DevOpsSchool

Course curriculum · November 2026

LEVEL 04 OF 04 · ADVANCED

Databricks Advanced Training

Design, tune and govern an enterprise Databricks lakehouse for performance, cost, security, generative AI and platform migration.

LIVE ONLINE Max 8-10 hours	**CORPORATE (DAY-WISE)** 5 days · Days 1-5	**SELF-PACED VIDEO** 1-year LMS access
BATCH SIZE Max 10 learners	**NEXT BATCHES** 1st of every month · weekends: first Saturday	**CERTIFICATE** Verifiable certificate of completion

Foundation	Essential	Intermediate	Advanced
1 day · Included	2 days · Included	3 days · Included	5 days · This curriculum

INSIDE THIS CURRICULUM

What you will achieve	Outcomes and level focus
Who should attend	Audience and prerequisites
Agenda	49 topics · Foundation to Advanced
Hands-on labs	4 labs
Formats and fees	Live, 1:1, self-paced, corporate
Trainer and reviews	Rajesh Kumar
Enrolment	Batches, payment, contacts

WHY DEVOPSSCHOOL

- Rated 4.6 on Google
- 1,500+ verifiable certificates issued in the public registry
- Trainer Rajesh Kumar: 20 years in industry
- Small batches: max 10 learners, every session recorded

Course page: www.devopsschool.com/courses/databricks/databricks-advance-training.html
Talk to us: WhatsApp +91 99057 40781 · +1 (469) 756-6329 · contact@DevOpsSchool.com · reply within 1 hour

OUTCOMES

What you will achieve

By the end of the Databricks course path, you will be able to:

1

Engineer reliable lakehouse pipelines

Build batch and streaming pipelines on Delta Lake with the medallion pattern and data quality expectations.

2

Govern data with Unity

Secure, share and audit data and models using Unity Catalog permissions, lineage and system tables.

3

Operate ML in production

Track, register, serve and monitor models with MLflow and Databricks Model Serving.

WHAT THE ADVANCED LEVEL ADDS

Performance, cost, security and enterprise lakehouse

Hands-on: Tune and govern a production lakehouse. Profile a slow job in the Spark UI, fix it with clustering and AQE, apply row filters and column masks, and report cost from system tables.

WHO SHOULD ATTEND

- Experienced Databricks practitioners responsible for quality, security and scale.
- Senior engineers, architects and leads who design and troubleshoot Databricks solutions.
- Teams making Databricks an organisation-wide standard.

BEFORE YOU START

- The Intermediate level, or equivalent hands-on experience.
- You can already: Automate Databricks delivery with Lakeflow Declarative Pipelines, Asset Bundles, streaming and fine-grained Unity Catalog security.
- Basic computer skills and a laptop with a stable internet connection for the live sessions and labs.
- This curriculum includes every earlier level, so it is complete on its own.

AGENDA · FOUNDATION TO ADVANCED

Databricks Advanced Training agenda

49 topics. Each level builds on the one before it, so this curriculum lists every earlier level in full. Days are for corporate (day-wise) delivery; live online batches cover the same agenda in max 8-10 hours.

DAY 1 — Foundation: Platform, Spark basics, Delta Lake and MLflow

Included from Foundation

1. Introduction to Databricks: overview of the platform and its components, workspace, clusters and notebooks architecture
2. Setting up Databricks: creating, configuring and managing clusters; runtime versions, autoscaling and auto-termination
3. Core features: notebooks, version control, jobs and collaboration across data engineering and data science
4. Apache Spark basics: Spark core components, driver and executors, RDDs, DataFrames and Spark SQL
5. Spark DataFrame API: how DataFrames are used for large-scale processing, data cleaning and manipulation
6. Spark SQL: performing queries on large datasets, temporary views and mixing SQL with Python cells
7. Building data pipelines: techniques for transforming and storing data using Spark
8. Working with Delta Lake: reliable and scalable data storage, reading and writing Delta tables
9. Machine learning with MLlib: data preparation, feature extraction, model training and evaluation
10. MLflow: managing machine learning experiments, model tracking and hyperparameter tuning
11. Collaboration in Databricks: sharing notebooks, Git integration for version control, team-based workflow
12. Data visualization: built-in notebook visualizations and building interactive dashboards to present insights
13. Hands-on: set up a workspace, run Spark queries, create a Delta Lake pipeline and build an MLlib model and dashboard

