

Summary

Databricks Data Engineer focused on industrial lakehouse and ML features. Builds reliable data products with Databricks, Delta Lake, DLT, PySpark, Unity Catalog, connects engineering decisions to measurable operations, and communicates clearly with cross-functional teams.

Work Experience

Databricks Data Engineer

ForgeSight Manufacturing

1/2022 - Present

- Designed Databricks and Delta Lake pipelines for industrial lakehouse and ML features at ForgeSight Manufacturing; while partnering with product, platform, and finance leads, processing 4.6 TB of daily sensor data, and raised on-time delivery from 92% to 99.6%.
- Reworked DLT models and partitioning for the Predictive Maintenance Lakehouse warehouse; coordinated a 9-person engineering group, cutting median query time from 24 seconds to 4.1 seconds.
- Instrumented schema, freshness, and lineage checks across 38 production tables for industrial lakehouse and ML features; worked with 14 analysts, reducing data incidents by 63%.
- Automated backfills, alerts, and runbooks around PySpark jobs in New York, NY; supported 12 markets and saved 86 on-call hours per quarter.

Education

Master of Science

Columbia University

9/2015 - 6/2019

Skills

Core stack

Databricks, Delta Lake, DLT,
PySpark, Unity Catalog

Delivery

Data modeling, Data quality,
CI/CD, Observability,
Documentation

Collaboration

Requirements workshops, Incident
response, Mentoring, Agile
delivery

- Designed Databricks and Delta Lake pipelines for industrial lakehouse and ML features at ForgeSight Manufacturing; with analysts, SREs, and domain owners across the delivery squad, processing 4.6 TB of daily sensor data, and raised on-time delivery from 92% to 99.6%.
- Reworked DLT models and partitioning for the Predictive Maintenance Lakehouse warehouse; coordinated a 6-person engineering group, cutting median query time from 24 seconds to 4.1 seconds.
- Instrumented schema, freshness, and lineage checks across 38 production tables for industrial lakehouse and ML features; worked with 5 application teams, reducing data incidents by 63%.
- Automated backfills, alerts, and runbooks around PySpark jobs in New York, NY; supported 4 business units and saved 86 on-call hours per quarter.

English

🌐 Languages

Bengali

Databricks Certified Data Engineer Professional

Professional Training Board 2024

🏆 Certifications

Data Reliability Workshop

Data Engineering Guild 2023

<> Projects

Predictive Maintenance Lakehouse

Built a governed data product for industrial lakehouse and ML features, with documented ownership, tested transformations, and an analyst-ready serving layer.