADF, ADLS, Synapse, Azure Databricks), AWS (S3, Glue, EMR)

Data Engineering & ETL: Data Modelling, ETL/ELT Pipelines, Delta Lake, DQ

Orchestration & DevOps: Airflow, Azure DevOps, Terraform, CI/CD Pipelines, Kubernetes, Docker

Databases & Caching: Elasticsearch, Redis, Oracle

EXPERIENCE

EXL Analytics[UK]

March 2025 - Present

Data Engineer

[Remote]Pune, Maharashtra

- Delivered a project to migrate legacy on-premise processes to the cloud using Big Data technologies (Spark), reducing processing time by 20%.
- Developed a **microservice** deployed on Kubernetes, integrating **FastAPI** to push messages into a **Kafka broker** for real-time tracking of processing status (success or failure).
- Partnered with **business stakeholders** during SIT/UAT phases, ensuring seamless production releases and eliminating high-severity post-deployment defects.
- Designed and **Built ad-hoc file integration** pipelines enabling business teams to seamlessly ingest flat files into the existing data lake architecture, reducing manual onboarding effort and turnaround time.
- Integrated and processed billing data from multiple sources including **AWS, Azure, GCP, and Oracle**, handling various file formats like JSON, CSV, Excel and **Parquet**.

Reliance Jio

December 2023 - March 2025

Data Engineer

Mumbai, Maharashtra

- Engineered a **Hive-driven alert system** to identify JioFiber outages, reducing downtime resolution time by **40%**, analyzing **100K+** daily complaints and device disconnections within **60 minutes**.
- Optimized overall process performance through **Spark performance tuning**, improving job run times by 20% and efficiently managing a **15Terabyte(TB)** dataset containing approximately **12 billion** records.
- Developed an innovative solution that improved validation for **50,000+ JioFiber WiFi** connections per month based on signal strength, streamlining onboarding processes.
- Designed and implemented advanced **scheduling capabilities** using Airflow for data pipeline orchestration, **reducing manual intervention time by 80%** and streamlining workflow efficiency.

Hexaware Technologies

March 2023 - August 2023

Data Analyst, Intern

Bhopal, India

- Developed a Spark-based ETL job to automate loading data from server to SQL, replacing the earlier manual Excel-based process and enabling faster analysis across key engagement KPIs.

CERTIFICATES / ACHIEVEMENTS

- Azure Certified **Fabric** Data Engineer Associate [\[Link\]](#)
- Databricks** Certified Data Engineer **Professional**
- Solved **700+**Problems on platforms like **LeetCode** and **GFG**

EDUCATION

Institute of Technology, Bhopal

Bachelor of Technology in Electronics Engineering

Bhopal, India

July 2019 - June 2023