DAY 2 — Essential: Unity Catalog, ingestion, medallion design and SQL warehouses

Included from Essential

14. Unity Catalog basics: metastore, catalogs, schemas, tables, volumes, managed versus external tables and GRANT statements
15. Data ingestion: COPY INTO, Auto Loader (cloudFiles) with schema inference and evolution, reading CSV, JSON and Parquet
16. Medallion architecture: bronze, silver and gold layers, naming conventions and data contracts between layers
17. Delta Lake operations: MERGE INTO upserts, UPDATE and DELETE, time travel, DESCRIBE HISTORY and RESTORE
18. Delta table maintenance: OPTIMIZE, liquid clustering, VACUUM retention and table properties
19. Databricks SQL: SQL warehouses (serverless, pro, classic), SQL editor, saved queries and query history
20. Lakeflow Jobs: multi-task workflows, task dependencies, parameters, schedules, retries and notifications
21. Notebook utilities: dbutils for files, widgets and secrets, %run, magic commands and parameterised notebooks
22. Compute choices: all-purpose versus job compute, serverless compute, cluster policies and Photon
23. Git folders: cloning repos, branches, commits and pull requests from the workspace
24. AI/BI dashboards and Genie spaces: datasets, visual widgets, filters, sharing and scheduled refresh
25. Working with files: workspace files, Unity Catalog volumes and cloud object storage paths

DAY 3 **Intermediate: Declarative pipelines, CI/CD, streaming and governance** Included from Intermediate

26. Lakeflow Declarative Pipelines (formerly Delta Live Tables): streaming tables, materialized views, expectations for data quality
27. Structured Streaming: readStream and writeStream, checkpoints, triggers, watermarks and stateful aggregations
28. Change data capture: Change Data Feed, APPLY CHANGES (AUTO CDC) and slowly changing dimensions type 1 and 2
29. Databricks Asset Bundles: databricks.yml, targets for dev, staging and prod, bundle validate, deploy and run
30. Databricks CLI and REST API: workspace automation, service principals and OAuth machine-to-machine tokens
31. CI/CD: GitHub Actions or Azure DevOps pipelines that test notebooks, run pytest and deploy bundles
32. Unity Catalog security: row filters, column masks, dynamic views, ownership, lineage and system tables for audit
33. Lakehouse Federation and Delta Sharing: querying external databases and sharing data across organisations
34. MLflow in depth: model registry in Unity Catalog, aliases, model signatures and Feature Engineering tables
35. Model Serving: deploying real-time endpoints, batch inference with ai_query and monitoring request payloads
36. Databricks Connect and VS Code extension: local development against remote compute
37. Troubleshooting: Spark UI stages and tasks, data skew, out-of-memory errors, failed jobs and driver logs

DAYS 4-5 **Advanced: Performance, cost, security and enterprise lakehouse** **New at this level**

38. Spark performance tuning: adaptive query execution, partition sizing, broadcast joins, caching and skew hints
39. Photon and storage optimisation: predictive optimization, deletion vectors, file sizing and clustering key choice
40. Cost management: system.billing tables, budgets, tagging, cluster policies, serverless versus classic cost trade-offs
41. Workspace architecture: account console, multiple workspaces, metastore per region, environment isolation
42. Network security: private connectivity, customer-managed VPC or VNet, IP access lists and customer-managed keys
43. Identity at scale: SCIM provisioning, account groups, service principals and least-privilege entitlement design
44. Data governance programme: Unity Catalog tags, attribute-based policies, lineage reviews and audit log monitoring
45. Reliability: job repair runs, idempotent pipelines, disaster recovery with Deep Clone and cross-region replication
46. Generative AI on Databricks: Vector Search, Foundation Model APIs, Mosaic AI Agent Framework and agent evaluation
47. Lakehouse Monitoring: data and model drift profiles, inference tables and quality alerts
48. Migration: moving Hive metastore tables to Unity Catalog with UCX, and legacy Hadoop or warehouse workloads to Databricks
49. Capstone: deliver a governed, cost-tracked lakehouse with streaming ingestion, ML model serving and CI/CD

PRACTICE

Hands-on labs

You build these yourself during the course, guided by the trainer.

