Varchasv Hoon

Mumbai, India
varchasvhoon02@gmail.com — +91-9582668467
linkedin.com/in/varchasvh — github.com/varchasvh

PROFESSIONAL SUMMARY

Data Engineer with proven ability to build scalable, cloud-native data systems. Skilled in designing end-to-end batch and streaming pipelines using PySpark, AWS (S3, Glue, RDS, Redshift), Kafka, and Airflow. Currently working with Citi Bank via LTIMindtree. Committed to delivering analytics-ready solutions in product-driven and FinTech environments.

EXPERIENCE

Data Engineer – LTIMindtree (Client: Citi Bank)   Apr 2024 – Present
Mumbai, India

- Developing scalable PySpark-based ingestion and transformation pipelines on AWS.
- Automating ETL workflows and modeling using SQL, Glue, and Airflow.
- Applied Spark partitioning, schema evolution, and performance tuning techniques.

PROJECTS

Batch Data Pipeline on AWS   May 2025 – Jun 2025
Personal Project

- Engineered ETL pipeline using **S3, Glue (PySpark)**, and **RDS (MySQL)**.
- Loaded transformed data into **Redshift** via secondary Glue job.
- Managed orchestration using **Apache Airflow** with retries and alerts.

Real-Time Streaming Pipeline   Jul 2025 – Aug 2025
Personal Project

- Simulated order stream via **Kafka**; processed with **Spark Structured Streaming**.
- Modeled and stored enriched data in **PostgreSQL** using dimensional schema.
- Containerized pipeline using **Docker Compose** for deployment.

TECHNICAL SKILLS

- **Languages:** Python, SQL, Bash
- **Big Data Frameworks:** PySpark, Apache Spark
- **Cloud & Storage:** AWS (S3, Glue, RDS, Redshift), IAM
- **Streaming:** Kafka, Spark Structured Streaming
- **Databases:** MySQL, PostgreSQL
- **Orchestration:** Apache Airflow
- **Tools:** Docker, Git, Jupyter, Faker

EDUCATION

B.Tech in Computer Science and Engineering   2020 – 2024
Galgotias University, Greater Noida   *CGPA: 8.05 / 10*

CERTIFICATIONS

- Databricks Certified Associate Developer for Apache Spark (Python) – Jun 2024
- AWS Certified Data Engineer – Associate (Planned)