FOUNDATION

From raw files to a dashboard

Create a workspace and cluster, load a dataset with Spark, write it to Delta Lake, train and log a simple MLlib model in MLflow, and publish a dashboard.

ESSENTIAL

Build a governed medallion pipeline

Ingest raw files with Auto Loader into bronze, clean them into silver with MERGE, aggregate gold tables in Unity Catalog and schedule it as a job.

INTERMEDIATE

Ship a pipeline through CI/CD

Define a streaming declarative pipeline with expectations in an Asset Bundle, test it in GitHub Actions and promote it from dev to prod.

ADVANCED

Tune and govern a production lakehouse

Profile a slow job in the Spark UI, fix it with clustering and AQE, apply row filters and column masks, and report cost from system tables.

HOW YOU CAN LEARN

Formats and fees

The same agenda in four formats. Fees are in Indian rupees as of November 2026; the course page shows them in your currency.

FORMAT	DURATION	WHAT YOU GET	FEE
Live & interactive (online)	Max 8-10 hours	Instructor-led public batch of max 10 learners with demos, labs, Q&A and assignment review. Weekday batches start on the 1st of every month, weekend batches on the first Saturday of every month.	**INR 34,999**
1:1 live training (online)	Max 8-10 hours, one-to-one	The same agenda, taught one-to-one by the trainer. Starts immediately, without waiting for the next batch; sessions planned around your working hours, at your pace.	**INR 99,999**
Self-paced video	1-year LMS access	All four Databricks levels, plus 20+ more courses and 50+ tools, at your own pace.	**INR 833/month** billed yearly at INR 9,996
Corporate (day-wise)	5 days for this level	A private batch for your team, on demand. Foundation 1 day, Essential 2, Intermediate 3, Advanced 5 (cumulative).	**Quote on request**

ASSESSMENT AND CERTIFICATE

Your progress is checked through hands-on labs and assignments at every level. On completion you receive a DevOpsSchool certificate of completion with a unique credential ID that anyone can verify at devopsschool.com/certificates.

INCLUDED WITH EVERY LIVE BATCH

- Every live session is recorded and stays in your LMS for a year.
- Can't attend your batch? Move to next month's batch.
- Free demo class before you pay.
- Free 10-15 minute call with an advisor or the trainer to check the course fits you.

YOUR TRAINER

Learn from Rajesh Kumar

Principal Architect, trainer and mentor in DevOps, SRE, DevSecOps and platform engineering, with 20 years across 8+ software MNCs, including Staff DevOps/SRE Engineer for a leading Japanese mobile-payment platform and Principal DevOps Architect and Manager in Europe.

Profile: www.rajeshkumar.xyz · Videos: youtube.com/@TheDevOpsSchool

20 years in industry	**200+** corporate trainings
10,000+ engineers mentored	**70+** organisations coached

REVIEWS

What learners say

Rated 4.6 on Google · more reviews at devopsschool.com/testimonials and in our video testimonials.

> “Rajesh is very good trainer. Rajesh was able to resolve our queries and question effectively. We really liked the hands-on examples covered during this training program.”
>
> Indrayani, India

> “The training was very useful and interactive. Rajesh helped develop the confidence of all.”
>
> Abhinav Gupta, Pune

START YOUR DATABRICKS TRAINING

Join the next batch

Weekday batches start on the 1st of every month and weekend batches on the first Saturday of every month, with max 10 learners each. 1:1 training can start immediately; corporate batches run on demand. We reply within 1 hour.

Pay by UPI, credit and debit cards, net banking or PayPal. Before you pay, read our refund policy. All levels of this course: www.devopsschool.com/courses/databricks/

ENROL OR RESERVE A SEAT
Databricks Advanced Training

WHATSAPP · INDIA
+91 99057 40781

WHATSAPP · USA
+1 (469) 756-6329

EMAIL
contact@DevOpsSchool.com